Information Fusion for Urban Road Extraction From VHR Optical Satellite Images

Zelang Miao, *Student Member, IEEE*, Wenzhong Shi, Alim Samat, *Student Member, IEEE*, Gianni Lisini, and Paolo Gamba, *Fellow, IEEE*

***Abstract*—This paper presents a novel method exploiting fusion at the information level for urban road extraction from very high resolution (VHR) optical satellite images. Given a satellite image, we explore spectral and shape features computed at the pixel level, and use them to select road segments using two different methods (i.e., expectation maximization clustering and linearness filtering). A road centerline extraction method, which is relying on the outlier robust regression, is subsequently applied to extract accurate centerlines from road segments. After that, three different sets of information fusion rules are applied to jointly exploit results from these methods, which offer ways to address their own limitations. Two VHR optical satellite images are used to validate the proposed method. Quantitative results prove that information fusion following centerline extraction by multiple techniques is able to produce the best accuracy values for automatic urban road extraction from VHR optical satellite images.**

***Index Terms*—Centerline, expectation maximization (EM), information fusion, linearness filter, RANdom SAmple Consensus (RANSAC).**

I. Introduction

URBANIZATION happens in just a few years in East Asia, according to the World Bank [1], as shown by the mass movement of people to cities and the emergence of urban settlements. The region will have more decades of urban growth as economies shift from agriculture to manufacturing and services [1]. Within this framework, updated road layers in geographic information systems (GIS) are critical for many urbanization issues, such as urban expansion estimation, urban planning, and traffic/population movement monitoring [2]. Although this is not an easy task, especially in urban areas, with the advent of modern very high resolution (VHR) optical satellite images, computer-aided road network extraction from remotely sensed images provides a new opportunity to meet this challenge.

Manuscript received May 21, 2015; revised October 27, 2015; accepted November 04, 2015. Date of publication November 24, 2015; date of current version April 22, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 41201451 and Grant 40901214, in part by the Ministry of Science and Technology of China (Project no: 2012BAJ15B04 and 2012AA12A305), and in part by the National Administration of Surveying, Mapping, and Geoinformation of China (Ling Jun Ren Cai). *(Corresponding author: Wenzhong Shi.)*

Z. Miao and W. Shi are with the Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: lswzshi@polyu.edu.hk).

A. Samat is with the Department of Geographic Information Science, Nanjing University, Nanjing 210023, China.

G. Lisini and P. Gamba are with the Department of Industrial and Information Engineering, University of Pavia, Pavia 27100, Italy.

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/JSTARS.2015.2498663

To date, road extraction from VHR satellite images has received a lot of attention, and various methods have been proposed. A comprehensive review is available in [3], where these methods are grouped into five categories: 1) seed point-based; 2) active contours; 3) machine learning; 4) object-based; and 5) other. The first category starts from user-defined road seed points, followed by road segments delineation, which exploits spectral/spatial characters of roads to recognize road network elements. In this category, a piecewise parabola model [4] was presented to delineate the road centerline network. This method first applies the piecewise parabola model around the seed point, followed by the least square matching to solve the parameters of the precise parabola to be extracted. Similar methods are presented in [5] and [6]. To reduce the seed point number, kernel density estimation incorporating the geodesic method is presented in [7]. For a specified pixel, road footprint [8] and angular texture signature [9] are defined to measure the road probability from the shape prospective. In [10] and [11], particle filtering is utilized to trace a segment initiated by the seed point. The shortage of particle filtering is that it fails to process road branches. To overcome this limitation, particle filtering and extended Kalman filtering were integrated in [12] to process complicated road cases. The single-image mode was discussed in the foregoing road extraction methods. As an extension, Dal Poz *et al.* [13] presented a semiautomatic method for rural road extraction from stereoscopic aerial images. This method derives the road seed points to formulate the object-space road model, followed by optimization using dynamic programming.

The second category method relies on the user defined initial contour to implement so-called active contour model to extract roads from satellite images. Two fundamental papers reporting the application of active contour in road extraction were given in [14] and [15]. Previous studies show that, in a variety of road extraction tasks, a single snake suffers from many limitations, such as the failure in disconnected road networks and enclosed regions. To tackle this limitation, a family of quadratic snakes [16] has been proposed for road extraction. This method combines advances in oriented filtering, thresholding, Canny edge detection, and gradient vector flow (GVF) energy, which surpasses consistently outperforms a single snake. Similarly, a higher order active contour model [17] is designed to solve the road discontinuity issue caused by shadows and trees. The parameter tuning and computational cost of active contour are addressed by Li *et al.* [18]. The substitution of the traditional regularization term by a Gaussian kernel benefits fewer parameters and larger time step, which in turn improve the result

smoothness and computational efficiency. To improve road extraction accuracy, active contour model has also been studied to integrate with other methods, such as graph cut [19], the multiresolution analysis [20], and Newton snakes [21]. A comparison of different active contour models for road extraction was given by Nakaguro *et al.* [22].

The third category is based on the theory of machine learning. With the spurious of machine learning, this powerful tool in the field of computer science has naturally extended its application area into the road extraction. Support vector machine (SVM), a powerful tool in machine learning, has come into wide use in road extraction [23]–[26]. Meanwhile, artificial neural networks are applied to extract roads, which concentrated on evaluating different structures of neural networks along with different measuring units and descriptors [27]. A multistage strategy for automatically extracting roads from high-resolution multispectral satellite images based on salient features was introduced by Das *et al.* [28]. This method incorporates the salient features of roads using P-SVM and dominant singular measure (DSM). The path classifier was studied for automatic delineation of linear features from images based on the global optimization with geometric priors [29]–[31]. The junction-point processes [32] and a higher-order conditional random field model [33] were also exploited to extract road network from the image. Particularly, Mnih and Hinton [34], [35] addressed the application issue of machine learning in road extraction from aerial images.

The fourth category is relying on the object level information to extract road pixels. These algorithms are primarily pixel-based and sometimes accompanied by textural information extracted from rectangular regions, in which structural and conceptual information is not properly exploited. To overcome the shortcomings of the pixel-based methods and to reduce the semantic gap, object-based algorithms have been developed [36]–[38]. Object-based methods are known to achieve better results than pixel-based methods in processing VHR satellite images [39]–[41]. A considerable number of studies have compared object-based approaches with traditional pixel-based classification methods [42]–[46]. Many of these studies have found that object-based methods typically produce higher classification accuracies than pixel-based methods do. A new work proposed by Huang and Zhang is based on SVM and multifeature model at both pixel and object levels [47], [48]. Peng *et al.* updated the outdated road maps by incorporating generic and specific prior knowledge into a multiscale phase field model [49]. Zarinpanjeh *et al.* used object-based analysis for road map updates [50]. Additionally, Grote *et al.* developed a method for road network extraction using object-based analysis [51]. In this category, to improve road extraction accuracy, various road spatial features derived from objects have been presented and then integrated with spectral features. These spatial characters include shape features [23], [25], [28], [52], directional mathematical morphological [25], [53], directional filter [54], [55], etc. It is worth to note that the methods in third and fourth categories frequently produce segments. The delineation of accurate road centerline from the classified image was addressed in [56]. Although integrating shape features in road extraction results in a good performance, it is challenging to obtain a universal linear feature extraction method for all situations [54], and this paper will further study this issue.

There are also some other interesting road extraction methods. A wavelet-based approach for road extraction from VHR satellite images was presented by Zhang and Couloigner [57]. Some new researches based on SAR imagery and LIDAR systems also have recently been presented [54], [55], [58]. Doucette *et al.* proposed an automated road centerline extraction method that exploits spectral content from high-resolution multispectral images [59]. The method is based on antiparallel edge centerline extraction and self-organized road mapping. A fast linear feature detector for road extraction was introduced by Shao *et al.* [60]. This method only considers ridge line (or bright ribbon) extractions that are mostly roads in aerial and satellite images. An interesting method that combines road color feature with road GPS data to detect road centerline was given in [61].

In general, semiautomatic methods can produce satisfactory result and thus are much closer to the operational level than automatic methods. The reason is that automatic methods generally cannot achieve satisfactory results in complicated scenes (i.e., dense urban environment). In contrast, the human interaction is more robust to these cases. Despite this advantage, however, semiautomatic delineation of roads from satellite images suffers problems of cost, accuracy, and efficiency, resulting in labor-intensive processes. Automated tools to assist analysts in this process would be of tremendous benefit for the production of timely road data set.

Previous studies, mostly focusing on road extraction in rural/suburb areas, show that road network extraction in urban environment is a much more challenging task because 1) road network in urban environment is generally dense; 2) many nonroad pixels, such as car parking and building, are spectrally similar to roads and are occasionally adjacent to roads; and 3) roads are frequently partially occluded by shadows and trees, leading to discontinuities in the ultimately extracted road network.

This is the reason why this study focuses on an automatic approach for urban road extraction from VHR optical satellite images. The presented approach addresses two of the challenges identified above: 1) automatic process and 2) dense urban environment. We also examine both qualitatively and quantitatively with many tests the benefits and limitations of fusing multiple extractions via different fusion rule sets, as opposed to using a single extraction procedure. The contributions of this paper can be thus summarized into three points.

1) The design of a novel framework for road segments extraction from VHR optical satellite images by introducing a new linearness filter applied to enhancing linear features in multispectral satellite images.
2) The implementation of a regression-based method for accurate road centerline extraction from candidate road areas identified by the previous steps.
3) The characterization of multiple information fusion rule sets that are available to combine road extraction results from multiple sources and with multiple techniques.

This paper is organized as follows. Section II presents the basic principles of the proposed method, while experimental

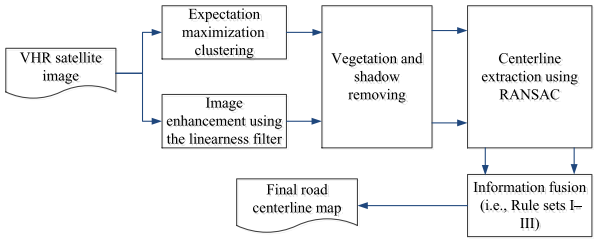


Fig. 1. Schematic flowchart of the proposed method.

results are reported in Section III, and discussions and conclusion are presented in Section IV.

II. Proposed Methodology

The objective of this study is to design a computationally efficient approach to extract accurate urban road networks from VHR optical satellite images. Fig. 1 graphically summarizes the main steps of the proposed method. The basic elements are a clustering method and a linearness filtering, implemented in a multiscale framework to improve the recognition of roads with different widths. The complete road network extraction and reconstruction is made by three steps.

1) Road segments are separately extracted using expectation maximization (EM) clustering and the linearness filter, respectively. This step provided candidate road segment sets.
2) Road centerlines are extracted from the candidate road segments by an outlier robust regression method.
3) Road centerline results produced by the different methods and/or the corresponding candidate road segments are combined using three information fusion rule sets to produce the final road centerline map.

Details of each step are described in the following sections.

A. EM Clustering

With the aid of the segmentation, the input image is subdivided into nonoverlapping regions. To each of these regions a set of parameters is assigned, e.g., the average gray and texture levels, its shape parameters, and its size. Since roads are homogeneous regions with respect to many of the above mentioned parameters, roads may be recognized by applying a suitable clustering method. In this paper, EM clustering, one of the most powerful segmentation algorithms, proposed by Dempster *et al.* [62], is used to segment the image. The EM algorithm was designed to estimate the parameters in statistical models. To cluster the VHR satellite image using EM, the multispectral satellite image is considered as a set of feature vectors $x_1, x_2, \ldots, x_N$ in a d-dimensional feature space driven from a Gaussian mixture

$$p\left(x;\mu_c,\Sigma_c\right)=\sum_{c=1}^{C} w_c p_c\left(x\right) \tag{1}$$

$$p_c\left(x\right)=\frac{1}{\left(2\pi\right)^{\frac{d}{2}}\left|\Sigma_c\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\left(x-\mu_c\right)^T\Sigma_c^{-1}\left(x-\mu_c\right)\right\} \tag{2}$$

(a) (b)

Fig. 2. (a) True color image of QuickBird image. The ground truth dataset, obtained by visual interpretation, is shown in yellow. (b) EM clustering result.

where C is the cluster number, p_c is the Gaussian distribution density with mean μ_c and covariance matrix Σ_c, and w_c is the weight of cluster C which satisfies $w_c \geq 0$ and $\sum_{(c=1)}^{C} w_c = 1$. The EM clustering technique is an iteration procedure and the mixture parameters in iteration are estimated by the following equations:

$$\mu_c^i=\frac{1}{w_c^{i-1}}\sum_{j=1}^{m_c^{i-1}}\left(x_{j.c}^{i-1}\right) \tag{3}$$

$$\Sigma_c^i=\frac{1}{w_c^{i-1}}\sum_{j=1}^{m_c^{i-1}}\left(x_{j.c}^{i-1}-\mu_c^i\right)\left(x_{j.c}^{i-1}-\mu_c^i\right)^T \tag{4}$$

$$w_c^i=\frac{m_c^{i-1}}{N} \tag{5}$$

where N is the number of pixel. The pixel is assigned to one of the C clusters according to the following equation:

$$x_j\in Q_c^i :|\ \Pr\left(c\mid x_j\right)=\max_l \Pr\left(l\mid x_j\right) \tag{6}$$

where $\Pr\left(c\mid x_j\right)=\dfrac{w_c^i\phi_c\left(x_j;\mu_c^i,\Sigma_c^i\right)}{\sum_{c=1}^{C}}\left(x_j;\mu_c^i,\Sigma_c^i\right)$.

Once the convergence is achieved, EM clustering is completed. In Fig. 2(a), we provide an example of a QuickBird image of the urban area, downloaded from a freely available image database assembled for road extraction purposes [63]. The spatial size of the test image is 400×400 pixels and the test image has four multispectral bands with the spatial resolution of 2 m/pixel. This image will be used throughout this paper to visualize intermediate results. Image segmentation results are thus presented in Fig. 2(b).

B. Linearness Filter

One different way to extract road candidate is to exploit one of their most important feature, i.e., the fact that roads are locally linear. To this aim, this paper presents a new linearness filter that combines the Hessian matrix, the local standard variance, and geometrical features.

1) Hessian Matrix-Based Filter: In the field of image processing, Hessian matrix [64] is an important tool that has various applications, because it is able to capture the local image structure [65]. Let us define $I:\Omega\rightarrow\mathbb{R},\Omega\subseteq\mathbb{R}^2$ a 2-D panchromatic image, where Ω and $\mathbb{R}$ denote the image domain

and real value, respectively. For each pixel (x, y) in I, its Hessian matrix $H(x, y)$ is given by

$$H(x,y) = \begin{bmatrix} G_\sigma * \frac{\partial^2 I}{\partial x \partial x} & G_\sigma * \frac{\partial^2 I}{\partial x \partial y} \\ G_\sigma * \frac{\partial^2 I}{\partial y \partial x} & G_\sigma * \frac{\partial^2 I}{\partial y \partial y} \end{bmatrix} \tag{7}$$

where G_σ is a Gaussian kernel centered at (x, y) with band width equal to σ. Eigen-decomposition [66] is subsequently applied to the Hessian matrix H, resulting in

$$\begin{aligned} H &= \left[\overrightarrow{e_1}\ \overrightarrow{e_2}\right] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \overrightarrow{e_1} \\ \overrightarrow{e_2} \end{bmatrix} \\ &= \lambda_1 \overrightarrow{e_1}\overrightarrow{e_1}^T + \lambda_2 \overrightarrow{e_2}\overrightarrow{e_2}^T \end{aligned} \tag{8}$$

whose eigenvalues are $\lambda_i, i \in \{1, 2\}$ ordered such as $\lambda_1 \geq \lambda_2$, and $\overrightarrow{e_i}$ represent the corresponding eigenvectors. The two parameters R and S are computed as

$$\begin{cases} R &= \left|\frac{\lambda_2}{\lambda_1}\right| \\ S &= \left(\lambda_1^2 + \lambda_2^2\right)^{\frac{1}{2}} \end{cases} \tag{9}$$

where R is essential to separate line and plane structures since only in the line-like case it becomes zero; while S is a measure of the second order linear-ness. Note that value of S becomes low for homogeneous backgrounds, a case when all eigenvalues tend to become small. To extract linear features, we start from the likelihood function p_E defined as

$$p_E(x,y) = \begin{cases} 0, & \text{if } \lambda_1 \geq 0 \\ e^{-\frac{R^2}{2\beta^2}}\left(1 - e^{-\frac{S^2}{2c^2}}\right), & \text{otherwise} \end{cases} \tag{10}$$

where $\beta = 0.5$ and $c = (1/2)\max\{S(x) | x \in \Omega\}$ as suggested by Frangi *et al.* [67]. Note that the likelihood function represents the probability that a pixel belongs to a linear feature.

Although the likelihood function p_E generally shows a good performance, it suffers from two major problems [68]. 1) edges and similar high-contrast quasi-linear structures can still have large likelihood values, resulting in false positives. 2) It generally produces false positives in plain regions with low contrast and low signal-to-noise ratio (SNR). Therefore, the estimation of linearness probability of a pixel based on Hessian matrix alone is not suitable and a more general filter is required. To this aim, it will be useful to develop other indexes based upon the information extracted from the linear feature segment to which a point belongs, instead of the point itself. If a composite linearness measure can be computed according to both p_E and these quantities, it should be able to give smooth and reliable responses inside linear features, while suppressing noise in the background. In this study, linear feature enhancement using local standard variances is introduced to achieve these goals. The central idea is to model linear features by using straight line segments around a pixel and extract useful information from local standard variance and a geometrical feature, the circularity index.

2) Local Standard Variance and the Circularity Index: By regarding a local window with the size σ, equivalent to the Gaussian kernel bandwidth in (7), the mean radiance $\bar{r}_i(x, y, \sigma, \theta)$ of the ith spectral band along the direction identified by an angle can be expressed as

$$\bar{r}_i(x,y,\sigma,\theta) = \frac{1}{2\sigma}\sum_{k=-\sigma}^{\sigma} V_I \tag{11}$$

where $V_I = p_i([x + k\cos(\theta)], [y + k\sin(\theta)])$, $\theta \in \{0°, 1°, 2°, \ldots, 179°\}$. Its standard variance (SD) can be defined as

$$s_i(x,y,\sigma,\theta) = \left(\frac{1}{2\sigma - 1}\sum_{k=-\sigma}^{\sigma}\left(V_I - \bar{r}_i(x,y,\sigma,\theta)\right)^2\right)^{\frac{1}{2}}. \tag{12}$$

For multispectral satellite images, the total SD $tSD(x, y, \sigma, \theta)$ is the sum of SD of each spectral band, which is provided by computing

$$tSD(x,y,\sigma,\theta) = \sum_{i=1}^{B} s_i(x,y,\sigma,\theta) \tag{13}$$

where B is the band number. The lowest total SD (TSD) $tSD_o(x, y, \sigma, \theta)$ and its direction $\theta_o(x, y, \sigma, \theta)$ are defined as

$$\begin{cases} tSD_o(x,y,\sigma,\theta) &= \arg\max_{tSD} tSD(x,y,\sigma,\theta) \\ \theta_o(x,y,\sigma,\theta) &= \arg\max_{\theta} tSD(x,y,\sigma,\theta). \end{cases} \tag{14}$$

From (14), it is observed that high tSD_o value indicates a likely presence of edge or noise at the pixel (x, y); while low tSD_o value indicates that there is a possible linear segment along $\theta_o(x, y, \sigma, \theta)$.

Additionally, in this work, we exploit the circularity index [64], originally defined for a 2-D binary X, as

$$C(X) = \frac{4S(X)}{\pi L^2(X)} \tag{15}$$

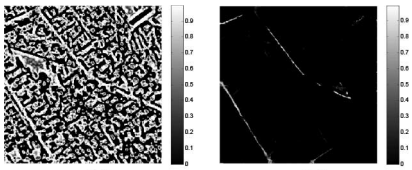
where $S(X)$ is the area of the object, $L(X)$ is the geodesic diameter of the object. The same operator can be extended to process gray level images, as presented in [69]. After applying the circularity operator, the gray values of pixels located in a circular structure do not change or change slightly. Accordingly, the probability that a pixel belongs to a circular structure can be computed as

$$s_p = |I - C| \tag{16}$$

where I represents a 2-D panchromatic image.

It is worth to note that a low value of s_p indicates high probability of a homogeneous structure. After obtaining the Hessian matrix, the lowest TSD, and the circularity index, a linearness filter can be defined as follows:

$$p_E(x,y) = \begin{cases} 0, \text{if } \lambda_1 \geq 0 \text{ or } tSD_o \geq 0.1 \text{ or } s_p \leq 10 \\ e^{-\frac{tSD_o^2}{3\sigma}} e^{-\frac{R^2}{2\beta^2}}\left(1 - e^{-\frac{S^2}{2c^2}}\right), \text{otherwise} \end{cases} \tag{17}$$



(a) (b)

Fig. 3. (a) True color image of QuickBird image. The ground truth dataset, obtained by visual interpretation, is shown in yellow; (b) EM clustering result.

where $tSD_o \geq 0.1$ represents possible edge/boundary regions, $s_p \leq 10$ denotes low contrast or low SNR regions, a feature that commonly appear in plane structure. It can be seen that these two conditions are complementary to the one by the Hessian matrix, and helps us to tackle its limitations to some extent. In practical applications, roads usually have different widths, and a multiscale processing is required. Given a set of scales Ω_σ, linearness filter responses at multiple scales are combined using a max rule

$$p_E(x,y) = \max\left\{ {p_E}^{\sigma_i}(x,y) \middle| \sigma_i \in \Omega_\sigma \right\} \tag{18}$$

where $\Omega_\sigma = \{3,5,7,9,11\}$ (in pixels) in this study, and can be easily adapted to images with different spatial resolutions, in case. Fig. 3(a) and (b) compares the results of the original linearness filter and the proposed one. It can be seen that the original linearness filter produces high response values at edge area and some background areas. By contrast, the proposed filter overcomes these limitations to some extent, and returns a much cleaner version of the linear features. After obtaining the linearness map, Otsu's method [70] is applied to produce a binary mask.

To improve the extraction, in this paper, a pixel-wise ratio operator, called R_{vs} [71], is used to detect vegetation and shadow areas. The index involves a comparison between the values of saturation (B_S) and intensity (B_I) bands, according to the very simple definition

$$R_{vs} = \frac{B_S - B_I}{B_S + B_I}. \tag{19}$$

The advantage of this ratio is that it can simultaneously detect vegetation and shadow areas. To detect vegetation and shadow areas, Otsus method [70] is applied to the histogram of the ratio map R_{vs}. After they are identified, these areas are removed from the preliminary road map produced by EM clustering.

C. Centerline Extraction From Road Segments

This step involves centerline extraction from candidate road segments. It must be pointed out that the road segments produced by previous steps have an irregular shape. They usually have undesired features caused by occlusions and the different gray values in a segment due to image noise (e.g., branches and holes). Regarding these factors, traditional methods, such as the thinning algorithm [64], do not usually result into satisfactory centerlines, which instead include spurious features that reduce their smoothness without retaining the correct spatial topology. Previous studies have shown that regression techniques do not produce spurs and retain the road centerline smoothness [56]. Indeed, regression methods are good options to extract centerlines from road segments. However, the observation data (i.e., the coordinates of the pixels in a candidate road segment) are corrupted by noise and false/outlier data (i.e., pixels on edge and holes). Accordingly, regression results may be strongly influenced by outliers (see Fig. 4). A robust estimation is thus needed to obtain satisfactory centerlines.

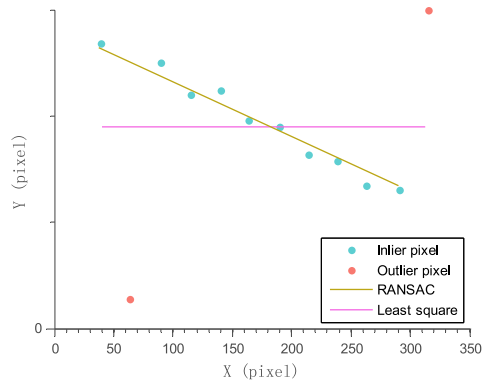


Fig. 4. Difference between least square regression and RANSAC.

To achieve stable and accurate results, this study relies on RANdom SAmple Consensus (RANSAC) [72] to extract centerlines from road segments, thanks to its efficiency and robustness to outliers. The RANSAC algorithm is a nonparametric method composed by two main steps:

1) a random selection of a subset of data points to start the model selection;
2) a classification of all data points as inliers or outliers by using this model.

Steps 1) and 2) are iteratively repeated and eventually the largest inlier set is selected, and the model re-estimated from it. Fig. 5 presents an example of road centerline extraction using the RANSAC. It can be seen that RANSAC eventually splits the observation data into two groups: 1) outlier group and 2) inlier group, outliers are directly discarded and inlier are used to extract the centerline.

In this step, road shape features can be integrated into RANSAC to remove miss-classified road pixels. Here, two shape metrics are employed, including: 1) area and 2) length. As the road is an elongated feature, roads usually do not have small area values. Therefore, segments with areas smaller than T_A are likely to be false positive and can be ignored.

Another road metric is its length. To check it, the segment centerline is removed if its length is smaller than the given threshold T_L, and operation that can be expressed by

$$L_i = \begin{cases} \text{RANSAC}(A_i), & \text{if length}(\text{RANSAC}(A_i)) \geq T_L \\ \varnothing, & \text{otherwise.} \end{cases} \tag{20}$$

In a real-world application, road candidates frequently have many branches, and RANSAC cannot be directly applied. To address this limitation, this study designed a modified RANSAC algorithm to extract centerlines from arbitrary road segments (reported in Algorithm 1). The central idea of this

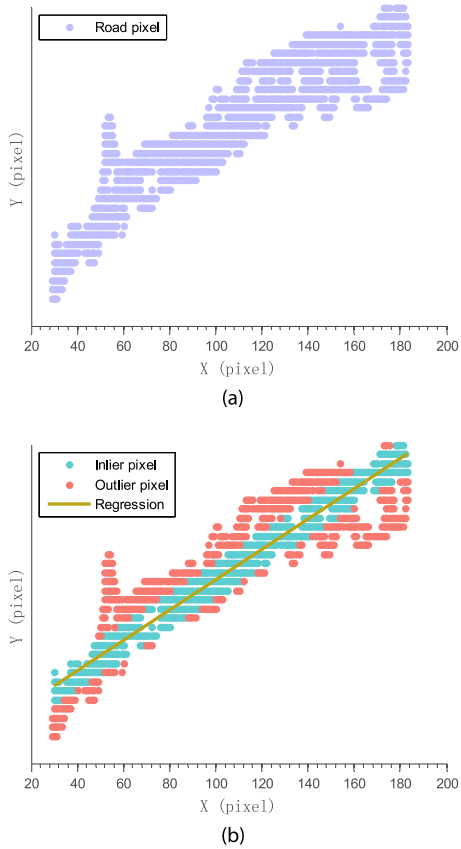


Fig. 5. Example of (a) road pixels identified in a candidate road segment and (b) road centerline extraction using RANSAC, where purple points are eventually recognized as outliers, blue points are considered as inliers, and the brown line is the regression result.

algorithm is to implement RANSAC in a recursive manner. For a complicated road segment (i.e., curved and with branches), Algorithm 1 divides the centerline into several straight lines. Although this selection slightly decreases the road centerline smoothness, it is better suited to store road centerlines in a GIS database. Additionally, this method makes it easier to implement information fusion techniques at the centerline level to integrate results from different extraction techniques, as discussed in the following section.

D. Information Fusion

After extracting candidate road segments using multiple methods, our goal now is to jointly analyze these results in the framework of information fusion. Previous studies show that information from multiple sources can complement each other and be beneficial. Clearly, fusing centerlines requires a set of rules. To make the wisest possible choice, in this study, three different rule sets are compared.

1) Object-Level Information Fusion: Object-level information fusion includes two different rule sets, named "rule set I" and "rule set II." The two sets jointly analyze objects obtained from EM clustering and the linearness filter to produce the final input to the RANSAC road centerline extraction. *Rule set I*. Suppose Ω_S and Ω_L are the sets of road segments extracted by

Algorithm 1. RANSAC for accurate road centerline extraction

Input
S represents a road segment.
Ω_S stands for the coordinates of pixels in S.
Ω_{inlier} denotes the coordinates of inliers produced by RANSAC.
N is the pixel number.
Output
L denotes the lines produced by RANSAC.

while $number\left(\Omega_{inlier}\right) \leq N$ **do**
  Perform RANSAC on Ω_S, producing a line l_i, inliers Ω^i_{inlier}, and outliers $\Omega^i_{outlier}$;
  $L\{i\} = l_i$; $\Omega_{inlier}\{i\} = \Omega^i_{inlier}$;
  $S \leftarrow LargestConnectedComponent\left(\Omega_S - \Omega_{inlier}\right)$;
  $\Omega_S \leftarrow \Omega\left(S\right)$;
  if $\frac{number(\Omega_{inlier})}{N} \geq 0.9$ **then**
    break;
  else
    continue.
  end if
end while

EM clustering and the linearness filter, respectively. Rule set I is defined as

$$R^I = \text{RANSAC}\left(\Omega_S \wedge \Omega_L\right). \tag{21}$$

Rule set II. It substitutes the logical AND operator in rule set I with the logical OR operator, and can be expressed as

$$R^{II} = \text{RANSAC}\left(\Omega_S \vee \Omega_L\right). \tag{22}$$

2) Centerline-Level Information Fusion: Rule set III. Differently from to rule sets I and II, rule set III is an ordered procedure implemented by applying a set of regularization steps, defined as

$$R^{III} = P_4 \cdot P_3 \cdot P_2 \cdot P_1 \left\{\text{RANSAC}\left(\Omega_S\right) \vee \text{RANSAC}\left(\Omega_L\right)\right\} \tag{23}$$

where P_1, P_2, P_3, and P_4 represent four spatial regularization techniques, introduced in the following paragraphs. a) P_1: Parallel centerlines are merged into one if the distance between them is small enough. In a 2-D Euclidean space, two segments are said to be parallel if their extensions to infinitely long lines never cross each other, as illustrated in Fig. 6(a). In discrete image processing, it may be a challenging task to quickly determine whether two straight segments are parallel or not. This study solves this problem by using an unsupervised clustering technique. Suppose L_i and L_j be two straight lines, and their corresponding end points be $\left\{M_i^1, M_i^2, M_j^1, M_j^2\right\}$. First, line segments are grouped into five categories based on their slopes. Then, two lines $\left(L_i, L_j\right)$ are judged to be parallel if and only if they fall into the same category. In this study, K-means algorithm [73] is selected for the clustering task due to its simplicity and efficiency.

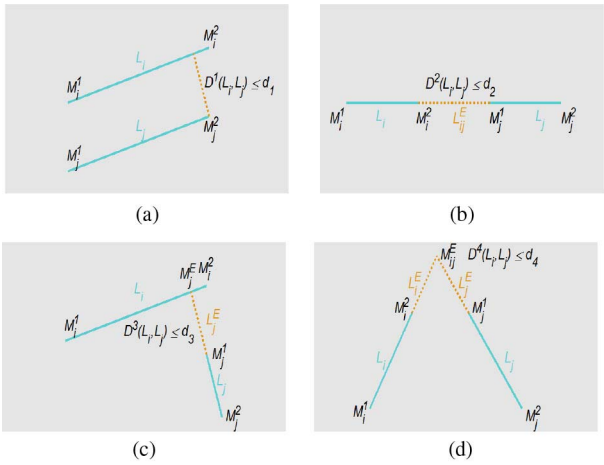


Fig. 6. Four spatial regularization strategies considered in rule set III. (a) Parallel segments. (b) Collinear segments. (c) Extension of one line segment intersecting the extension of another. (d) Segments that intersect each other.

As mentioned, once two parallel lines are individuated, they are merged. To this aim, if the distance between the two segments is smaller than a predefined threshold d_1, then the shorter segment is discarded. Hence, the spatial regularization rule P_1 for the parallel case is given by

$$P_1\left(L_i, L_j\right) = \begin{cases} \arg\max\limits_{l} \left(L_i, L_j\right), & \text{if } D^1 \le d_1 \\ \left(L_i, L_j\right), & \text{otherwise} \end{cases} \tag{24}$$

where $D^1 = \left(\left\|M_j^1 M_j^{1\prime}\right\| + \left\|M_j^2 M_j^{2\prime}\right\|\right)/2$ and $M_j^{2\prime}$ are projected points of M_j^1 and M_j^2 on L_i, respectively. b) P_2: Centerlines which are collinear/coincident are merged. Two linear segments are said to be collinear if they lay on the same straight line, as illustrated in Fig. 6(b). In this case, if the smallest distance between their end points is smaller than predefined threshold d_2, then the two segments will be connected by extending the longer one to touch the other one

$$P_2\left(L_i, L_j\right) = \begin{cases} LSF\left(L_i, L_j\right), & \text{if } D^2 \le d_2 \\ \left(L_i, L_j\right), & \text{otherwise} \end{cases} \tag{25}$$

where $D^2 = \left\|M_i^2 M_j^1\right\|$ and LSF stands for least square fitting (LSF) [73], a technique robust to the deviation between L_i and L_j. c) P_3: A centerline intersecting the extension of another one close to one of its ends [see Fig. 6(c)], is trimmed to the intersection point. In particular, when two segments or their extension form congruent adjacent angles to each other, they are said to be perpendicular. In this case, if the length of the extension line is smaller than a predefined threshold d_3, then these two lines will be connected. The spatial regularization strategy P_3 for this case is thus defined as

$$P_3\left(L_i, L_j\right) = \begin{cases} \left(L_i, L_j, L_j^E\right), & \text{if } D^3 \le d_3 \\ \left(L_i, L_j\right), & \text{otherwise} \end{cases} \tag{26}$$

where L_j^E is the extension line of L_j; $D^3 = \left\|M_j^1 M_j^E\right\|$, where M_j^E is the extension point of L_j. d) P_4: Two centerlines intersecting each other when extended are elongated until their intersection. To this aim, recall that the general equation of a straight line is $ax + by + c = 0$, where a, b, and c are real numbers. Considering two straight lines,

$$a_1 x + b_1 y + c_1 = 0 \tag{27}$$

$$a_2 x + b_2 y + c_2 = 0. \tag{28}$$

If $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$, then the two lines are intersecting. In this case, if the maximum length of two extension lines is smaller than a predefined threshold d_4, the two lines will be extended and connected [see Fig. 6(d)]. The spatial regularization strategy for the intersection case can be expressed as

$$P_4\left(L_i, L_j\right) = \begin{cases} \left(L_i, L_j, L_i^E, L_j^E\right), & \text{if } D^4 \le d_4 \\ \left(L_i, L_j\right), & \text{otherwise} \end{cases} \tag{29}$$

where L_i^E and L_j^E are extension lines of L_i and L_j, respectively; $D^4 = \max\left\{\left\|M_i^2 M_{ij}^E\right\|, \left\|M_j^1 M_{ij}^E\right\|\right\}$ where M_{ij}^E is the intersection point of L_i and L_j.

E. Parameter Tuning

The proposed method depends on several parameters (e.g., six thresholding values), and the choice of these parameters has an impact on the performance of the algorithm. It has formally been proved in the No Free Lunch theorem [74] that it is impossible to tune a parametric algorithm such that it has optimal settings for all possible natural image scenes. Accordingly, in this paper, we carry out an empirical analysis on parameter tuning. As a matter of fact, settings that significantly outperform the values obtained by this empirical analysis might be possible, but they seem not easy at all to obtain.

Specifically, the developed empirical model depends only on one information, i.e., the average road width R_w in the scene. The six thresholds used in this work are computed according to: 1) $T_A = 15 * R_w$; 2) $T_L = 2 * R_w$; 3) $d_1 = 5 * R_w$; 4) $d_2 = 4 * R_w$; 5) $d_3 = 3 * R_w$; and 6) $d_4 = \lfloor 2.5 * R_w \rfloor$, where $\lfloor \cdot \rfloor$ rounds a number to the nearest integer less than or equal to it.

III. Experimental Results

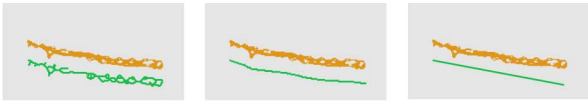
In this section, the proposed method is validated by using two multispectral satellite images. For quantitative evaluation, three measures are computed: Completeness $= TP/(TP + FN)$, Correctness $= TP/(TP + FP)$, and Quality $= TP/(TP + FP + FN)$ [75], where TP, FN, and FP represent true positives, false negatives, and false positives, respectively. The experiments run under MATLAB R2010b 32 b on a PC with Intel Core 2 CPU at 2.26 GHz, 2-GB RAM equipped with Window 7. For a fair comparison, the parameters were selected when the best performance is produced via trial-and-error test. The parameters used are provided in Table I.

A. Comparison of Different Centerline Extraction Algorithms

The advantage of the proposed approach for road centerline extraction from the classified image has been checked first. To

TABLE I
PARAMETERS USED FOR EACH TEST SITE

Test site	Parameter values
1	$C = 5,\ R_w = 20$
2	$C = 4,\ R_w = 15$



(a) (b) (c)

Fig. 7. Centerline results by (a) the thinning algorithm [66], (b) MARS [56], and (c) RANSAC, respectively. The background is shown in gray, while road pixel in brown, and centerlines in green. The extracted centerlines have been shifted from their true position to make the underlying pixels belonging to the road segments visible.

this end, the proposed method is compared with two methods existing in technical literature, that are 1) the thinning algorithm [66] and 2) the multivariate adaptive regression splines (MARS) [56]. The results of the comparison are reported in Fig. 7.

Generally speaking, the edges of road segments have a zigzag shape due to the complicated features adjacent to road, while there are many holes in the road segments caused by image noise, as illustrated by brown pixel in Fig. 7. These factors make the accurate centerline extraction from the road segment a challenging problem. As it can be observed from results by the three centerline extraction methods mentioned above, the thinning algorithm is the one producing more unwanted elements, such as spurs, bridges and loops [see Fig. 7(a)]. Although the presence of spurious elements does not affect the interpretation of road centerline, it reduces its smoothness and causes inaccuracies in road network patterns. By contrast, both MARS and RANSAC can provide results more accurate than the thinning algorithm, as illustrated in Fig. 7(b) and (c). Indeed, with MARS and RANSAC, there are no spurs, bias, or phase transitions. Additionally, MARS provides smoother result than RANSAC. This is due to that MARS relies on a regression of higher order than RANSAC. Note that the MARS result is a curve, while the result of RANSAC is a straight line, which in turn make it easier to store these data into a GIS. Thus, from the practical viewpoint, RANSAC is much more suitable for GIS than MARS. Additionally, RANSAC results can be directly used as inputs to rule set III.

It can be concluded that RANSAC achieves the best balance between accuracy and data storage efficiency. Hence, it is selected in this study to delineate road centerlines from classified images.

B. Comparison of Different Information Fusion Rules

In this step, the performance of the three different information fusion rule sets (i.e., rule set I, rule set II, and rule set III) is checked. For a fair comparison, the best performance of each rule is defined as the final output of this rule. Fig. 8 presents the results of all of the information fusion rules. The comparison among results from different information fusion rules is reported in Table II. Fig. 8 shows that rule set I provides much

(a) (b) (c)

Fig. 8. (a)–(c) The information fusion results of rule set I, rule set II, and rule set III, respectively. TP, FP, and FN are shown in red, green, and blue, respectively.

TABLE II
QUANTITATIVE EVALUATION RESULTS OF DIFFERENT INFORMATION FUSION RULES

Fusion method	Completeness	Correctness	Quality
Rule set I	0.60	1	0.60
Rule set II	0.84	0.80	0.69
Rule set III	0.82	0.92	0.77

cleaner extractions than the other two rule sets, and it produces the most correct performance. This coincides with the best correctness value of rule set I in Table II. The rationale for this result is that the logical AND operation aims at retaining as much correct information as possible. From Table II, it can be seen that that rule set II outperforms the other two rule sets in term of completeness (because of the logical OR). However, a comparison of Fig. 8(a) and (c) shows that rule set III can retain the correct roads produced by rule set I, while at the same time also inheriting the completeness property of rule set II. A comparison between Fig. 8(b) and (c) shows that rule set III can remove more false positives than rule II. This is due to the regularization strategies that make full use of the information provided by the segmentation method and the linear-ness filter. The results in Table II suggest that the information fusion rule at the centerline level produces the best balance between completeness and correctness, making it the method of choice from now on.

C. Experiment 1

The proposed method was tested first on a QuickBird image of the urban area of Pavia (Italy). The spatial size of the test image is 1024×1024 pixels and Fig. 9(a) shows the study area. The test image has four multispectral bands with the spatial resolution of 2 m/pixel and the proposed method is validated on the original multispectral bands with pan-sharpening. Fig. 9(b)–(c) presents the results of different road extraction strategies. From Fig. 9, it can be seen that EM fails to extract road in some parts (see A–D green rectangles), where roads are not properly shown due to the shade of buildings, trees, or their shadows. The EM clustering method is not always able to correctly extract shaded roads, as it is relying on only spectral information. By contrast, the linearness filter is successful in recovering road information in the same shaded areas. The reason for this success is that, although the spectral information of the shaded road change, its shape feature is retained, and this makes the linearness filter be able to delineate it. Similarly, Fig. 9 also shows that the linearness filter cannot extract road pixels in some areas

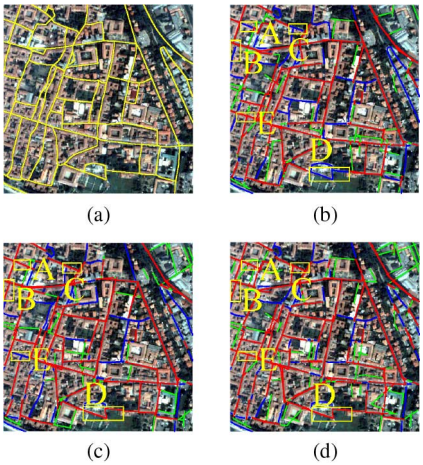


Fig. 9. Comparison results of different road extraction strategies on the Pavia test area. (a) Ground truth. (b) EM result. (c) Linearness filter result. (d) Information fusion result. TP, FP, and FN are shown in red, green, and blue, respectively.

TABLE III
QUANTITATIVE EVALUATION RESULTS OF DIFFERENT ROAD EXTRACTION STRATEGIES ON PAVIA TEST AREA

Method	Completeness	Correctness	Quality
EM	0.50	0.66	0.40
Linearness filter	0.54	0.65	0.42
Fusion	0.69	0.63	0.49

(see the E green rectangle), while the EM clustering is successful in those areas. Thus, EM and the linearness filter provide complementary information: by fusing these two methods, the overall performance improved [see Fig. 9(d)]. Accordingly, it can be seen that some discontinuities are successfully eliminated using the proposed road tracking rules. However, there are still some wrongly connected results, proving that automatic road tracking is still a challenging issue. An operational solution is to interactively postprocess road centerline segments by means of semiautomated methods.

Table III reports a quantitative assessment of different road extraction strategies. It can be seen that, with respect to completeness, the fusion method exhibits an obvious advantage over its competitors. In addition, EM outperforms the linearness filter and the fusion method in the metric of correctness. Compared to EM, the correctness value of the fusion method has slightly decreased, due to some wrongly regularized result produced by rule set III. From the viewpoint of quality, Table III clearly illustrates that the fusion method yields a substantial improvement with respect to the other two methods. The fusion method improves the quality of EM clustering by 9%, and the linearness filter results by 7%. Therefore, the fusion method already yields the best quality performance, producing the best tradeoff between completeness and correctness.

D. Experiment 2

The second experiment was performed on a dataset acquired by QuickBird on Xuzhou City, Peoples Republic of China. A subset (1024 pixels by 1024 pixels) of the whole pan-sharpened

Fig. 10. Comparison of the results of different road extraction strategies on the Xuzhou test area. (a) Ground truth. (b) EM results. (c) Linearness filter results. (d) Fusion result. TP, FP, and FN are shown in red, green, and blue, respectively.

TABLE IV
QUANTITATIVE EVALUATION RESULTS OF DIFFERENT ROAD EXTRACTION STRATEGIES ON XUZHOU TEST AREA

Method	Completeness	Correctness	Quality
EM	0.69	0.64	0.50
Linearness filter	0.66	0.52	0.41
Fusion	0.84	0.61	0.54

image was selected as the experimental area, and presented in Fig. 10(a). The road centerline extraction results using EM segment only, the linearness filter only, and the proposed method are shown in Fig. 10(b)–(d), respectively. In Fig. 10, the ground truth dataset is shown in yellow; while extraction results in red.

Table IV reports a quantitative evaluation of the results. The fusion rule produces the largest completeness value while EM the largest correctness value. This is because the fusion result starts from different road extraction results and thus has higher probability to detect portions of the true road network. Similarly to what happens in the first test site, the fusion method decreases correctness compared to EM clustering, because of the accumulation of errors from multiple road results and some wrongly regularized results produced by rule set III. With respect to the quality metric, it can be seen that the values of EM clustering, the linearness filter, and the fusion method are 0.50, 0.41, and 0.54, respectively. In these three methods, the tracking rules improve the quality performance from 4% to 13%, which is compatible with the conclusion of the first experiment. Hereby, the fusion method surpasses the other two competitors. This again verifies the superiority of the fusion method in improving road extraction accuracy.

E. Comparison With State-of-the-Art Methods

To assess the results of the presented method, in this area, a comparison was carried out with the two state-of-the-art methods: 1) Song's method [23] and 2) Huang's method [48]. Fig. 11 shows the comparison results of these two road extraction methods. As it can be seen, road features are extracted well by the proposed method. This performance is also confirmed

Fig. 11. Comparison results of different road extraction methods on two test areas: the first column show the results of Song's method [23], while the second column the results of Huang's method [48] on Pavia and Xuzhou, respectively. TP, FP, and FN are shown in red, green, and blue, respectively.

TABLE V
COMPARISON OF DIFFERENT ROAD EXTRACTION METHODS

Study area	Method	Completeness	Correctness	Quality
Pavia	Song [23]	0.72	0.45	0.38
	Huang [48]	0.57	0.50	0.37
	Proposed	0.69	0.63	0.49
Xuzhou	Song [23]	0.97	0.32	0.32
	Huang [48]	0.94	0.48	0.46
	Proposed	0.84	0.61	0.54

quantitatively by road extraction accuracy in terms of three measures (i.e., completeness, correctness, and quality). Table V lists the quantitative evaluation results for the three methods. As shown in Table V, for the first study area, the quality values yielded by the three methods are equal to 0.38, 0.37, and 0.49, respectively. The fail of Song's method is that the misclassified road pixels are connected with the correct road pixels. Therefore, it is challenging to use shape features to postclassify SVM classification result (i.e., separating correct road pixels from misclassified road pixels). The use of thinning algorithm in Huang's method leads to low correctness value, which in turn decreases the quality metric. Thus, the proposed method yields the best quality value in the first test case. For the second study area, although Song's method and Huang's method produce higher completeness values than the proposed method, the lower correctness values significantly decrease their extraction performance. Therefore, from the viewpoint of quality measure, the proposed method outperforms Song's method and Huang's method in both two test cases, which confirms the suitability of the proposed road extraction method for optical satellite images.

IV. DISCUSSION AND CONCLUSION

Urban road network remains one of the trickiest and most difficult features to be extracted: roads may be shadowed by trees and buildings; they are neither distinct from their surroundings nor completely homogeneous. Hence, they are difficult to be discriminated using automated approaches. Therefore, one of the underlying goals of this research was to show how to reduce inevitable issues in urban road network extraction combining results from multiple extraction methods. The proposed approach exploits information fusion rule sets, and is suited to VHR optical satellite images. The approach was tested on two images of densely built urban areas, where road networks are spectrally similar to their surroundings and partially occluded by trees and shadows.

The results reveal several insights for urban road network extraction that are relevant to future studies.

A. Integrating Local Information Can Help Linear Feature Enhancement

One typical road feature is that roads are locally straight; this feature has been quantified by means of a novel linearness filter. In image processing, Hessian matrix-based filters, such as Frangis filter [67], are commonly used for this task. The problem, however, is that they rely on image gradients or high-order derivatives, thus their responses are sensitive to noise and often too weak to discriminate road and nonroad pixels in low contrast regions. In addition, the Hessian matrix-based filters usually produce high responses around boundaries of different land cover classes, yielding difficulties in precisely localizing the exact boundary of a linear feature. Imprecise boundary localization could instead result in inaccurate quantification of road segments. This paper focuses on addressing these two challenging problems.

Specifically, we have studied the use of local information around a pixel together with its local structure derived from Hessian matrix. Experimental results show that the proposed approach helps discarding most false positives in low contrast, low SNR regions and at boundaries between land cover classes. We believe this approach provides far more cleaner result than the one proposed in [67].

B. Regression Can Extract Accurate Road Centerline

The efficacy and efficiency of the proposed regression method for road centerline detection highlight its excellent performance. Centerline extraction from road segments using this regression method has two advantages. 1) It is convenient for data storage. 2) The extracted centerline has no spurs and well retains the smoothness. The regression method can help to overcome spurs due to their spectral similarity to the surroundings, while holes are due to missing pixels. Indeed, our work provides evidence that it may be possible to extract smooth centerlines from road segments with coarse boundary and holes.

While the regression method is highly suited to detect centerlines from straight road candidates, this analysis also reveals biases that should be considered. First, the proposed method relies on a recursive way to process curved road segments, so the technique is limited in some regions where road segments have branches. Although it is difficult to process complicated road segments, the proposed method can produce preliminary centerlines that retain correct spatial topology. It is possible to overcome this limitation by considering a semi-automatic method, and including a user interaction option [48]. This allows greater freedom, but also leads to increased user input

and greater time consumption, which may in turn restrict the utility of the method in some situations.

C. Information Fusion Helps Yield High Road Delineation Accuracy in Complex Urban Scene

While the amounts and quality of remote sensing data are unprecedented, there is still no one perfect data source for the difficult task of mapping urban expansion. This work has shown that combining results extracted from multiple methods is advantageous given the complexity and heterogeneity of urban road network. Depending on the spectral and spatial characters, urban road networks can be extracted to some extent via automatic means. Our results show that linearness filter produces a higher accuracy when the urban road network is suffering the similar and/or different spectral similarity of surroundings, while the segmentation method achieves higher accuracy when the urban road network is homogeneous. Therefore, a road candidate may be extracted by the former, while another one by the latter extraction method. By exploiting these complimentary datasets and fusing the information obtained from both of them, a more accurate extraction is achieved.

While this study focuses on the fusion of two methods (i.e., EM clustering and the linearness filter) relying on a single data source (i.e., optical satellite images), future works will integrate multiple methods and multiple data sources, such as optical satellite images, LIDAR, and SAR.

ACKNOWLEDGMENT

The authors gratefully acknowledge Dr. G. C. Iannelli from the University of Pavia for providing some of the data sets used in this work and assisting in analysing them.

REFERENCES

[1] World Bank Group, *East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth*. Washington, DC, USA: World Bank, 2015.

[2] World Bank and the Development Research Center of the State Council, *Urban China: Toward Efficient, Inclusive, and Sustainable Urbanization.* Washington, DC, USA: World Bank, 2014.

[3] 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.

[4] X. Hu, Z. Zhang, and C. V. Tao, "A robust method for semi-automatic extraction of road centerlines using a piecewise parabolic model and least square template matching," *Photogramm. Eng. & Remote Sens.*, vol. 70, no. 12, pp. 1393–1398, 2004.

[5] R. Kimmel, "Numerical geometry of images: Theory, algorithms, and applications," 1st ed. New York, NY, USA: Springer-verlag, 2004.

[6] X. Lin, J. Zhang, Z. Liu, J. Shen, and M. Duan, "Semi-automatic extraction of road networks by least squares interlaced template matching in urban areas," *Int. J. Remote Sens.*, vol. 32, no. 17, pp. 4943–4959, 2011.

[7] Z. Miao, B. Wang, W. Shi, and H. Zhang, "A semi-automatic method for road centerline extraction from VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1856–1860, Nov. 2014.

[8] 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.

[9] X. Lin, J. Zhang, Z. Liu, and J. Shen, "Semi-automatic extraction of ribbon roads form high resolution remotely sensed imagery by cooperation between angular texture signature and template matching," *Int. Arch. Photogramm. Remote Sens.*, vol. 34, pp. 309–312, 2008.

[10] J. Zhou, W. F. Bischof, and T. Caelli, "Robust and efficient road tracking in aerial images," in *Proc. Joint Workshop ISPRS German Assoc. Pattern Recog., Vienna, Austria, Aug. 29–30*, 2005, pp. 35–40.

[11] M. Bicego, S. Dalfini, G. Vernazza, and M. Murino, "Automatic road extraction from aerial images by probabilistic contour tracking," in *Proc. Int. Conf. Image Process. (ICIP)*, 2003, vol. 3, pp. III–585-8.

[12] 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.

[13] A. P. Dal Poz, R. A. Gallis, J. F. da Silva, and E. F. Martins, "Object-space road extraction in rural areas using stereoscopic aerial images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 654–658, Jul. 2012.

[14] 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. 985–994, 1997.

[15] I. Laptev, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, "Automatic extraction of roads from aerial images based on scale space and snakes," *Mach. Vis. Appl.*, vol. 12, no. 1, pp. 23–31, 2000.

[16] R. Marikhu, M. N. Dailey, S. Makhanov, and K. Honda, "A family of quadratic snakes for road extraction," in *Proc. Comput. Vis. (ACCV)*, 2007, pp. 85–94.

[17] M. Rochery *et al.*, "New higher-order active contour energies for network extraction," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2005, vol. 2, pp. II–822-5.

[18] Z. Li, W. Shi, Q. Wang, and Z. Miao, "Extracting man-made objects from high spatial resolution remote sensing images via fast level set evolutions," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 883–899, Feb. 2015.

[19] M. Rajeswari, K. Gurumurthy, S. Omkar, J. Senthilnath, and L. P. Reddy, "Automatic road extraction using high resolution satellite images based on level set and mean shift methods," in *Proc. 3rd Int. Conf. Electron. Comput. Technol. (ICECT)*, 2011, vol. 2, pp. 424–428.

[20] R. Péteri, J. Celle, and T. Ranchin, "Detection and extraction of road networks from high resolution satellite images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2003, vol. 1, pp. I–301-4.

[21] M. Butenuth and C. Heipke, "Network snakes: Graph-based object delineation with active contour models," *Mach. Vis. Appl.*, vol. 23, no. 1, pp. 91–109, 2012.

[22] Y. Nakaguro, S. S. Makhanov, and M. N. Dailey, "Numerical experiments with cooperating multiple quadratic snakes for road extraction," *Int. J. Geograph. Inf. Sci.*, vol. 25, no. 5, pp. 765–783, 2011.

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

[24] 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.

[25] W. Shi, Z. 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.

[26] 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.

[27] M. Mokhtarzade and M. V. Zoej, "Road detection from high-resolution satellite images using artificial neural networks," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 9, no. 1, pp. 32–40, 2007.

[28] S. Das, 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.

[29] E. Türetken, G. González, C. Blum, and P. Fua, "Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors," *Neuroinformatics*, vol. 9, nos. 2–3, pp. 279–302, 2011.

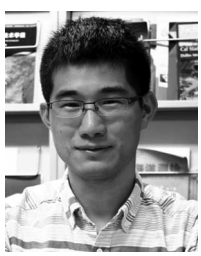
[30] E. Türetken, F. Benmansour, and P. Fua, "Automated reconstruction of tree structures using path classifiers and mixed integer programming," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2012, pp. 566–573.

[31] E. Turetken, F. Benmansour, B. Andres, H. Pfister, and P. Fua, "Reconstructing loopy curvilinear structures using integer programming," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2013, pp. 1822–1829.

[32] D. Chai, W. Forstner, and F. Lafarge, "Recovering line-networks in images by junction-point processes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2013, pp. 1894–1901.

[33] J. D. Wegner *et al.*, "A higher-order CRF model for road network extraction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2013, pp. 1698–1705.

[34] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images," in *Proc. Comput. Vis. (ECCV)*, 2010, pp. 210–223.
[35] V. Mnih and G. Hinton, "Learning to label aerial images from noisy data," in *Proc. 29th Int. Conf. Mach. Learn. (ICML'12)*, 2012, pp. 567–574.
[36] M. Baatz and A. Schäpe, "Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation," *in Angewandte Geographische Informationsverarbeitung XII*. Berlin, Germeny: Herbert Wichmann Verlag, 2000, pp. 12–23.
[37] M. Herold, J. Scepan, A. Müller, and S. Günther, "Object-oriented mapping and analysis of urban land use/cover using IKONOS data," in *Proc. 22nd Earsel Symp. Geoinf. Eur. Wide Integr.*, 2002, pp. 4–6.
[38] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3, pp. 239–258, 2004.
[39] T. Blaschke and J. Strobl, "What is wrong with pixels? Some recent developments interfacing remote sensing and GIS," *GeoBIT/GIS*, vol. 6, no. 1, pp. 12–17, 2001.
[40] C. Cleve, M. Kelly, F. R. Kearns, and M. Moritz, "Classification of the wildland–urban interface: A comparison of pixel-and object-based classifications using high-resolution aerial photography," *Comput. Environ. Urban Syst.*, vol. 32, no. 4, pp. 317–326, 2008.
[41] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 1, pp. 2–16, 2010.
[42] E. Baltsavias, "Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3, pp. 129–151, 2004.
[43] A. S. Laliberte *et al.*, "Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico," *Remote Sens. Environ.*, vol. 93, no. 1, pp. 198–210, 2004.
[44] R. C. Frohn, K. M. Hinkel, and W. R. Eisner, "Satellite remote sensing classification of thaw lakes and drained thaw lake basins on the North Slope of Alaska," *Remote Sens. Environ.*, vol. 97, no. 1, pp. 116–126, 2005.
[45] M. Ehlers, M. Gaehler, and R. Janowsky, "Automated techniques for environmental monitoring and change analyses for ultra high resolution remote sensing data," *Photogramm. Eng. & Remote Sens.*, vol. 72, no. 7, pp. 835–844, 2006.
[46] J. Im, J. Jensen, and J. Tullis, "Object-based change detection using correlation image analysis and image segmentation," *Int. J. Remote Sens.*, vol. 29, no. 2, pp. 399–423, 2008.
[47] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.
[48] X. Huang and L. Zhang, "Road centreline 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, 2009.
[49] 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.
[50] N. Zarrinpanjeh, F. Samadzadegan, and T. Schenk, "A new ant based distributed framework for urban road map updating from high resolution satellite imagery," *Comput. & Geosci.*, vol. 54, pp. 337–350, 2013.
[51] A. Grote, C. Heipke, and F. Rottensteiner, "Road network extraction in suburban areas," *Photogramm. Rec.*, vol. 27, no. 137, pp. 8–28, 2012.
[52] Y. Han, H. Kim, J. Choi, and Y. Kim, "A shape–size index extraction for classification of high resolution multispectral satellite images," *Int. J. Remote Sens.*, vol. 33, no. 6, pp. 1682–1700, 2012.
[53] S. Valero, J. Chanussot, J. A. Benediktsson, H. Talbot, and B. Waske, "Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images," *Pattern Recog. Lett.*, vol. 31, no. 10, pp. 1120–1127, 2010.
[54] 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.
[55] M. Negri, P. Gamba, G. Lisini, and F. Tupin, "Junction-aware extraction and regularization of urban road networks in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2962–2971, Oct. 2006.
[56] 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.
[57] Q. Zhang and I. Couloigner, "A wavelet approach to road extraction from high spatial resolution remotely-sensed imagery," *Geomatica*, vol. 58, no. 1, pp. 33–39, 2004.
[58] 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.
[59] P. Doucette, P. Agouris, A. Stefanidis, and M. Musavi, "Self-organised clustering for road extraction in classified imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 55, no. 5, pp. 347–358, 2001.
[60] Y. Shao, B. Guo, X. Hu, and L. Di, "Application of a fast linear feature detector to road extraction from remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 3, pp. 626–631, Sep. 2011.
[61] C. Cao and Y. Sun, "Automatic road centerline extraction from imagery using road GPS data," *Remote Sens.*, vol. 6, no. 9, pp. 9014–9033, 2014.
[62] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. Ser. B Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.
[63] VPLab. Downloads [Online]. Available: http://www.cse.iitm. ac.in/sdas/vplab/satellite.html, accessed on May 15, 2015.
[64] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*, 2nd ed. Knoxville, TN, USA: Gatesmark, 2009.
[65] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.
[66] W. H. Press, *Numerical Recipes* 3rd Edition: The Art of Scientific Computing. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[67] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Proc. Med. Image Comput. Comput. Assisted Intervent. (tMICCAI'98)*, 1998, pp. 130–137.
[68] Y. Yuan, Y. Luo, and A. Chung, "VE-LLI-VO: Vessel enhancement using local line integrals and variational optimization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1912–1924, Jul. 2011.
[69] V. Morard, E. Decenciere, and P. Dokládal, "Efficient geodesic attribute thinnings based on the barycentric diameter," *J. Math. Imag. Vis.*, vol. 46, no. 1, pp. 128–142, 2013.
[70] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, nos. 285–296, pp. 23–27, 1975.
[71] A. O. Ok, C. Senaras, and B. Yuksel, "Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1701–1717, Mar. 2013.
[72] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[73] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
[74] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
[75] C. Wiedemann, C. Heipke, H. Mayer, and O. Jamet, "Empirical evaluation of automatically extracted road axes," in *Proc. CVPR Workshop Empirical Eval. Techn. Comput. Vis.*, Los Alamitos, CA, USA, 1998, pp. 172–187.



Zelang Miao (S'15) received the B.S. degree in surveying engineering and the M.S. degree in geodesy and surveying engineering from China University of Mining and Technology, Xuzhou, China, in 2009 and 2014, respectively. He is currently working toward the Ph.D. degree in land surveying and geoinformatics at Hong Kong Polytechnic University, Kowloon, Hong Kong.

He is a Visiting Ph.D. Student at the Department of Industrial and Information Engineering, University of Pavia, Pavia, Italy, from March to May 2015. His research interests include pattern recognition, land cover and land use mapping, and global/regional urbanization.

Mr. Miao serves as a Referee for Remote Sensing, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, *Journal of Applied Remote Sensing*, and *Journal of Circuits, Systems and Signal Processing*.

Wenzhong Shi received the Ph.D. degree in GIS and remote sensing from the University of Osnabrck, Vechta, Germany, in 1994.

He is a Chair Professor in GIS and Remote Sensing, Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Kowloon, Hong Kong. He has authored over 100 research papers and 10 books. His 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 International Society for Photogrammetry and Remote Sensing in 2012.

Alim Samat (S'13) received the Ph.D. degree in cartography and geographic information system from Nanjing University, Nanjing, China, 2015.

Currently, he is the Assistant Researcher with the State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China. His research interests include remote sensing image applications for urban and terrain deformation, machine learning, and hyperspectral and polarimetric SAR remote sensing information processing.

Dr. Samat also serves as the Reviewer for several international journals including the IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTION ON CYBERNETICS, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and the *International Journal of Remote Sensing*.

Gianni Lisini received the Ph.D. degree in electronics engineering from the University of Pavia, Pavia, Italy, in 2005.

In 2003, he was a Visiting Researcher at the ENST Telecom Paris, Paris, France, working on road extraction from satellite urban data. He joined the Remote Sensing Group, Department of Electrical, Computer, and Biomedical Engineering, University of Pavia, Pavia, Italy. He has taken part to several research projects on related topics, including the EU FP7 project "Services and Applications for Emergency Response" (SAFER), which aims at implementing preoperational versions of the Emergency Response Core Service in the frame of the GMES initiative (Global Monitoring for Environment and Security). Currently, he has a research grant at IUSS-Pavia under EXTRAS, an FP7 Cooperation project, for studies about implementation of algorithms aimed at the systematic search of X-ray variable and transient sources in archival data. He is the author or coauthor of over 15 paper on international, peer-reviewed scientific journals, over 50 papers at international conferences, and 3 book chapters.

Paolo Gamba (SM'00–F'13) received the Laurea degree (*cum laude*) and the Ph.D. degree in electronic engineering from the University of Pavia, Italy, in 1989 and 1993, respectively.

He is an Associate Professor of Telecommunications with the University of Pavia, Pavia, Italy, where he also leads the Telecommunications and Remote Sensing Laboratory. He has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, EO data, and risk management. He has authored more than 110 papers in international peer-reviewed journals and presented more than 250 research works in workshops and conferences.

Dr. Gamba served as a Editor-in-Chief of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS from 2009 to 2013, and as a Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from October 2005 to May 2009. Currently, he is the Chair of the Chapters Committee of the same society. He has been the Organizer and Technical Chair of the Biennial GRSS/ISPRS Joint Workshops on Remote Sensing and Data Fusion over Urban Areas since 2001. He also served as a Technical Co-Chair of the 2010 IEEE Geoscience and Remote Sensing Symposium, Honolulu, Hawaii, July 2010, and Technical Co-Chair of the 2015 IEEE Geoscience and Remote Sensing Symposium, in Milan, Italy. He has been the Guest Editor of Special Issues of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN REMOTE SENSING APPLICATIONS, *ISPRS Journal of Photogrammetry and Remote Sensing*, *International Journal of Information Fusion*, and *Pattern Recognition Letters* on the topics of Urban Remote Sensing, Remote Sensing for Disaster Management, and Pattern Recognition in Remote Sensing Applications.