

Joint Enhancing Filtering for Road Network Extraction

Yu Zang, Cheng Wang, *Senior Member, IEEE*, Yao Yu, Lun Luo, Ke Yang, and Jonathan Li, *Senior Member, IEEE*

Abstract—In this paper, we propose a task-oriented enhancing technique for extracting road networks from satellite images. By exploiting an approximate estimation of the potential road edges for guidance, we developed a joint enhancing filtering framework to generate a version of the input image that facilitates road network extraction. First, an adaptive smoothing scheme is designed to suppress the interference of noise or heavy textures, such as residential areas or terrain boundaries. By combining this scheme with the proposed novel anisotropic shock filter, the edges of the potential road regions can be kept sharp and clear. Through abundant experimental comparisons with state-of-the-art filtering techniques and quantitative evaluations using data from various satellite sensors, the performance of the proposed approach is comprehensively evaluated. The experimental results demonstrate that our system can address heavy high contrast textures and provide a meaningful improvement in the feature detection for road extraction.

Index Terms—Guided image filtering, joint enhancing filtering, road network extraction.

I. INTRODUCTION

Road network extraction is a long-standing problem in remote sensing image processing. High-quality automatic road extraction results are valuable for the recognition of vehicles, buildings, or other road-related objects, and are also of great significance for reducing the effort required to acquire transportation data.

Road network extraction has been attracting considerable attention in the past decade. With the objective of achieving high-quality and robust road network extraction, the authors of existing studies addressed this problem from different perspectives. However, literature review revealed that in most of the studies, a common preprocessing step is involved to allow the better extraction of the road features. Specifically, Yuan *et al.* [1] employed a Laplacian of the Gaussian filter to

Manuscript received April 3, 2016; revised July 18, 2016 and September 19, 2016; accepted October 11, 2016. Date of publication December 7, 2016; date of current version February 23, 2017. This work was supported in part by the Natural Science Foundation of China under Grants 61501387, 61401382, and 61371144, the Postdoctoral Science Foundation of China under Grant 149621, and the High Resolution Earth Observation Civil-Military Integration Demonstration Project under Grant GFZX0404080102.

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

L. Luo and K. Yang are with the China Transport Telecommunications and Information Center, Beijing 100011, China.

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

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

Digital Object Identifier 10.1109/TGRS.2016.2626378

approximately segment the road-like regions. Ünsalan *et al.* [2] proposed a nonlinear median filtering for preprocessing to avoid the undesired details in the image, Shi *et al.* [3] proposed adaptive morphological operators to construct morphological profiles (MPs), and Zang *et al.* [4] employed a Gaussian filter together with a bilateral filter (BLF) to suppress false positive detections. These studies imply a consensus that preprocessing plays an important role in terms of achieving high-quality road network extraction. However, most of the applied filters are not associated with the targets, such as road-like structures, in road network extraction. Thus, an interesting problem arises, that is, whether it is possible to design more specific filters for specific tasks.

On the other hand, the most recently developed structure-aware [5]–[7] filtering techniques introduced new heuristics for road extraction. Recently, more effort has been invested in seeking a more flexible metric that suits structure-aware smoothing, such as those using the oscillations of local extrema [5], total variation (TV) with intensity gradients [6], and patch-based local covariance on a batch of features [7]. These methods exhibit commendable strength in dealing with structures free of a high/low contrast assumption, but the major challenge of these filters is that sometimes it is difficult to determine which structures are the most salient, which leads to unpleasant artifacts. However, for a particular application, such as road network extraction, the textures need to be smoothed while the main structures need to be enhanced in order to generate more accurate filtering results.

Inspired by these two types of methods, the objective of this paper is to design a task-related enhancing filtering technique that can be incorporated into road network extraction schemes. For many road extraction methods, it is easy to acquire a road structure measurement [3], [4]. By exploiting this approximate estimation as guidance, we propose a joint enhancing filtering framework to adaptively smooth textures, such as residential areas or terrain boundaries, while keeping the potential road edges sharp and clear. Benefiting from the explicit guidance, our system is able to address heavy high contrast textures and provide meaningful filter results for the feature detection involved in road extraction. We evaluated our approach using data from various satellite sensors and comprehensively compared it with previous state-of-the-art filters. The experimental results show that the proposed method can provide meaningful improvement in the feature detection for road extraction. It is also noted that, except the road network extraction, the major contribution of this paper relies on a task-related filtering framework, which can be easily extended to other remote

sensing applications such as coastline detection or oil spilling detection by simply replacing the explicit guidance.

II. RELATED WORK

In recent years, in many studies, the road network extraction problem has been investigated from different viewpoints, and noteworthy progress has been made. In previous surveys [8]–[11] and the most recent road extraction studies [1]–[4], [12], we found an interesting tendency that an increasing number of researchers prefer to include a pre-processing filter in their systems, to guarantee better feature extraction performance and high detection correctness.

Early road extraction studies were aimed at detecting roads using their visual or geometric features. With the assumption that road regions often appear as thin, low curvature, and high contrast structures, various filters, such as morphological [13], directional [14], and Kalman filters [15], were applied for road line feature extraction [16], [17] or primitive line connection [18]–[21].

Recently, road network extraction approaches have tended to employ more complex systems to further improve the detection performance, and more specific filters have been employed to provide better feature detection. Poullis and You [12] employed Gabor filtering for geospatial feature inference classification. Based on this, followed by orientation-based segmentation, in their method, road centerlines are extracted to describe the road network. Considering the observation that the pixels on the sides of the road boundaries have large responses, while the pixels within the roads have small responses to the Laplacian of the Gaussian, Yuan *et al.* [1] extracted roads automatically by clustering the well-aligned pixels according to a proposed locally excitatory globally inhibitory oscillator network. Ünsalan and Sirmacek [2] and Sirmacek and Ünsalan [22] utilized adaptive median filtering to eliminate small noise terms and proposed a graph-based topology analysis scheme to refine the road map in which spectral, shape, and gradient features are combined to generate approximate road primitives [2]. In the most recent road detection study [3], a road centerline extraction scheme was developed, based on a pretrained spectral–spatial classifier, which significantly improves the detection robustness. It is worth noting that, to suppress the interference of the undesired textures and overcome the blur effect of feature descriptor mathematical morphology (MM), general adaptive neighborhood-based MM [23] was applied to form the MPs. Ziems *et al.* [24] proposed a fusing scheme to combine the results of different models, in which different road detection methods are employed. Benefiting from its multimodel character, this approach presents an impressive robustness and detection performance. By exploiting various enhancing or smoothing filters, the methods in these studies are able to provide a better performance than those in which the extraction algorithm is applied directly to the original remote sensing images.

On the other hand, the development of structure-aware image processing techniques provides more interesting heuristics for road network extraction. In early filtering techniques, local intensity contrast was widely used for characterizing

dominant structures [25]–[29], of which the BLF proposed in [25] is a classical example. BLF takes both the spatial and intensity contrast into account, presenting several appealing features for edge-aware processing. Following this idea, to address the “halo”-free results, Farbman *et al.* [30] proposed a weighted least squares optimization scheme to preserve regions with large gradient magnitudes. Xu *et al.* [27] measured salient structures using the L_0 gradient-based optimization scheme. Li *et al.* [29] proposed a mix-domain filter, combining a set of nonlinear filters on each pixel and synthesizing the responses to obtain a globally optimized result. However, local contrast-based smoothing methods suffer from the major limitation that they possibly fail to smooth high contrast textures. Thus, they are not suitable for handling some challenging images with complicated textures, such as residential areas, for road detection.

To resolve the limitation of contrast-based smoothing techniques, various structure-aware filters have been developed to address high contrast textures. Subr and Soler [5] utilized local extrema distribution for the measurement of structures and details, where the details are assumed to be captured by the oscillations between local minima and maxima, and then, they employed empirical mode decomposition (EMD) to sandwich the underlying multiscale base layers. Zang *et al.* [31] employed a space-filling curve to thread all the image pixels in a 1-D sequence and computed local extrema and envelopes in the curve instead of a 2-D image grid. Such a domain reduction significantly simplifies the EMD process, thus gaining an extremely fast performance for structure-aware image smoothing based on local extrema. Inspired by the classical TV model [32], Buades *et al.* [33] proposed the *local TV* measurement to decompose an image into structures and oscillatory texture elements. Following this study, Xu *et al.* [6] proposed *relative TV* (RTV) to measure the structures of an image and obtain a plausible structure-aware smoothing effect by solving a global optimization problem. More recently, Karacan *et al.* [7] proposed patch-based local covariances of some prescribed features for classifying salient structures and textures. Using this metric, contrast-independent smoothing results are easily obtained by computing the local means of each pixel with the similarity as weights.

III. JOINT ENHANCING FILTERING

The proposed *joint enhancing filtering* utilizes the estimation of potential road-like structures to guide the preprocessing of the image, thus suppressing the interference of high contrast textures while enhancing the road-like regions.

Fig. 1 shows the framework of our approach. The entire system comprises two primary parts: the adaptive smoothing part is aimed to smooth heavy textures, such as buildings or residential areas while preserving potential road-like structures; the enhancement part is aimed to compensate for the possible blurriness of the road areas, thus providing sharp and clear road edges. Both parts are guided by an estimated road structure model.

To handle the high contrast textures better, inspired by the idea of image EMD [5], [34], the smoothing part is applied

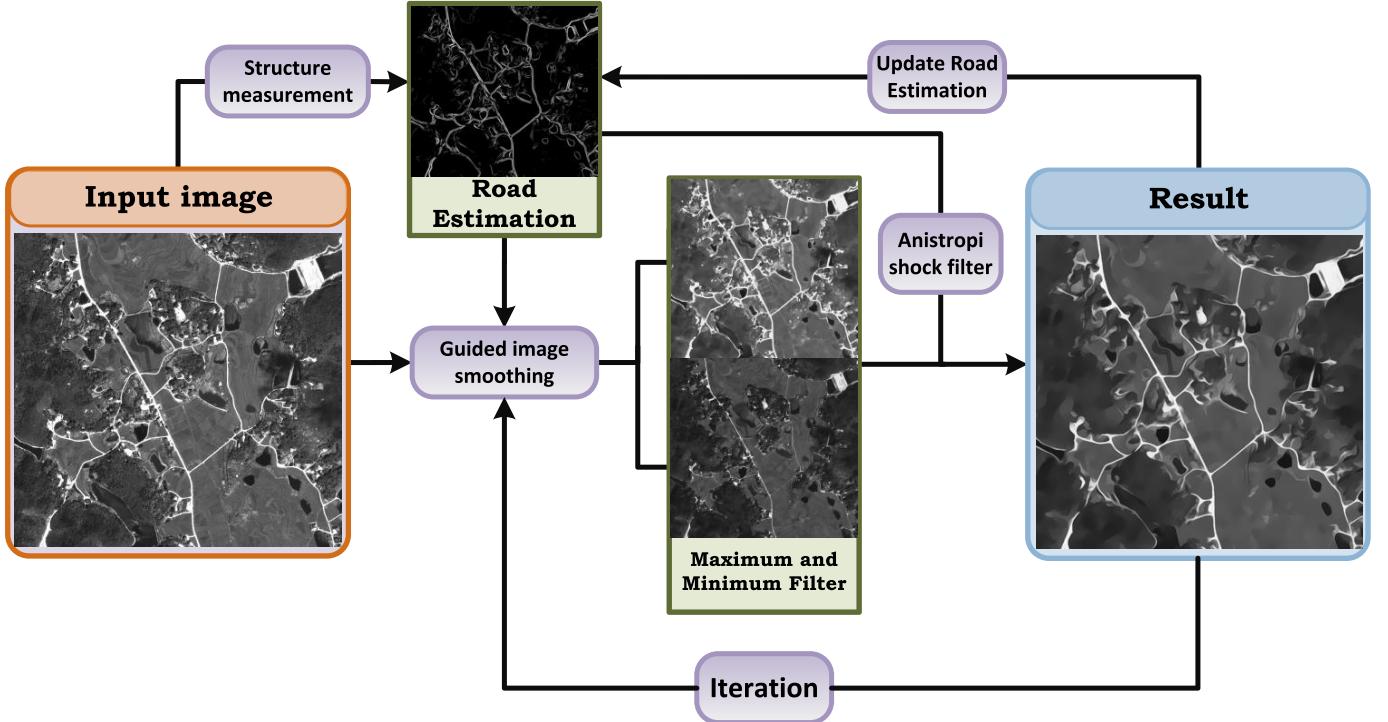


Fig. 1. Framework of our approach.

on the “maximum envelope” and “minimum envelope” of the image (which are approximated by the maximum and minimum filters in our approach). The results are then combined by the proposed anisotropic shock filter scheme to further enhance the road areas. The filtering results can then be employed to generate more accurate road structure estimations, thus leading to a better performance in an iterative manner.

A. Adaptive Image Smoothing Guided by Road-Like Structures

1) *Definition*: In most studies on road detection [2], [3], the image is filtered following a general weighted averaging framework, i.e., the value of a filtered pixel is calculated by averaging its neighbors:

$$I'_p = \int_{q \in N(p)} W_{p,q}(I) I_q dq \quad (1)$$

where p and q are pixel indexes, $N(p)$ is the neighborhood of pixel p , the weight function $W_{p,q}(I)$ is related to the input I , and different forms of $W_{p,q}(I)$ provide different filtering performances. Fig. 2(a)–(c) shows some examples. Fig. 2(a) is the input from Geoeye satellite, which includes heavy high contrast textures of the buildings; Fig. 2(b) is the result of the classic Gaussian filter, where different structures are smoothed to the same degree; and Fig. 2(c) is the result of the most recently developed covariance filter [7], where the textures are not sufficiently smoothed and the road edges are not preserved well.

On the other hand, for road extraction, current schemes [2]–[4] frequently provide approximate estimations or feature descriptors of road-like structures, which are very

valuable for the preprocessing. Based on this observation, inspired by the idea of guided image filtering [35], the objective of the proposed adaptive smoothing technique is to smooth the image by the guidance of the road-like structures

$$I'_p = \int_{q \in N(p)} W_{p,q}(I, G) I_q dq \quad (2)$$

where $W_{p,q}(I, G)$ is a joint function composed of not only the input image, but also explicit guidance.

2) *Guidance Construction*: Given the input image I , the guidance map of our approach, denoted by G , is based on the aperiodic directional structure measurement (ADSM) of I (the details of the construction of this measurement are given in [4]). Specifically, the magnitude of each point (corresponding to the pixels in the image) in the map, which provides an intuitive description of the probability that each pixel belongs to the road edges, is calculated by rescaling the ADSM value to the range $[0, 1]$. The direction of each point is defined by the direction of the minimum eigenvector of the structure tensor centered at the corresponding pixel in image I , denoted by ξ_p .

Fig. 2(d) shows the visualized result of the magnitude of the guidance map, which approximately represents the potential road edges. Fig. 2(e) visualizes the direction field of the guidance map generated using the line integral convolution technique [36]. It can be seen that the road structures show better consistency than do the textures.

3) *Adaptive Image Smoothing*: With this guidance, our adaptive smoothing process can be formulated as

$$I'_p = \frac{1}{K} \int_{q \in N(p)} w_{p,q}^g(\sigma_g, I) g(\sigma_d, G) I_q dq. \quad (3)$$

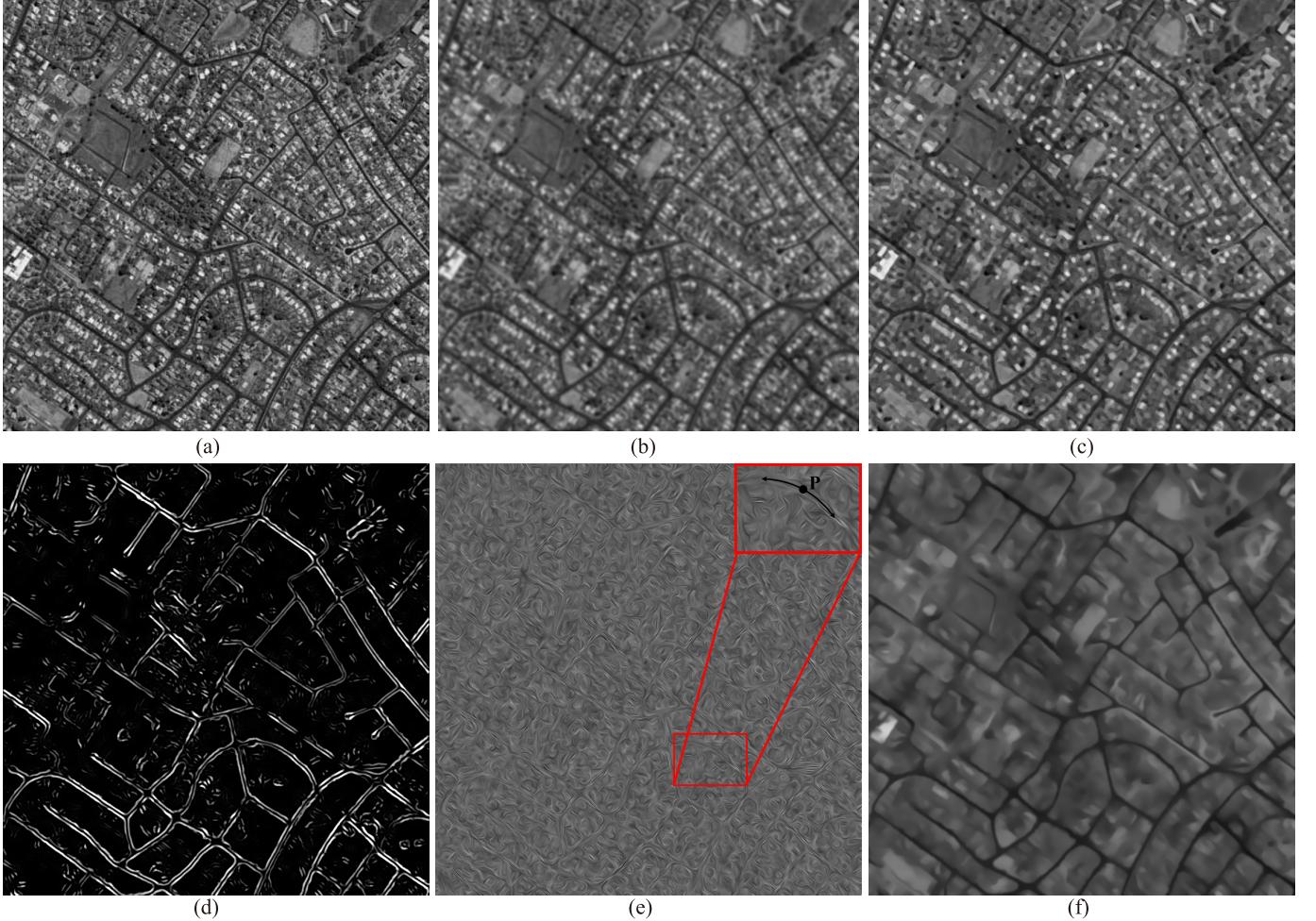


Fig. 2. Test examples of our approach. (a) Input from Geoeye satellite. (b) Result of classic Gaussian filter. (c) Result of the latest covariance filter. (d) Visualized result of the magnitude of the guidance map. (e) Visualized direction field of the guidance map. (f) Adaptive smoothing result of our approach.

Here K is the normalizing factor to guarantee the sum of the weight equals 1. $w_{p,q}^g(\sigma_g, I)$ is the Gaussian kernel, which encourages neighbors that are closer to the center pixel p to have larger weights, and is defined as

$$w_{p,q}^g(\sigma_g, I) = \exp\left(-\frac{|x_p - x_q|^2}{2\sigma_g^2}\right) \quad (4)$$

where x_p and x_q denote the coordinates of pixels p and q , respectively, and σ_g controls the shape of the filter kernel. $g(\sigma_d, G)$ is a function of the guidance G , and is formulated as

$$g(\sigma_d, G) = \exp\left(-\frac{|\langle \vec{\xi}_p \times \vec{\xi}_q \rangle| \cdot |(G_p - G_q)|^2}{2\sigma_d^2}\right) \quad (5)$$

where G_p and G_q denote the magnitude of the guidance map, and $\vec{\xi}_p$ and $\vec{\xi}_q$ are the minimum eigenvectors of the structure tensor centered at pixels p and q , respectively. The magnitude of the cross product $|\langle \vec{\xi}_p \times \vec{\xi}_q \rangle|$, with range $[0, 1]$, decreases when two vectors align closely, i.e., the included angle between them approaches 0 or π , and equals 1 as the two vectors become orthogonal. This kernel is designed to encourage neighbors that are similar (i.e., with similar road probabilities and local directions) to the center pixel p

to have larger weights, and therefore, for the potential road structures, the filter kernel is more likely to follow consistent local directions, as shown in the zoomed-in patch of Fig. 2(e). Fig. 2(f) shows our adaptive smoothing result, in which the heavy textures are sufficiently smoothed, while the road-like structures are well preserved.

B. Structure-Aware Enhancement via Anisotropic Shock Filter

The adaptive smoothing is aggressive, as shown in Fig. 2(f), and the heavy high contrast textures are effectively smoothed. To obtain the sharp edges, we propose an anisotropic shock filter for deblurring of road-like structures. The earliest *shock filter* is proposed in [37] as

$$\mathbf{I}_t = -\text{sign}(\Delta \mathbf{I}) |\nabla \mathbf{I}| \quad (6)$$

where t counts the iterative evolution in the filtering process, with initial condition $\mathbf{I}_0(x, y)$ equal to the initial image $\mathbf{I}(x, y)$. The underlying structures are sharpened at the zero crossings of $\Delta \mathbf{I}$. As an improvement in the adaptivity of this shock filter, Weickert [38] proposed a coherence-enhancing shock filter, which replaces the Laplacian operator $\Delta \mathbf{I}$ by the

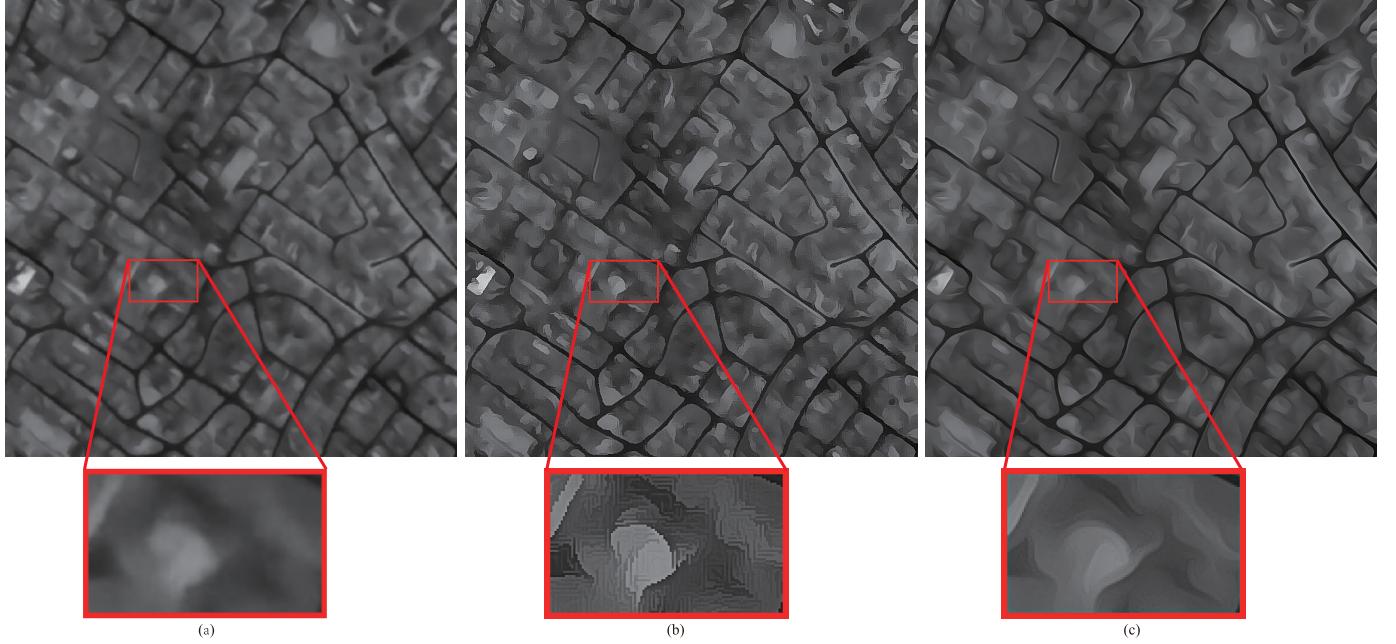


Fig. 3. Enhancing result of our approach. (a) Adaptive smoothing result. (b) Result of traditional shock filter. (c) Our enhancing result.

second derivative along the direction of the domain eigenvector of the structure tensor $\tilde{\xi} = (c, s)$ as

$$\mathbf{I}_t = -\text{sign}(c^2 \mathbf{I}_{xx} + 2cs \mathbf{I}_{xy} + s^2 \mathbf{I}_{yy}) |\nabla \mathbf{I}|. \quad (7)$$

The essence of the shock filter is that it creates isotropic shocks to enhance the regions that are selected by a certain edge detector (determined by the $\text{sign}(\cdot)$ function). A limitation of such a filter is that shocks may be constructed in texture or almost smooth regions, resulting in maze-like artifacts [39] [see the zoomed-in patch of Fig. 3(a)]. Drawing up underpins of the shock filter, our approach is aimed to add extra guidance to facilitate the creation of shocks with different degrees, i.e., anisotropic shocks. Specifically, this scheme can be formulated as

$$\mathbf{I}_t = f(G, \mathcal{E}) |\nabla \mathbf{I}|. \quad (8)$$

Guided by the potential road structures estimation G , the function $f(\cdot)$ indicates not only where to create shocks but also what degree of shocks should be created.

As noted in [40], the partial differential equation shown in 6 can be approximated by combining the erosion and dilation operator [depending on the $\text{sign}(\cdot)$ function], and therefore, our anisotropic shock filter is defined as

$$I' = W(G_p, \mathcal{E}) D(I_p) + (1 - W(G_p, \mathcal{E})) E(I_p) \quad (9)$$

where the adaptive smoothing result I is acquired as described in Section III-A, $D(I)$ and $E(I)$ denote the dilation and erosion versions of image I , respectively, and \mathcal{E} is the edge detector inherited from Weickert's shock filter, calculated as $\mathcal{E} = c^2 \mathbf{I}_{xx} + 2cs \mathbf{I}_{xy} + s^2 \mathbf{I}_{yy}$. Using the guidance map G , the degrees of shocks are measured by the weight function $[W(G_p, \mathcal{E})]$, which is essentially based on the hyperbolic

tangent function

$$W(G_p, \mathcal{E}) = T \left(\frac{1 - G_p \cdot \text{sign}(\mathcal{E})}{2} \right)$$

and

$$T(x) = \frac{(1 + \tanh(\lambda(x - 0.5)))}{2} \quad (10)$$

where λ is a factor to control the enhancing degree, with a larger value indicating sharper edges. In our experiment, value of λ is set at 6.0, while it can be adaptively adjusted for different cases in the typical range [2.0, 10.0]. The goal of our scheme is to encourage the enhancement of potential road-like structures while suppressing other regions, such as texture or flatter structures. Specifically, for road-like structures, i.e., $G_p \rightarrow 1$, (9) is equivalent to

$$I = \begin{cases} D(I) & \text{if } \mathcal{E} < 0 \\ E(I) & \text{if } \mathcal{E} > 0 \\ (D(I) + E(I))/2 & \text{if } \mathcal{E} = 0. \end{cases} \quad (11)$$

Therefore, shocks are created to enhance the structures, While for regions with $G_p \rightarrow 0$, (9) is degraded to $I = (D(I) + E(I))/2$, and no enhancement is applied.

Fig. 3 shows the enhancing result of our approach. Fig. 3(a) is the adaptive smoothing result. Fig. 3(b) is the result of the traditional shock filter; as shown in the highlighted patch, complex maze-like artifacts appear in the smooth region. Fig. 3(c) shows our enhancing result; as viewed in the zoomed-in patch, the road-like structures are significantly enhanced, while the artifacts are effectively removed.

IV. DISCUSSION

A. Iterative Scheme

The entire system is implemented in an iterative scheme, and the iterations of the structure estimation and filtering

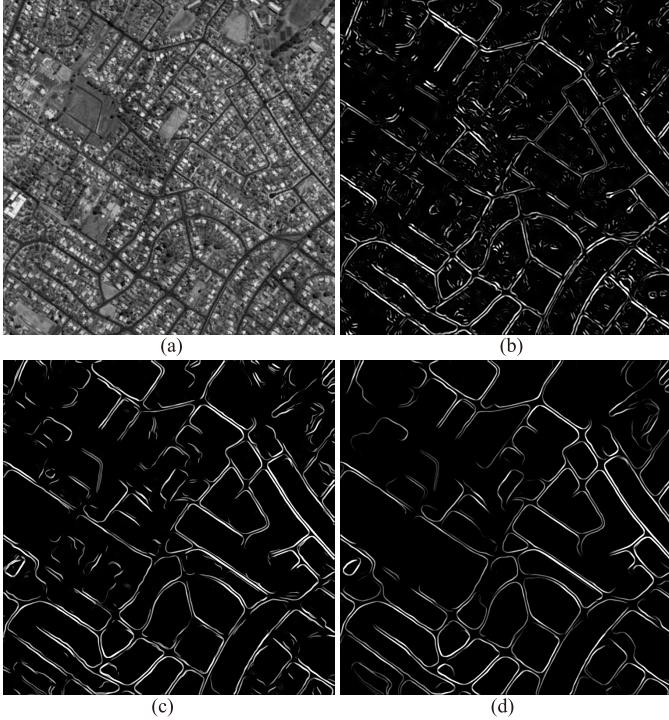


Fig. 4. Road structure estimations with different iteration times. (a) Input with heavy textures. (b) and (c) Road structure estimations after one to three iterations of our approach.

process are separate. First, the enhancing results can be in turn applied to generate more accurate road structure estimation, as shown in Fig. 4, where Fig. 4(a) is the input with heavy textures and Fig. 4(b) and (c) are road structure estimations generated by the results after one to three iterations. As the iteration progress, the textures are better suppressed and the estimation is more accurate.

The iteration of the structure estimation is independent of the filtering process, i.e., we first obtain a satisfactory road structure estimation by several iterations (typically the iteration number is set at 2 in our experiments), and then, use this estimation together with the original input to obtain the final enhancing result.

B. Parameter Settings

Most of the user specified parameters are set at empirical values. Specifically, for the smoothing phase, with an input image resolution of 4 m, the radius of $N(p)$ in 3 is set at 10, which is demonstrated sufficient large for most of the cases at this resolution. As the increase of the image resolution, the size of $N(p)$ could decrease correspondingly. σ_g in 4 is set at 5.0. For the enhancing phase, λ in 10 is set at 6.0.

The major parameters to control the effect are the iteration number of the filtering process and σ_d in 5. For the iterations, our results show that one to three times are sufficient for most cases. For the parameter σ_d , the typical values fall in the range [5, 20], which is test appropriate for most of the cases at resolution 4 m. We explored the output space of the proposed approach; the corresponding results are shown in Fig. 5. Here, we show examples of three typical values of the iteration

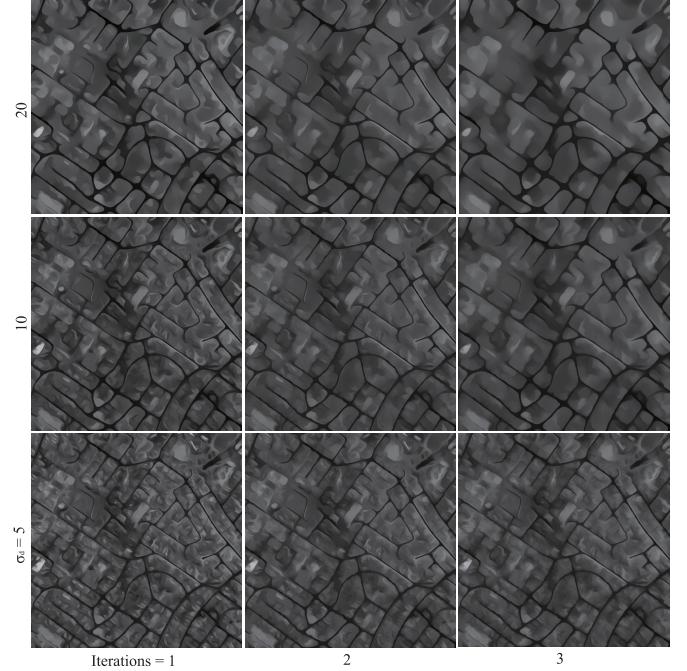


Fig. 5. Results of our approach with different parameters.

number and σ_d , and list nine images to show the manner in which these two parameters affect the results. As the number of iterations and σ_d increase, the results become more smooth with fewer details, which implies higher *correctness*. However, when the values of these parameters are set extremely high, the results may suffer slightly decrease of *completeness* in the following road extraction phase. As a result, to trade off these two statistics of the results, settings of two iterations together with $\sigma_d \in [10, 20]$ are applied most frequently in our experiments.

C. Time Performance

We implemented our approach based on C++ programming on a PC with a 3.2-GHz Intel Core i5-3200 CPU and 4-GB RAM. The algorithm is very efficient and with linear complexity $O(kn)$ corresponds to the total number of pixels. Specifically, for a million pixel image (size of 1000×1000), the entire process of our approach takes about 5 s. On the other hand, the system is highly parallelized, and the performance on a quad-core PC increases approximately 250%.

V. RESULTS

Three groups of experiments conducted to evaluate the algorithm comprehensively are described in this section. Specifically, we first present the filtering results of experiments based on inputs from different sensors, such as GeoEye, Ikonos, Pleiades satellites, and aerial image, where various terrains, such as urban areas, rural regions and mountains, are involved. On the other hand, we also evaluate the performance and parameter setting of our system on images with different resolutions. In Section V-B, we compare our approach with the most recent structure-aware image filtering techniques

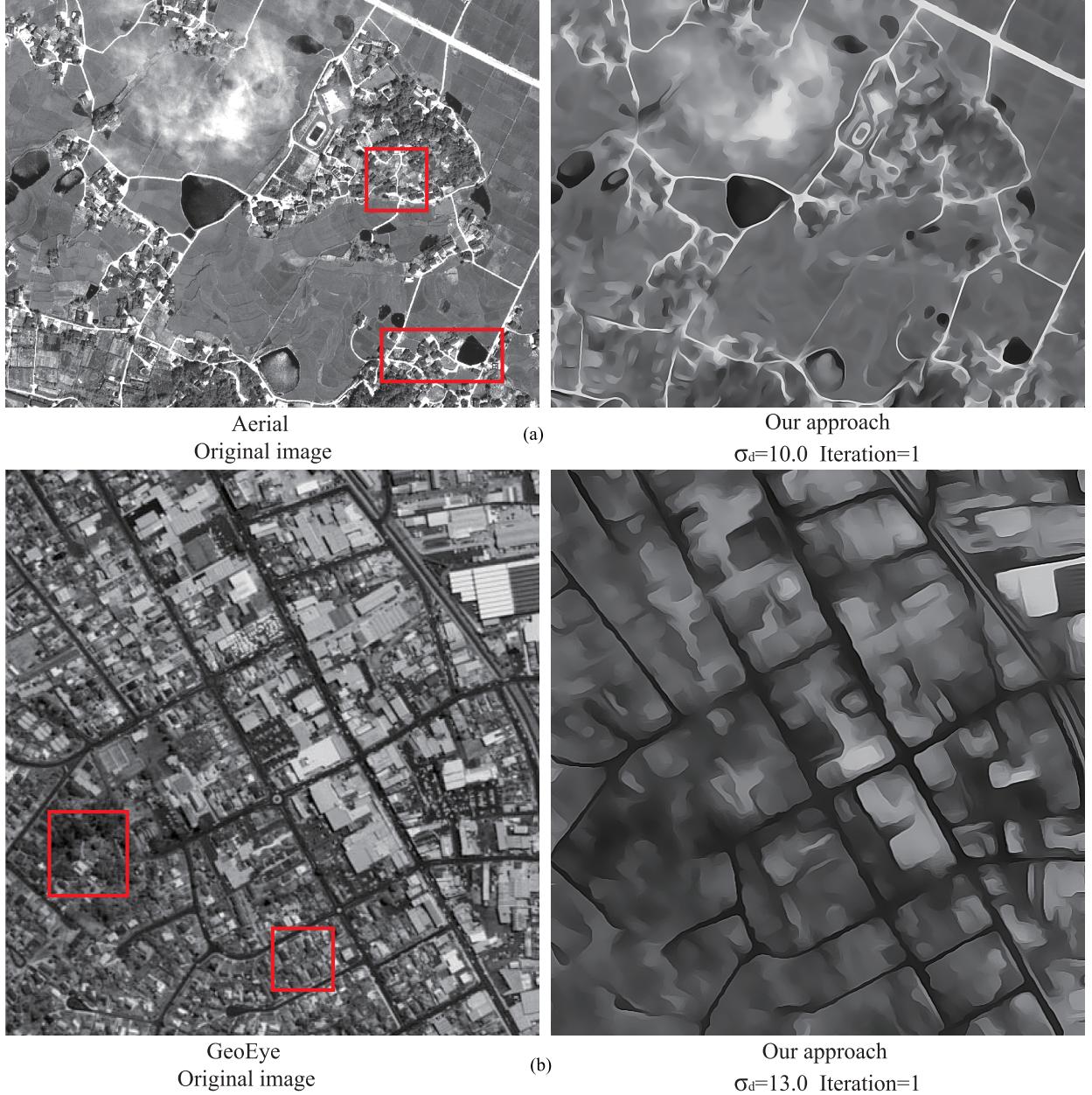


Fig. 6. Results of various sensors. (a) Result of aerial image. (b) Result of Geoeye image.

to demonstrate the effectiveness of our approach. Then, in Section V-C, experimental quantitative evaluations are provided to show how our approach affects current road network extraction methods.

A. Evaluation Using Images From Various Sensors

To test the adaptiveness of our approach, we first conducted an evaluation in which our method is applied to images from four different sensors, consisting of aerial image, and those from the Geoeye, Ikonos, and Pleiades satellites. The resolution of the Ikonos image is 1 m, and the resolution of the rest images is 0.5 m. Various terrains, such as urban, and rural region, and mountain, are included. The corresponding results are presented in Figs. 6 and 7. For the images from different

sensors, the potential interferences of road extraction are sufficiently smoothed, while the road edges are well preserved. It is also viewed that, despite the fact that the intensities of the road regions in different images are various, the filtering results are similar.

Specifically, in the case of the aerial image [Fig. 6(a)], which includes several villages, the major challenge is to preserve the narrow roads mixed in the residential area while smoothing the buildings and trees, as shown in the highlighted patch. In the city regions in the Geoeye image, heavy textures are involved in both the green belt and the residential area. These textures are sufficiently smoothed in the result of our approach, as shown in the highlighted patch. For the Ikonos image, the low spectral contrast roads (as shown in the highlight area), which can easily be destroyed by previous

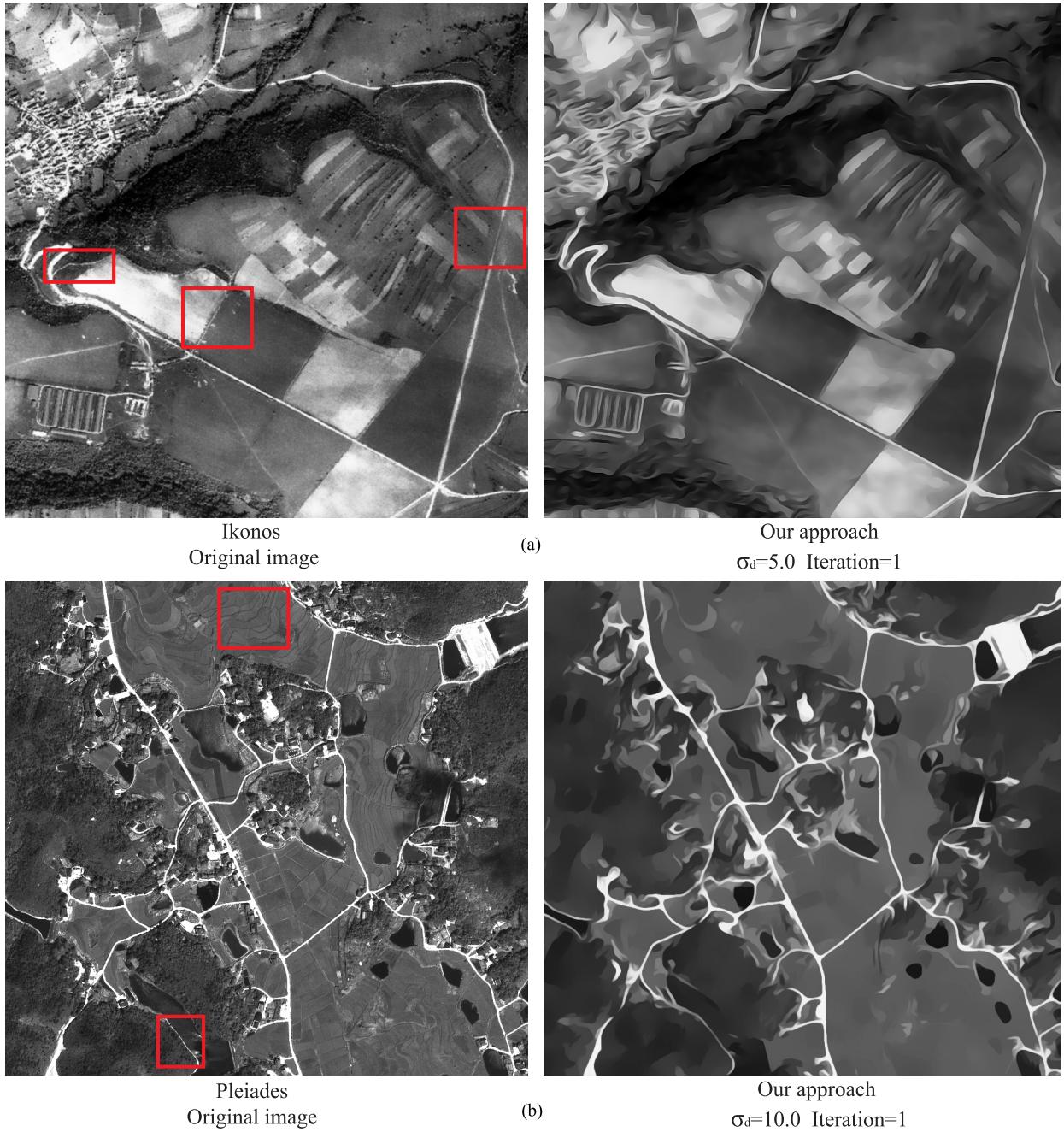


Fig. 7. Results of various sensors. (a) Result of Ikonos image. (b) Result of Pleiades image.

filters, such as the Gaussian or BLF, are well preserved or enhanced in our result. For the Pleiades image, the boundaries of terrains and the narrow roads occluded by trees, which may influence the quality of the road extraction result, are smoothed and connected by our approach.

We also evaluate the performance of our approach on images with different resolutions. As shown in Fig. 8, images in the first row (from left to right) are taken from ZiYuan3 (with resolution 0.5 m), GaoFen3 (with resolution 1 m), and Pleiades 1A (with resolution 0.5 m) satellites. The sizes of these images are $15\,872 \times 18\,653$, $23\,609 \times 31\,899$, and $28\,648 \times 37\,929$, respectively. Images from the second to the

fourth rows are three examples corresponding to the ZiYuan3, GaoFen3, and Pleiades 1A satellites. Here, the image from Pleiades satellite are downsampled to 4 m to cover different resolutions in the experiment.

In the first experiment, we aim to evaluate the performance of the road detection method [4] on the inputs with (denoted by "Y" in Table I) or without (denoted by "N" in Table I) our filter as preprocessing. Specifically, three types of results are presented, including the filtering results (with the parameter settings $\sigma_d = 25.0$, Iterations = 2 for resolution 0.5 m; $\sigma_d = 20.0$, Iterations = 1 for resolution 1 m; and $\sigma_d = 10.0$, Iterations = 1 for resolution 4 m), as shown in the

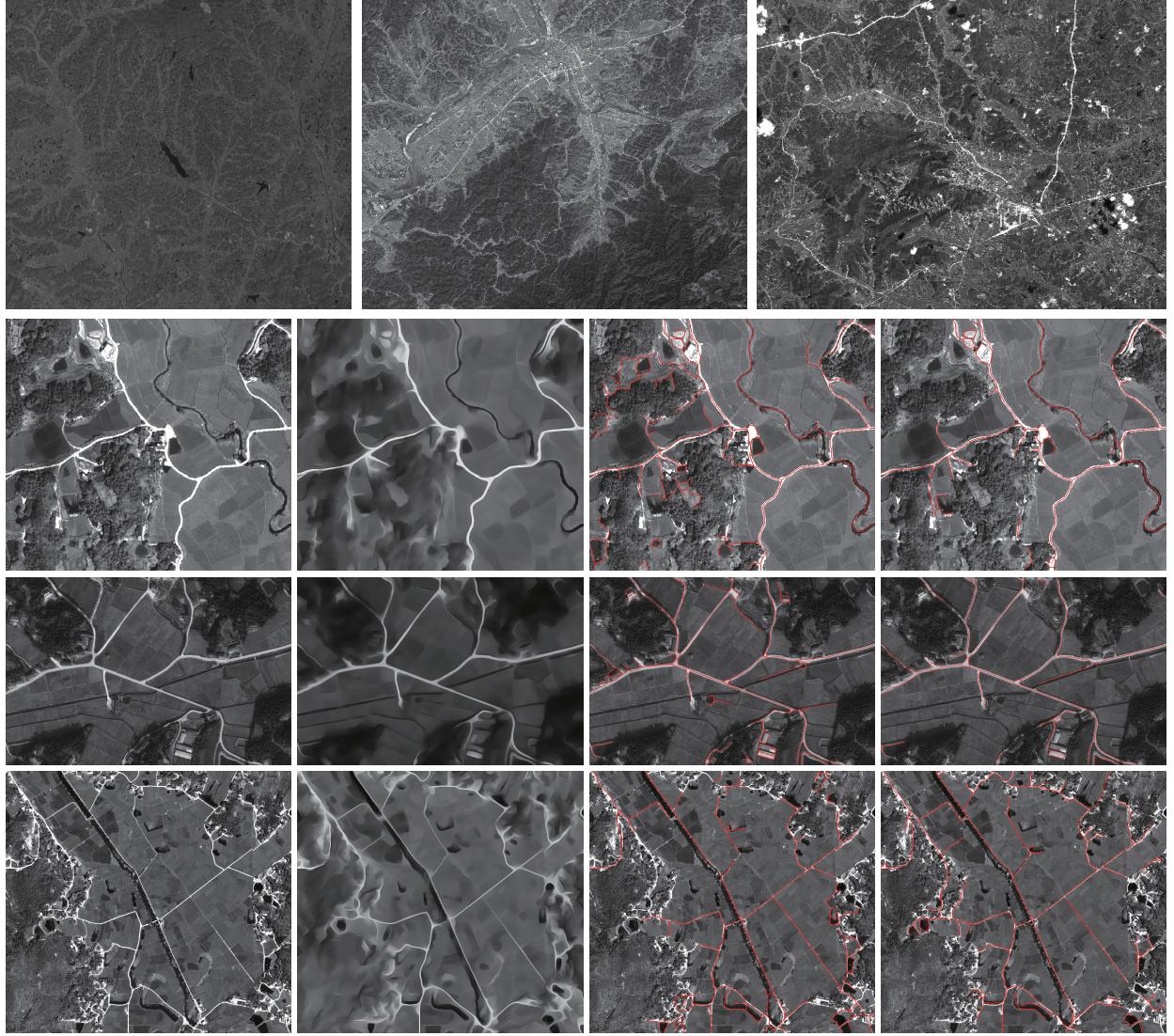


Fig. 8. Results of inputs with different resolutions. The first row (from left to right) is input images with resolutions 0.5, 1, and 4 m (which is generated by downsampling the original 0.5-m image). Images in the first column of rows 2–4 are the typical samples corresponding to images with resolutions 0.5, 1, and 4 m. Images in the second column are the filtering results of our approach. Road detection results of [4] on the original input and our filtering result are shown in the third and fourth columns. The road networks are highlighted as red for visualization.

TABLE I
CORRESPONDING STATISTICS OF IMAGES WITH VARIOUS RESOLUTIONS
BY COMBINING WITH OUR FILTER (DENOTED BY Y) OR NOT
(DENOTED BY N)

Inputs	Comp		Corr		Quality	
	N	Y	N	Y	N	Y
ZiYuan_sample	0.8592	0.8435	0.7642	0.8387	0.6791	0.7257
GaoFen_sample	0.8849	0.8806	0.8153	0.8862	0.7371	0.7911
Pleiades_sample	0.8345	0.8401	0.8057	0.8698	0.6946	0.7462
ZiYuan_all	0.8122	0.8098	0.7008	0.7812	0.6030	0.6600
GaoFen_all	0.7965	0.7997	0.7049	0.8107	0.5973	0.6738
Pleiades_all	0.7786	0.7831	0.7135	0.8232	0.5931	0.6703

second column; road detection results of [4] on the original input image, as shown in the third column; and road detection results of [4] on our filtering result, as shown in the last column. The road networks are highlighted as red in the input image.

Table I shows some quantified results, where both the samples (which are denoted by XX_sample) and the whole image (which are denoted by XX_all) are applied for testing. The reference road network is obtained by ground survey. It can be viewed that the *correctness* is significantly improved for either the samples or the whole image, especially for the resolution 1 and 4 m, where 15.01% and 15.37% improvements are observed (for the whole image testing). On the other hand, for the *completeness*, slightly decrease (about 0.29%) is observed for resolution 0.5 m. While for resolutions 1 and 4 m, due to the enhancing scheme of our approach, such values even receive small increasing, about 0.40% and 0.57% improvements for 1 and 4 m are observed. The reason we guess is that some of the occlusions are connected by our filtering process.

On the other hand, we also evaluate how the parameters affect the results for images with different resolutions.

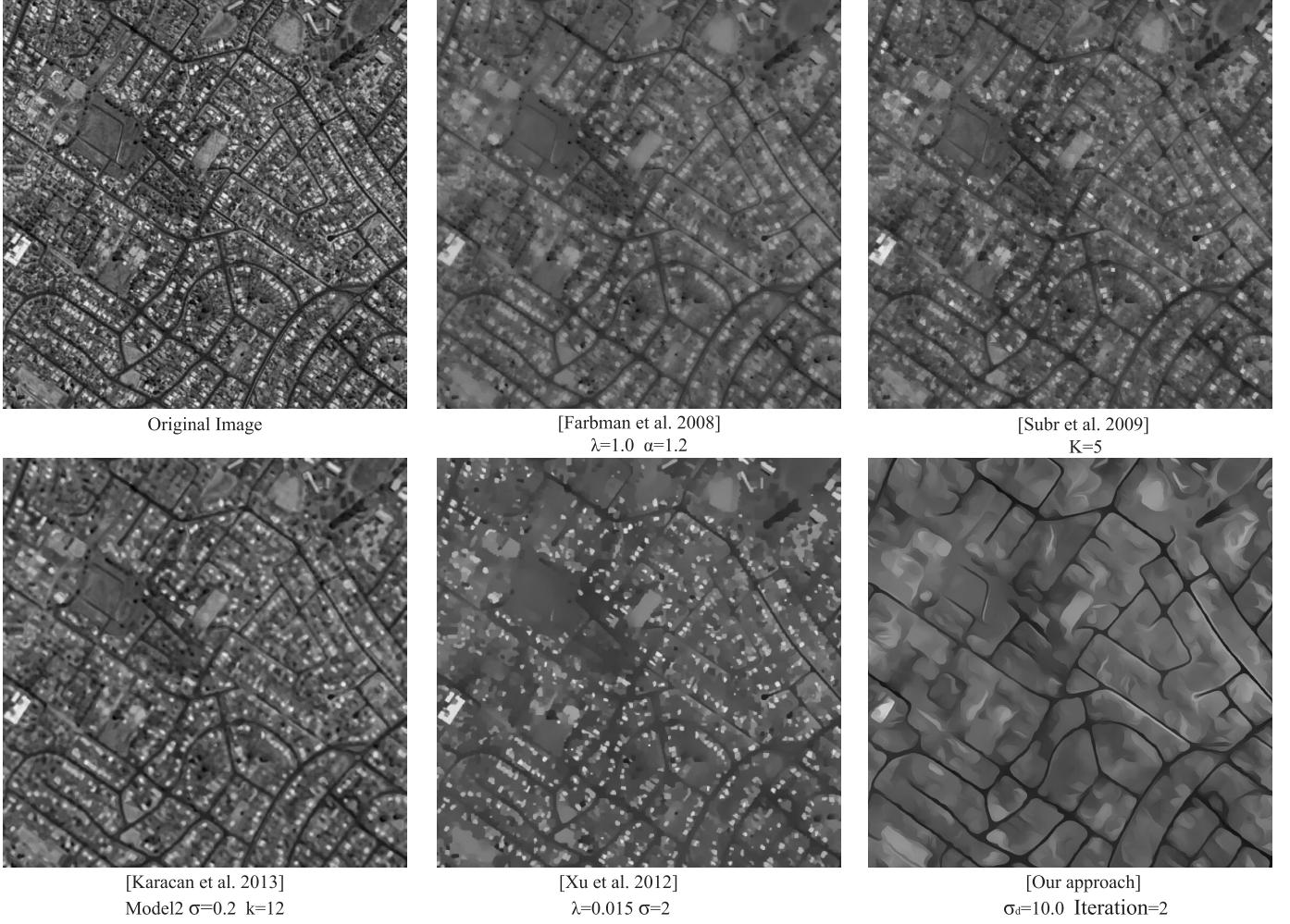


Fig. 9. Comparison with state-of-the-art structure-aware image smoothing methods using Geoeye image.

TABLE II
CORRESPONDING STATISTICS OF RESULTS WITH
VARIOUS PARAMETER SETTINGS

Iter = 1	$\sigma_d = 10.0$		$\sigma_d = 20.0$		$\sigma_d = 25.0$	
	Comp	Corr	Comp	Corr	Comp	Corr
ZiYuan	0.8183	0.7427	0.8149	0.7538	0.8116	0.7607
GaoFen	0.8014	0.7839	0.7997	0.8107	0.7728	0.8182
Pleiades	0.7831	0.8232	0.7526	0.8316	0.7344	0.8392

Here the whole images taken from ZiYuan3, GaoFen3, and Pleiades 1A satellites along with six groups of parameter settings are applied for testing, also the image from Pleiades satellite are downsampled to 4 m, and the ground truths are taken from ground survey.

The corresponding results are shown in Tables II and III; it can be viewed that, when the iterations = 1, as the value of σ_d increases, the *completeness* suffers just slightly decrease for resolutions 0.5 and 1 m; however, for images with resolution 4 m, the *completeness* significantly decreases when the value of σ_d changes from 20.0 to 25.0. For the *correctness*, as σ_d changes from 20.0 to 25.0, such a measurement for resolutions 1 and 4 m tends to be convergent, while for images with resolution 0.5 m, such a measurement increases continuously.

TABLE III
CORRESPONDING STATISTICS OF RESULTS WITH
VARIOUS PARAMETER SETTINGS

Iter = 2	$\sigma_d = 10.0$		$\sigma_d = 20.0$		$\sigma_d = 25.0$	
	Comp	Corr	Comp	Corr	Comp	Corr
ZiYuan	0.8125	0.7596	0.8107	0.7773	0.8098	0.7812
GaoFen	0.7721	0.7927	0.7509	0.8235	0.7397	0.8284
Pleiades	0.7439	0.8304	0.7254	0.8395	0.7088	0.8531

Then, when applying two iterations, i.e., iterations = 2, the *completeness* suffers apparently decreasing as σ_d increases for images with resolutions 1 and 4 m. For images with resolution 0.5 m, as the increase of σ_d , the *completeness* does not change much, while the *correctness* improves stably.

B. Comparisons With Typical Structure-Aware Image Filters

In this experiment, we conducted an extensive comparison with previous state-of-the-art structure-aware image smoothing methods, including those presented in [5]–[7] and [31]. Two different types of images are used for testing: the Geoeye image with heavy textures and the Pleiades image with occlusions and shadows, as shown in Figs. 9 and 10. In the Geoeye

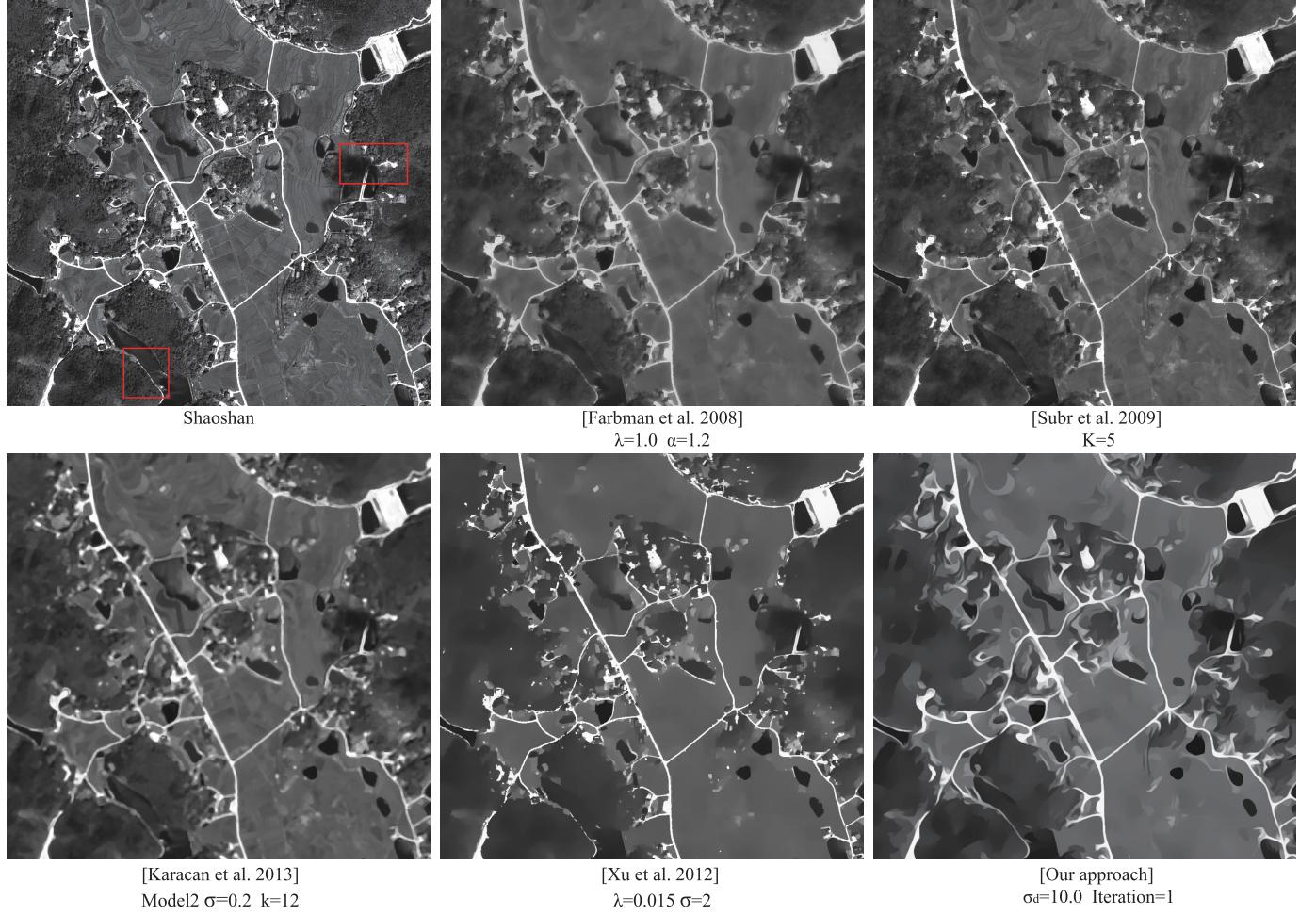


Fig. 10. Comparison with state-of-the-art structure-aware image smoothing methods using Pleiades image.

image, a large number of high contrast buildings may lead to a high false positive rate for road extractions, and previous filters either destroy the road edges [5], [6] or cannot easily address these heavy textures [5]–[7], [30]. Our result, in contrast, shows the ability to smooth these textures while enhancing the road edges. The Pleiades image is not relatively challenging; however, as shown in the highlighted area, occlusions and shadows occur in some of the road regions. In the results presented in [30] and [7], the roads are to some degree blurred, and the occlusions are not well addressed in the results presented in [5], [30], and [6]. Our approach performs well when occlusions are present; however, like other methods, it fails to handle the cases where shadows occur (as shown in the red rectangle in the right). For these cases, in future work, the effectiveness of applying semantic information could be investigated.

C. Integration With Most Recent Road Network Extraction Approaches

The goal of the experiments described in this section is to quantitatively evaluate the effectiveness of our approach. First, the Geoeye image shown in Fig. 9(a), which is challenged for various road extraction methods, is selected as the input.

The three most recent road extraction methods are implemented for testing. Fig. 11 shows the corresponding results. Fig. 11(a), (c), and (e) shows the road extraction results of the methods in [2]–[4] separately. Due to the graph optimization, the method in [2] is able to provide impressive detection *completeness*; in contrast, by conducting a multistage false filtering scheme, the results of [3] present high detection *correctness*; the method in [4], by focusing on the structure of potential road regions, can efficiently generate high-*quality* detection results. Fig. 11(b), (d), and (f) shows the results by combining the proposed joint enhancing filtering.

To quantitatively evaluate the extraction results, ground truth data are manually generated, and classic measurements criteria are applied [41], [42]

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

where TP, FN, and FP denote true positive, false negative, and false positive (corresponding to the green, red, and blue lines in the images), respectively. Table IV shows the corresponding

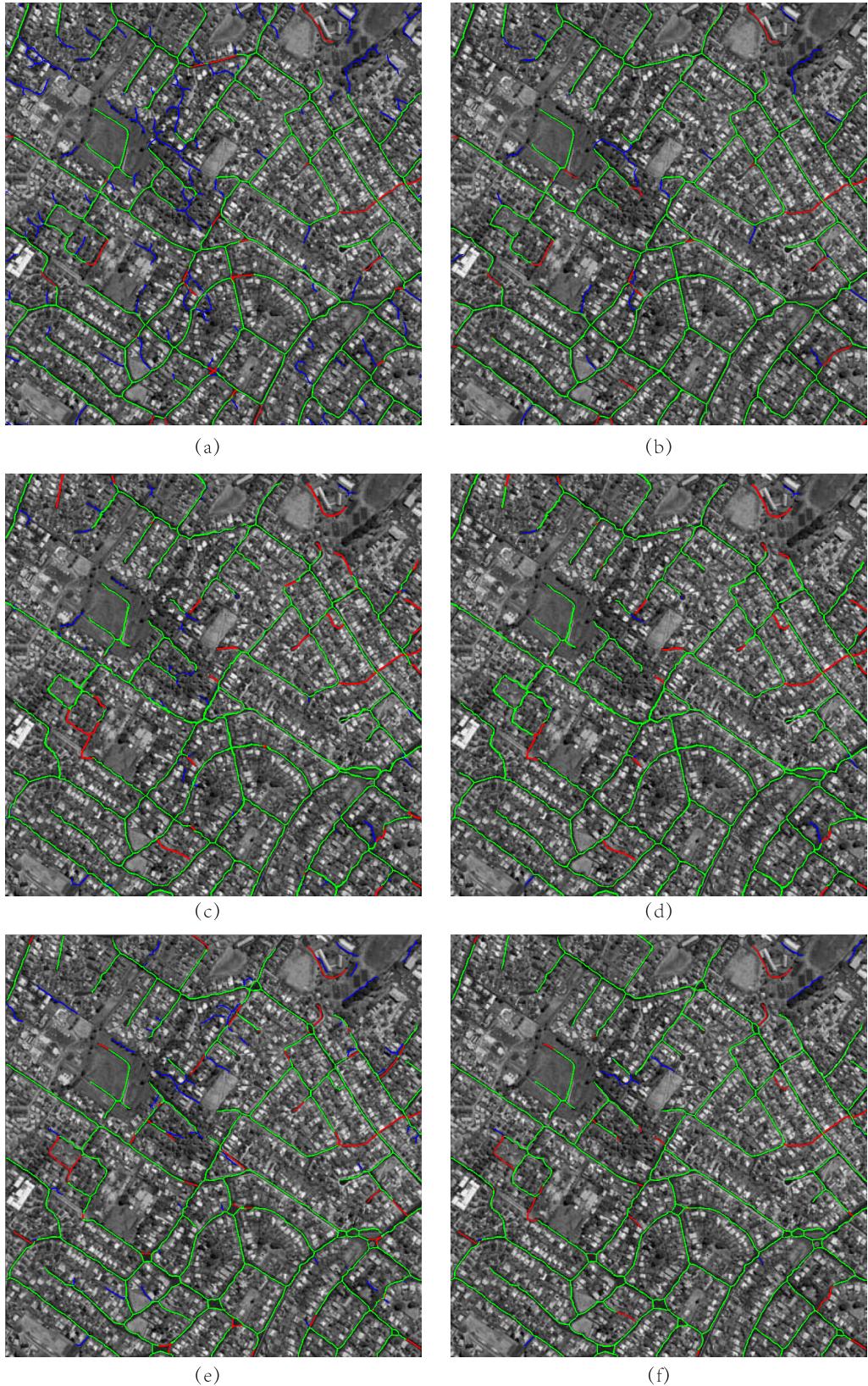


Fig. 11. Road extraction results by combining our approach and current road extraction methods. (a), (c), and (e) Road extraction results of [2]–[4] separately. (b), (d), and (f) Results combining the proposed joint enhancing filtering.

statistics of the various methods with (denoted by “Y”) and without (denoted by “N”) our filter. Specifically, the *completeness* measurements of the methods in [2]–[4] increase

by approximately 0.35%, 3.78%, and 3.14% respectively, which is a fair result, while for the performance of *correctness*, a significant improvement is observed (19.97%, 5.92%,

TABLE IV

CORRESPONDING STATISTICS OF VARIOUS ROAD EXTRACTION METHODS COMBINING WITH OUR FILTER (DENOTED BY Y) OR NOT (DENOTED BY N)

Methods	Completeness		Correctness		Quality	
	N	Y	N	Y	N	Y
Ünsalan [2]	0.8991	0.9022	0.7417	0.8898	0.6847	0.8115
Shi [3]	0.7904	0.8203	0.8992	0.9524	0.7261	0.7880
Zang [4]	0.8522	0.8790	0.8538	0.9395	0.7437	0.8320

TABLE V

CORRESPONDING STATISTICS OF [2] ALONG WITH CLASSIC STRUCTURE-AWARE FILTERING TECHNIQUES: BLF [25], RTV [6], RCF [7], AND OUR APPROACH

Ünsalan [2]	Completeness	Correctness	Quality
Ori	0.8027	0.6633	0.5703
BF [26]	0.7961	0.6802	0.5793
RTV [6]	0.7901	0.7103	0.5975
RCF [7]	0.7623	0.6937	0.5703
Our approach	0.8019	0.7891	0.6604

TABLE VI

CORRESPONDING STATISTICS OF [3] ALONG WITH CLASSIC STRUCTURE-AWARE FILTERING TECHNIQUES: BLF [25], RTV [6], RCF [7], AND OUR APPROACH

Shi [3]	Completeness	Correctness	Quality
Ori	0.7303	0.7694	0.5992
BF [26]	0.7397	0.7723	0.6073
RTV [6]	0.7277	0.7938	0.6120
RCF [7]	0.7005	0.7745	0.5818
Our approach	0.7387	0.8526	0.6550

TABLE VII

CORRESPONDING STATISTICS OF [4] ALONG WITH CLASSIC STRUCTURE-AWARE FILTERING TECHNIQUES: BLF [25], RTV [6], RCF [7], AND OUR APPROACH

Zang [4]	Completeness	Correctness	Quality
Ori	0.7786	0.7135	0.5931
BF [26]	N/A	N/A	N/A
RTV [6]	0.7693	0.7597	0.6187
RCF [7]	0.7507	0.7390	0.5934
Our approach	0.7831	0.8432	0.6832

and 10.04% for the methods in [2]–[4]). Benefiting from this, the overall road extraction *quality* of each method improved by 18.52%, 8.52%, and 11.87%, respectively.

The proposed approach is also evaluated using a remote sensing image, having a resolution 0.5 m, of Shaoshan City recorded by the Pleiades-1A satellite. The size of the entire image is 28 648 * 37 929, and the reference is acquired by ground survey. The experiment is similar to the evaluation above (Table IV), where previous road extraction methods (presented in [2]–[4]) together with classic structure-aware filtering techniques, BLF [25], RTV [6], and region covariances filter [7] (RCF), are employed for testing. The image is divided into 1000 * 1000 patches with 30% overlap. The average statistics are shown in Tables V–VII. The table shows the original results of the different road extraction methods,

and the remaining rows show the updated extraction results when the methods are combined with the different filters.

The results in [2] show high *completeness*. The different filters decrease this measurement to varying degrees, specifically, RCF leads to the greatest decrease, i.e., -5.03%, compared with the original results, while our approach provides the smallest decrease, -0.10%. On the other hand, for *correctness*, BF provides the smallest improvement (+2.55%), while our approach significantly increases this measurement (+18.97%) and the overall detection *quality* (+15.80%). Compared with the method in [3], BF and our approach provide positive gain in the *completeness*, while RTV and RCF slightly decrease this measurement. Then, for the *correctness*, each method provides different degrees of improvement: 0.38%, 3.17%, 0.66%, and 10.81% are observed for BF, RTV, RCF, and our approach, respectively. For the method in [4], since a BLF-based preprocessing step is involved in the original method, the row of “BF” is blank, and for the *completeness*, only our result provides a positive gain, +0.58%, and for the *correctness*, 6.48%, 3.57%, and 18.18% improvements are observed for RTV, RCF, and our approach, respectively.

VI. CONCLUSION

In this paper, we presented a joint filtering framework developed to generate a version of the input image that facilitates the road network extraction. Using the proposed adaptive smoothing scheme and anisotropic shock filter, our system is able to smooth heavy high contrast textures while keeping the structures of potential road edges sharp and clear. Abundant comparisons and quantitative evaluations using various satellite sensors demonstrated that the proposed approach can efficiently provide a meaningful improvement in road network extraction.

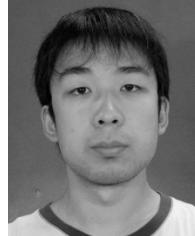
ACKNOWLEDGMENT

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

REFERENCES

- [1] 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.
- [2] C. Ünsalan and B. Sirmacek, “Road network detection using probabilistic and graph theoretical methods,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4441–4453, Nov. 2012.
- [3] 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.
- [4] Y. Zang, C. Wang, L. Cao, Y. Yu, and J. Li, “Road network extraction via aperiodic directional structure measurement,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3322–3335, Jun. 2016, doi: 10.1109/TGRS.2016.2514602.
- [5] F. Subr, C. Soler, and F. Durand, “Edge-preserving multiscale image decomposition based on local extrema,” *ACM Trans. Graph.*, vol. 28, no. 5, Dec. 2009, Art. no. 147.
- [6] L. Xu, Q. Yan, Y. Xia, and J. Jia, “Structure extraction from texture via relative total variation,” *ACM Trans. Graph.*, vol. 31, no. 6, Nov. 2012, Art. no. 139.
- [7] L. Karacan, E. Erdem, and A. Erdem, “Structure-preserving image smoothing via region covariances,” *ACM Trans. Graph.*, vol. 32, no. 6, 2013, Art. no. 176.

- [8] A. Baumgartner, C. Steger, H. Mayer, and W. E. L. Fur, "Multi-resolution, semantic objects, and context for road extraction," in *Proc. Semantic Modeling Acquisition Topogr. Inf. Images Maps*, 1997, pp. 140–156.
- [9] 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.
- [10] C. Ünsalan and K. L. Boyer, "A system to detect houses and residential street networks in multispectral satellite images," *Comput. Vis. Image Understand.*, vol. 98, no. 3, pp. 423–461, 2005.
- [11] 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.
- [12] C. Poullis and S. You, "Delineation and geometric modeling of road networks," *ISPRS J. Photogram. Remote Sens.*, vol. 65, no. 2, pp. 165–181, 2010.
- [13] A. Katartzis, H. Sahli, V. Pizurica, and J. Cornelis, "A model-based approach to the automatic extraction of linear features from airborne images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 9, pp. 2073–2079, Sep. 2001.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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, 2011.
- [18] C. Wiedemann and S. Hinz, "Automatic extraction and evaluation of road networks from satellite imagery," *Int. Arch. Photogram. Remote Sens.*, vol. 32, no. 3, pp. 95–100, 1999.
- [19] C. Wiedeman and H. Ebner, "Automatic completion and evaluation of road networks," in *Proc. Int. Arch. Photogram., Remote Sens.*, 2000, pp. 976–986.
- [20] S. Hinz and C. Wiedemann, "Increasing efficiency of road extraction by self-diagnosis," *Photogram. Eng. Remote Sens.*, vol. 70, no. 12, pp. 1457–1466, 2004.
- [21] J. Yang and R. S. Wang, "Classified road detection from satellite images based on perceptual organization," *Int. J. Remote Sens.*, vol. 28, no. 20, pp. 4653–4669, 2007.
- [22] B. Sirmacek and C. Ünsalan, "A probabilistic framework to detect buildings in aerial and satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 211–221, Jan. 2011.
- [23] J. Debayle and J. C. Pinoli, "General adaptive neighborhood image processing: Part I: Introduction and theoretical aspects," *J. Math. Imag. Vis.*, vol. 25, no. 2, pp. 245–266, Sep. 2006.
- [24] M. Ziems, U. Breitkopf, C. Heipke, and F. Rottensteiner, "Multiple-model based verification of road data," *ISPRS Ann. Photogram., Remote Sens. Spatial Inf. Sci.*, vol. I-3, pp. 329–334, Aug. 2012.
- [25] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
- [26] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, 2002.
- [27] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L_0 gradient minimization," *ACM Trans. Graph.*, vol. 30, no. 6, Dec. 2011, Art. no. 174.
- [28] E. S. L. Gastal and M. M. Oliveira, "Adaptive manifolds for real-time high-dimensional filtering," *ACM Trans. Graph.*, vol. 31, no. 4, Jul. 2012, Art. no. 33.
- [29] X. Y. Li, Y. Gu, S. M. Hu, and R. R. Martin, "Mixed-domain edge-aware image manipulation," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1915–1925, May 2013.
- [30] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 27, no. 3, Aug. 2008, Art. no. 67.
- [31] Y. Zang, H. Huang, and L. Zhang, "Efficient structure-aware image smoothing by local extrema on space-filling curve," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 9, pp. 1253–1265, Sep. 2014.
- [32] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [33] A. Buades, T. M. Le, J.-M. Morel, and L. A. Vese, "Fast cartoon + texture image filters," *IEEE Trans. Image Process.*, vol. 19, no. 8, pp. 1978–1986, Aug. 2010.
- [34] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A, Math. Phys. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
- [35] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [36] B. Cabral and L. C. Leedom, "Imaging vector fields using line integral convolution," in *Proc. ACM SIGGRAPH* Anaheim, CA, USA, vol. 27, Aug. 1993, pp. 263–270.
- [37] S. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM J. Numer. Anal.*, vol. 27, no. 4, pp. 919–940, 1990.
- [38] J. Weickert, "Coherence-enhancing shock filters," in *Pattern Recognition (Lecture Notes in Computer Science)*, vol. 2781. Sep. 2003, pp. 1–8.
- [39] J. E. Kyprianidis and H. Kang, "Image and video abstraction by coherence-enhancing filtering," *Comput. Graph. Forum*, vol. 30, no. 2, pp. 593–602, Apr. 2011.
- [40] F. Guichard and J.-M. Morel, "A note on two classical enhancement filters and their associated PDE's," *Int. J. Comput. Vis.*, vol. 52, no. 2, pp. 153–160, May 2003.
- [41] C. Wiedemann, C. Heipke, H. Mayer, and O. Jamet, "Empirical evaluation of automatically extracted road axes," in *Proc. Empirical Eval. Techn. Comput. Vis.*, 1998, pp. 172–187.
- [42] C. Wiedemann and H. Ebner, "Automatic completion and evaluation of road networks," in *Proc. Int. Arch. Photogram. Remote Sens.*, 2000, pp. 979–986.



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

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



Cheng Wang (M'07–SM'16) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2002.

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

Dr. Wang is the Co-Chair of ISPRS WG I/3, a Council Member of the Chinese Society of Image and Graphics, and a member of SPIE and IEEE GRSS.



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

Her research interests include remote sensing image processing and road network extraction.



Navigation.

Lun Luo received the master's degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 1999.

He is currently a Senior Engineer with the China Transport Telecommunications and Information Center, Beijing China. His current research interests include remote sensing image processing, InSAR signal processing, and multisensor fusion.

Mr. Luo is the Council Member of the Academic Committee on Remote Sensing, China Institute of



Ke Yang received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2015.

He is currently an Engineer with the China Transport Telecommunications and Information Center, Beijing China. His research interests include remote sensing image processing and big data analysis.



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

He is currently a Full Professor and the Head of the Mobile Sensing and Geodata Science Laboratory, Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON, Canada. He is also with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen, China. He has coauthored more than 300 publications, over 150 of which were published in refereed journals including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (TITS), the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTAR), the ISPRS Journal of Photogrammetry and Remote Sensing, the International Journal of Remote Sensing, Photogrammetric Engineering and Remote Sensing, and Remote Sensing Environment. His current research interests include information extraction from LiDAR point clouds and from earth observation images.

Dr. Li is the Chair of the ISPRS Working Group I/2 on LiDAR, Air- and Spaceborne Sensing Systems for 3-D Mapping (2016–2020), the Chair of the ICA Commission on Sensor-driven Mapping (2015–2019), an Associate Editor of the IEEE TITS and the IEEE JSTAR, and an Associate Editor-in-Chief of Sensors.