INTEGRATED SPECTRAL AND GEOMETRICAL INFORMATION FOR ROAD EXTRACTION FROM VHR SATELLITE IMAGES

MIAO ZELANG

Ph.D

The Hong Kong Polytechnic University

2016

The Hong Kong Polytechnic University
Department of Land Surveying and Geo-Informatics

Integrated Spectral and Geometrical Information for Road Extraction from VHR Satellite Images

MIAO Zelang

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

November 2015

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been awarded of any other degree or diploma, except where due acknowledgement has been made in the text.

Signed: _____

Name of student: MIAO Zelang

Abstract

With the advent of modern satellite sensors, it is possible to produce massive satellite images that provide rich land cover information. Object extraction, one of the fundamental images processing technologies, plays an important role to process these satellite images. With auxiliary of object extraction technologies, it is possible to generate useful information and knowledge from satellite images. Among various geo-information, road network extraction has been received much attention, partly due to its fundamental role in modern society and partly due to its challenging. Although this topic has been researched for decades, the results extracted are commonly unsatisfactory due to the extremely complicated natural scene and thus this topic is still not well resolved. Overall, road extraction from satellite images is still in its infancy and there is still large room to deep the insight into this topic.

An obvious trend of road extraction in the field of remote sensing is shifting from low/middle resolution satellite images to Very High Resolution (VHR) satellite images. Compared to low/middle spatial resolution satellite images, VHR satellite images can provide much more structural details of road feature. Naturally, road extraction from VHR satellite images can exploit spatial information in addition to spectral information, which is only available information source for low resolution satellite images. Road extraction from VHR satellite images has been receiving increasing attention in recent years.

This thesis focuses on the road delineation from Very High Resolution (VHR) satellite images by exploiting multiple road information, such as geometrical and spectral information. Specifically, several metrics are proposed to measure road geometrical characters in VHR satellite images. To make full use of available road information, a framework has been designed to combine the multiple information sources (i.e., geometrical and spectral features) to improve road extraction accuracy. A new method has been designed to shift traditional road extraction methods from pixel-based to object-based. By

doing so, it is convenient to measure road features at object level that in turn improves road extraction accuracy as well as computational efficiency. Traditional road centerline extraction methods suffer from 'spur' problem. To tackle this limitation, two methods have been proposed to extract accurate road centerlines from classified road images. The presented work relies on advanced computer vision methods, such as tensor voting and subspace constrained and mean shift (SCMS). The proposed method does not require rigorous road topology hypothesis in advance and thus has high generality. An information fusion method has been presented to combine multiple road extraction results produced by different methods or from different sensors. By contrast to state-of-the-art, the new work is designed from the viewpoint of computational geometry and sheds new insight on fusion of various road results from multiple methods or different sensors. Finally, a seed point based semi-automatic method has been presented to eliminate road gaps to improve the completeness of road network extracted. The presented method can be used to process big road gap, which is a challenging work for most cutting-edge technologies. The experimental results demonstrate the usefulness and feasibilities of the proposed method in this thesis.

List of Publication

1. **Z. Miao**, W. Shi, A. Samat, G. Lisini, and P. Gamba. Information fusion for urban road extraction from VHR optical satellite images. *IEEE Journal of Selected Topics on Geoscience and Remote Sensing*. 2015, DOI: 10.1109/JSTARS.2015.2498663. (accepted, in press)

2. **Z. Miao** and W. Shi. A new methodology for spectral-spatial classification of hyperspectral images. *Journal of Sensors*,vol. 2016, Article ID 1538973, 12 pages, 2016. doi:10.1155/2016/1538973.

3. A. Samat, J. Li, S. Liu, P. Du, **Z. Miao**, J. Luo. Improved hyperspectral image classification by active learning using pre-designed mixed pixels. *Pattern Recognition*, DOI: 10.1016/j.patcog.2015.08.019. (accepted, in press)

4. **Z. Miao**, W. Shi, P. Gamba, and Z. Li. An object-based method for road network extraction in VHR satellite images. *IEEE Journal of Selected Topics on Geoscience and Remote Sensing*. 2015. DOI: 10.1109/JSTARS.2015.2443552. (accepted, in press)

5. **Z. Miao**, B. Wang, W. Shi, and H. Wu. A method for accurate road centerline extraction from a classified image. *IEEE Journal of Selected Topics on Geoscience and Remote Sensing*. 2015, 7(12): 4762-4771.

6. **Z. Miao**, B. Wang, W. Shi, H. Wu, and Y. Wan. Use of GMM and SCMS for Accurate Road Centerline Extraction from the Classified Image, *Journal of Sensors*, vol. 2015, Article ID 784504, 13 pages, 2015. doi:10.1155/2015/784504

7. 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 Transactions on Geoscience and Remote Sensing*, 2015, 53(2): 883-899.

8. **Z. Miao**, B. Wang, W. Shi, and H. Zhang. A semi-automatic method for road centerline extraction from VHR images. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(11): 1856-1860.

9. W. Shi, **Z. Miao***, and J. Debayle. An integrated method for urban main road centerline extraction from optical remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(6): 3359-3372.

10. **Z. Miao** and W. Shi. Road centerline extraction from classified images by using the geodesic method. *Remote Sensing Letters*. 2014, 5(4): 367-376.

11. W. Shi, **Z. Miao***, Q. Wang, and H. Zhang. Spectral-spatial classification and shape features for urban road centerline extraction. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(4): 788-792.

Acknowledgements

First and foremost, my special and endless thanks go to my chief supervisor, Professor Wenzhong Shi. Not only did my academic research benefit from his inspiration and guidance, but much of my remote sensing image processing knowledge comes from him. He constantly supported and encouraged me through the whole years I've worked on this, including both MSc and PhD study. He has read most of my journal manuscripts, providing valuable insights and feedback. Besides, he taught me the most crucial lessons of how to deepen my insights into research strategy and frame an interesting question. He set me on the right path.

I owe great thanks to Professor Paolo Gamba of University of Pavia, Italy. He hosted me as a visiting PhD student for three month, which was an awesome and unforgettable experience. The regular discussion with him once every week was absolutely enlightening and has greatly built my scientific career. When I was finishing some papers, he also helped polish my horrible, sleep-deprived writing into a form that was at least minimally readable and acceptable. Thank you for helping to make my work as good as it could be.

Particular thanks go to Mrs Elaine Anson of Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Dr. Hua Zhang from China University of Mining and Technology and Mr. Y. Li of University of Bath. During my PhD study period, they taught me the elegant scientific writing with grace and humor and also encouraged me to cheer up when I was down and frustrated, which warmed me up. Thank you.

Many of those scholars have kindly answered my questions on the topic of road extraction via email and provided me their constructive suggestions. Therefore, part of this dissertation grew from them, including Prof. Helmut Mayer from The University of the Bundeswehr in Munich, Prof. Jocelyn Chanussot from Grenoble Institute of Technology, and Prof. Hao Wu from Wuhan University of Technology. Thank you all.

Many of my friends have given me ideas, insights, and cutting-edge technologies about remote sensing. The list is long and is still growing longer everyday: Bin Wang, Zhongbin Li, Qunming Wang, Xing Yan, Anshu Zhang, Cheng Yang, Susu Deng, Qianxiang Xu, Han Xu, Mi Jiang, Chisheng Wang, Linghong Ke, Wu Zhu, Yiliang Wan, Teng Liu, Rui Jin, Ya Gao, Sicong Liu, Alim Samto, Yi Liu, Xinxin Wang. I would also convey my thanks to the following friends for their continuous help: Yuanling Zhao, Pan Shao, Xuzhe Lv, Didi Li, Liping Cai, Ming Hao, Pengfei He, Lipeng Gao, Zhenxuan Li, Yaqing Xu, Rui Zhang, Di Wu. I would also like to thank my friends in Pavia: Sheng Fu, Liangjun Chen, Hua Fan, Wenyu Liu, Kaiyan Ma, Wensi Zhang, Pengxiang Zheng, Gianni Cristian Iannelli, Niccolò Ricardi, Gianni Lisini, Amparito Morales Figueroa, Mostapha Harb, Daniele De Vecchi, Guillermo Albrieu Llinás, Andreas Salentinig, Taqi Ad-din, Agarwal Chanchal and Arun Kumar.

This PhD study is funded by the three year Studentship of The Hong Kong Polytechnic University. Without this great financial support, I would not have the chance to study at Hong Kong. I would also like to thank all the staff members in the Department of Land Surveying and Geo-Informatics for their kind help and support during my PhD study.

Last but not least, I am deeply indebted to my parents and all the other family members. Their unconditional love and great patience accompany me through my whole life.

Contents

List of Figures

List of Tables

Abbreviations

VHR	Very High Resolution
KDE	Kernel Density Estimation
SVM	Support Vector Machine
EM	Expectation Maximization
SCMS	Subspace Constrained Mean Shift
CCA	Connected Component Analysis
GMM	Gaussian Mixture Model
OFF	Object-based Frangi's Filter
OSF	Object-based Shape Filter
ISODATA	Iterative Self Organizing Data Analysis Technique Algorithm
SRM	Statistical Region Merging
TV	Tensor Voting
SA	Spectral Angle
SD	Standard Variance
PCA	Principal Component Analysis
RANSAC	RANdom SAmple Consensus
SNR	Signal-to-noise Ratio
MARS	Multivariate Adaptive Regression Splines
MP	Morphological Profile

Chapter 1

Introduction

1.1 Background

Urbanization is happening in just a few years in East Asia, according to the World Bank (2015), 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 (2015). Within the advent of modern acquisition sensors (i.e., Ziyuan-3, Ikonos, and QuickBird, etc.), Very High Resolution (VHR) satellite images have become increasingly available and thus it is possible to monitor urbanization from space. Figure 1.1 presents three satellite images over Xuzhou city, China, captured by different sensors.

Although we can obtain massive satellite images, the useful knowledge is still limited. This is in part due to it is unable to timely process these images. To meet this challenge, object extraction plays an important role in processing these satellite images to produce useful information. Among various object features, road extraction from satellite images has received much attention in the past decades. This is because that updated road layers in Geographic Information Systems (GIS) are critical for many urbanization issues, such

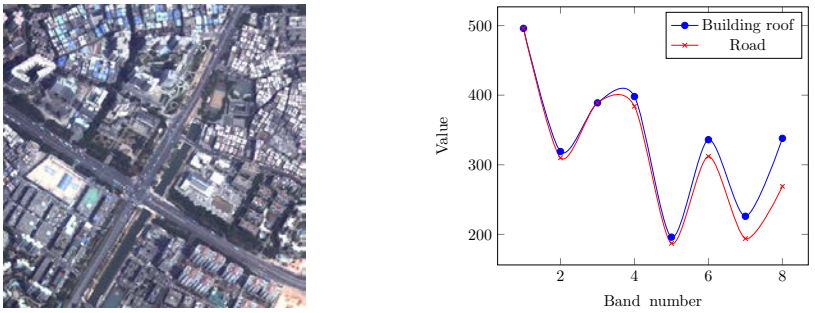
(a) (b) (c)

FIGURE 1.1: (a)-(c) show the satellite images over Xuzhou, China, captured by Landsat, Ziyuan-3, and QuickBird, respectively.

as urban expansion estimation, urban planning, and traffic/population movement monitoring (2014). Although this is not an easy task, especially in urban areas, computer-aided road network extraction from remotely sensed images provides a new opportunity to meet this challenge. Meanwhile, road extraction with the aid of computer can reduce manual work load and improve efficiency. After years of development, there is still no compelling evidence that the state-of-the-art road extraction methods can produce reliable and satisfactory results for any situation (i.e., rural, urban/rural, and urban environments) or any spatial resolution. Numerous road extraction methods have been proposed but it remains unclear as to when these methods will be operational, as there is still rare commercial software regarding this topic. Therefore, the road extraction problem remains a tremendous obstacle in the field of remote sensing. The is because that a number of factors complicate the road extraction task, as summarized in the following section.

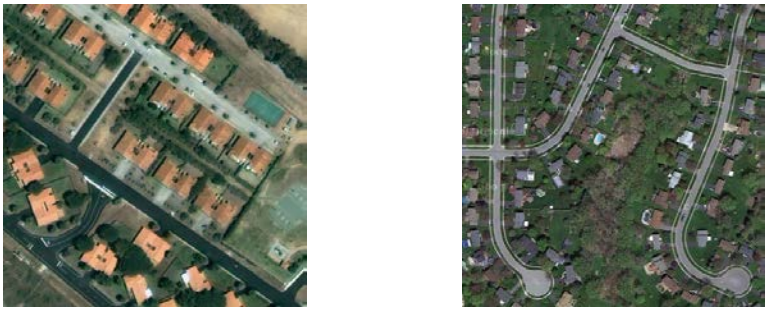
Spectral similarity. The extraction of roads is particularly problematic in urban areas due to the spectral similarity of roads and other impervious surfaces such as buildings, as illustrated in Figure 1.2. Indeed, the spectral separability of asphalt road surfaces and bituminous roofs is still not easy since they tend to share similar spectral properties.

Material change. In real world scenarios, roads are made of various construction materials, such as cement and asphalt. Generally, different materials



(a) (b)

FIGURE 1.2: (a) The test with three seed points which are shown in red. (b) The spatial structural element of mathematical morphology.



(a) (b)

FIGURE 1.3: (a) The test with three seed points which are shown in red. (b) The spatial structural element of mathematical morphology.

lead to different spectral characters. For instance, cement roads have high intensity values on multispectral images, while asphalt roads low intensity values (see Figure 1.3(a)). Hence, the variety of road material threats to the reliability and accuracy of state-of-the-art methods, as most of methods can only process one of these two cases (i.e., high intensity road or low intensity road).

Occupy. This problem is further exacerbated by issues related to the occlusion of road surfaces by trees, shadow and the presence of vehicles (see Figure 1.3(b)). The occlusion of trees and shadow commonly leads to discontinuities of the road segments extracted. Meanwhile, vehicles on the road result in redundant 'holes' of the road segments extracted. This issue is particularly common in dense urban environment, such as Hong Kong.

To overcome aforementioned challenges, the integrating of spectral and spatial

information plays an important role. Given the extreme difficulty of distinguishing road feature from other man-made structures (such as car park and building roof), deriving complementary information (i.e. geometrical feature) from VHR satellite image is a rational option. In any case, understanding to what extent the combination of spectral and geometrical information improves the road extraction performance in VHR satellite image requires further research and dialogue.

1.2 State-of-the-art road extraction methods

To date, road extraction from VHR satellite images has received a lot of attention, and various methods have been proposed. These methods can be roughly classified into two groups: 1) automatic and 2) semi-automatic. In general, semi-automatic 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, semi-automatic 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. A comprehensive review is available in (Mena, 2003), where road extraction methods are further categorized loosely into: 1) seed point based method, 2) active contours, 3) machine learning, 4) object-based method, and 5) other road extraction methods. However, many include methods that cross domains (for example, by exploiting the combination of machine learning and object-based method). The methods are outlined as follows.

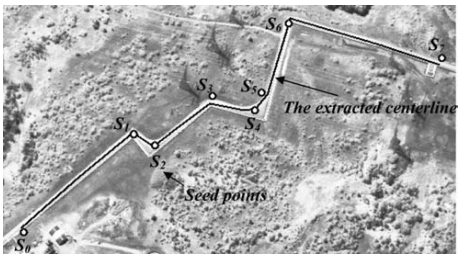


FIGURE 1.4: An example of seed point based method. The source image is from (Hu et al., 2004).

1.2.1 Seed point based method

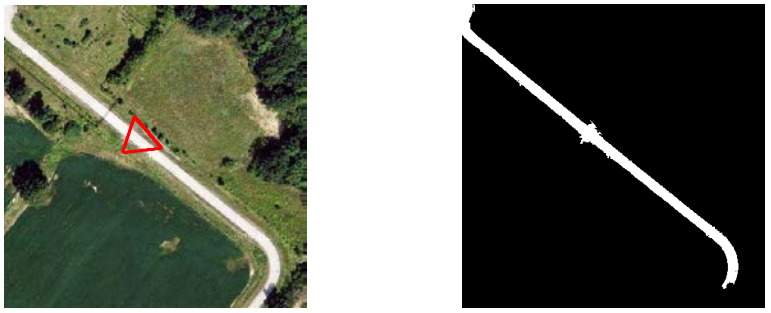
The first category starts from user-defined road seed points, followed by road segments delineation, which exploits spectral/spatial characters of roads to recognize road network elements. In this category, a piecewise parabola model (Hu et al., 2004) was presented to delineate the road centerline network. This method firstly 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 (Kimmel, 2004, Lin et al., 2011). To reduce the seed point number, kernel density estimation incorporating the geodesic method is presented in (Miao et al., 2014b). For a specified pixel, road footprint (Hu et al., 2007) and angular texture signature (Lin et al., 2008) are defined to measure the road probability from the shape prospective. In (Bicego et al., 2003, Zhou et al., 2005), 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 (Movaghati et al., 2010) to process complicated road cases. The single-image mode was discussed in the foregoing road extraction methods. As an extension, Dal Poz *et al.* (2012) presented a semi-automatic 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. Figure 1.4 illustrates an example of seed point based method.

1.2.2 Active contour

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 (Gruen and Li, 1997, Laptev et al., 2000). 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 (Marikhu et al., 2007) 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 (Rochery et al., 2005) 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.* (2015). 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 (Rajeswari et al., 2011), the multi-resolution analysis (Péteri et al., 2003), and Newton snakes (Butenuth and Heipke, 2012). A comparison of different active contour models for road extraction was given by Nakaguro *et al.* (2011). An example of road extraction using active contour is given in Figure 1.5.

1.2.3 Machine Learning

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. A general framework of machine learning is given in Figure 1.6. Support Vector



(a) (b)

FIGURE 1.5: (a) The original test image with initial contour which is shown in red triangle. (b) The active contour extraction result.

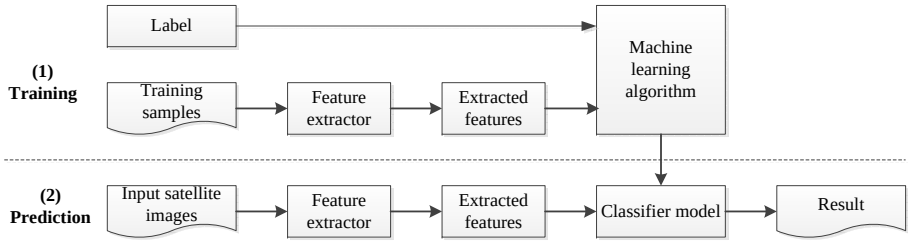


FIGURE 1.6: A general framework of machine learning.

Machine (SVM), a powerful tool in machine learning, has come into wide use in road extraction (Fauvel et al., 2008, Shi et al., 2014a, Song and Civco, 2004, Tarabalka et al., 2009). 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 (Mokhtarzade and Zoej, 2007). A multistage strategy for automatically extracting roads from high-resolution multispectral satellite images based on salient features was introduced by Das *et al.* (2011). This method incorporates the salient features of roads using P-SVM and Dominant Singular Measure (DSM). The path classifier was studied to automatic delineation linear features from images based on the global optimization with geometric priors (Türetken et al., 2011, 2012, Turetken et al., 2013). The junction-point processes (Chai et al., 2013) and a higher-order conditional random field model (Wegner et al., 2013) were also exploited to extract road network from the image. Particularly, Mnih and Hinton (2010, 2012) addressed the application issue of machine learning in road extraction from aerial images.

1.2.4 Object-based method

The fourth category is relying on the object level information (see Figure 1.7) 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 (Baatz and Schäpe, 2000, Benz et al., 2004, Herold et al., 2002). Object-based methods are known to achieve better results than pixel-based methods in processing VHR satellite images (Blaschke, 2010, Blaschke and Strobl, 2001, Cleve et al., 2008). A considerable number of studies have compared object-based approaches with traditional pixel-based classification methods (Baltsavias, 2004, Ehlers et al., 2006, Frohn et al., 2005, Im et al., 2008, Laliberte et al., 2004). 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 (2009, 2013) is based on SVM and multi feature model at both pixel and object levels. Peng *et al.* (2008) updated outdated road maps by incorporating generic and specific prior knowledge into a multi-scale phase field model. Zarinpanjeh *et al.* (2013) used object-based analysis for road map updates. Additionally, Grote *et al.* (2012) developed a method for road network extraction using object-based analysis. 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 (Das et al., 2011, Han et al., 2012, Shi et al., 2014a, Song and Civco, 2004), directional mathematical morphological (Shi et al., 2014a, Valero et al., 2010), directional filter (Gamba et al., 2006, Negri et al., 2006), *et al.* 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 (Miao et al., 2013). Although integrating shape features in road extraction results in a good performance, it is challenging to obtain a universal linear feature

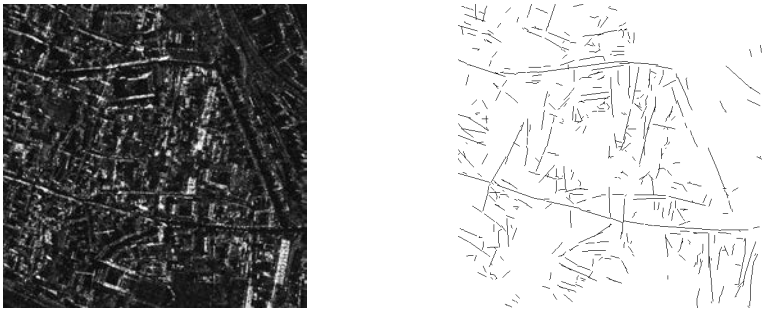
(a) (b)

FIGURE 1.7: (a) A test optical satellite image. (b) The derived objects where the boundary between objects is shown in red.

extraction method for all situations (Gamba et al., 2006), and this thesis will further study this issue.

1.2.5 Other road extraction methods

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 (2004). Some new researches based on SAR imagery and LIDAR systems also have recently been presented (Gamba et al., 2006, Negri et al., 2006, Poullis and You, 2010). An example of road extraction using SAR image is reported in Figure 1.8. Doucette *et al.* (2001) proposed an automated road centerline extraction method that exploits spectral content from high-resolution multi-spectral images. The method is based on anti-parallel edge centerline extraction and self-organized road mapping. A fast linear feature detector for road extraction was introduced by Shao *et al.* (2011). 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 (Cao and Sun, 2014).



(a) (b)

FIGURE 1.8: (a) A test SAR image. (b) The road extraction result.

1.3 Objectives of the thesis

Based on the aforementioned issues, four objectives of this thesis have been identified as follows:

- To design several measures to evaluate the road shape features.
- To propose a practical and effective method to extract road segments from the VHR satellite images.
- To design a framework to characterize accurate road centerlines from classified road images.
- To propose a method to connect large road gaps to eliminate road discontinuities.

1.4 Thesis outline

The outline of the thesis is summarized as Figure 1.9. The structure of this thesis is organized as follows:

Chapter 2 presents a new approach for urban main road centerline extraction from optical satellite images integrating spectral-spatial classification, Local Geary's C, road shape features, locally weighted regression and tensor voting. This method has three main contributions: 1) initial road network extraction

by fusion of spectral-spatial classification based on GANMM and homogeneous property measured by Local Geary's C; 2) false roads removal based on roads shape feature to extract reliable roads; 3) road centerline extraction based on locally weighted regression and road network generation using tensor voting. This work has been previously published in (Shi, Miao, and Debayle, 2014a) and (Shi, Miao, Wang, and Zhang, 2014b). Chapter 7 concludes remarks, followed by a brief look at future research for improving road network extraction algorithms presented in this thesis.

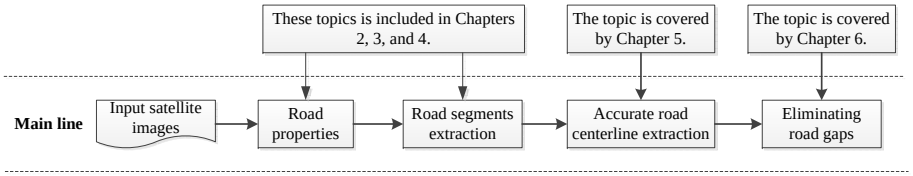


FIGURE 1.9: The organization of the thesis

Chapter 3 explores a novel object-based road extraction method for VHR satellite images. Specially, the main contributions of this chapter are: 1) a novel proposal for two object-based filters to enhance road features in VHR satellite images; 2) the design of a hybrid feature set to extract road features; and 3) the integration of tensor voting (TV), active contour methods and the geometry information to simultaneously fill road discontinuities and improve the road smoothness. This work has been published in (Miao et al., 2015).

Chapter 4 addresses the issue of automatic dense urban road extraction. The main contributions are four information fusion strategies designed from the viewpoint of computational geometry.

Chapter 5 explores ways of delineating accurate centerlines from classified road maps. We investigate two ways of performing accurate road centerline extraction, called Feature Point based Subspace Constrained Mean Shift (F-SCMS) method and Gaussian Mixture Model based Subspace Constrained Mean Shift (GMM-SCMS) method. This chapter includes work previously published in (Miao et al., 2014a) and (Miao et al., 2014c).

Chapter 6 presents a novel algorithm to eliminate large road gaps to improve the road extraction performance. The main contribution of this chapter is that the road gap elimination problem is converted to a seed point based road extraction issue. This chapter is mainly based on (Miao et al., 2014b).

Chapter 2

An Integrated Method for Road Extraction

2.1 Introduction

The usefulness of VHR satellite images for land use and land cover mapping is widely recognized. Due to the their spatial resolution and level of thematic information content, VHR satellite imagery present great potential to meet this requirement. The advantages of this category of imagery are especially visible in spatially and spectrally complex areas; like for example diverse agriculture landscape with small and narrow fields. Due to the spatial resolution of VHR images, the classification problems with mixed pixels widely known for lower resolution images are in this case on second-rate level.

The shape features derived from VHR satellite images in image analysis have already significantly confirmed their usefulness in quite wide range of applications. This chapter will introduce several shape features which are suitable for the road extraction.

2.2 The proposed method

The proposed method is presented in this section. The organisation of this method is shown in Figure 2.1. The details of the above steps follow in detail below.

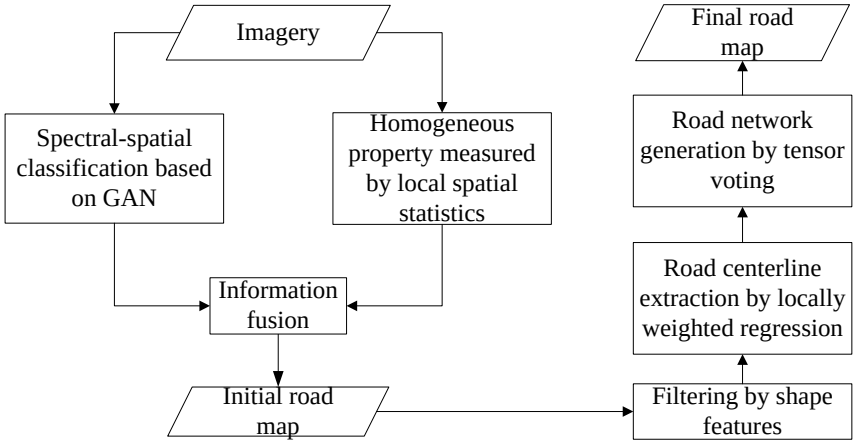


FIGURE 2.1: Flowchart of the proposed methodology.

2.2.1 Spectral-spatial classification

The purpose of the spectral-spatial classification step is to segment the imagery into two groups: road groups and non-road groups. In this study, general adaptive neighbourhood mathematical morphology (GANMM) is investigated for spectral-spatial classification.

The GANMM has been introduced by Debayle and Pinoli (2006, 2009, 2012). The central idea of GANMM is to substitute traditional fixed-shape structural elements (SEs) by adaptive SEs. Suppose $D \subseteq R^2$ is the spatial domain of an image and I is the natural set of image mappings from D into $\mathbb{R}$. In this chapter, the general adaptive neighbourhood (GAN) sets $V_m^f(x)$ are defined as

$$V_m^f(x) = C_{f^{-1}([f(x)-m\ ,\ f(x)+m])}(x) \tag{2.1}$$

where f is a criterion mapping, $m \in \mathbb{R}^+$ is a homogeneity tolerance, x is a pixel within the spatial support D of the image $f \in I$. The set $V_m^f(x)$ is

(a) (b) (c)

FIGURE 2.2: (a) The test with three seed points which are shown in red. (b) The spatial structural element of mathematical morphology. (c) The spatial structural element of GAN.

characterized by the following two properties:

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

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

$$\forall (m, f, x) \in \mathbb{R}^+ \times C \times D \; R_m^f(x) = \bigcup_{z \in D} \left\{ V_m^f(z) \,\middle|\, x \in V_m^f(z) \right\} \tag{2.2}$$

Figure 2.2 illustrates the difference of spatial element used in MM and GANMM. It can be seen that the adaptive SE used in GANMM is tailored to the real spatial structure of the nature image. In the field of image processing, mathematical morphology (MM) is one of commonly used image processing technologies. MM can be used to smooth the noise image. In fact, the closing operator in MM can be taken as a non-linear smoothing filter. Morphology has become popular in recent years. The basic operations are available in many

image analysis software packages. Morphology lends itself to efficient parallel hardware implementations and computers using this are available. For detail background of MM, we referred interesting readers to (Serra, 1986). Despite its simplicity and efficiency, MM suffers from two factors: 1) the predefined fixed structural element (SE), and 2) the window size selection. In most cases, the image filtering result by MM cannot retain the spatial structure of the original image that leads to useful information loss.

By applying the adaptive SEs, the adaptive dilation and adaptive erosion are then defined as,

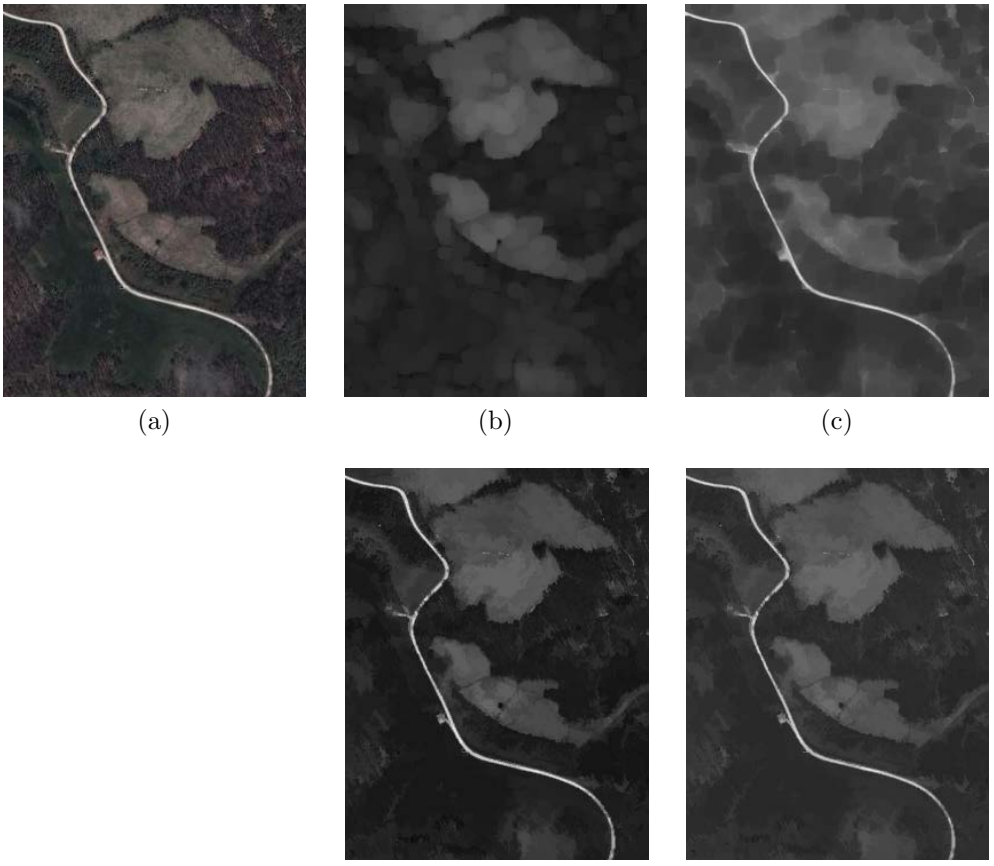
$$\mathrm{D}_m^f(f)(x) = \sup_{\omega \in R_m^f(x)} f(\omega) \, \mathrm{E}_m^f(f)(x) = \inf_{\omega \in R_m^f(x)} f(\omega) \tag{2.3}$$

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

$$\mathrm{C}_m^f(f)(x) = \mathrm{E}_m^f \circ \mathrm{D}_m^f(f)(x) \, \mathrm{O}_m^f(f)(x) = \mathrm{D}_m^f \circ \mathrm{E}_m^f(f)(x) \tag{2.4}$$

Figure 2.3 presents a comparison of classical and GAN-based MM. As can be seen, results of classical MM are blurred and the image spatial structure is also damaged by using classical MM. However, GANMM does not produce blurred effects; hence the image spatial structure is well retained. The results also indicate that GANMM produces smooth results which reduce the local spectral variance caused by image noise. One limitation of GANMM is that its computational cost is high for processing large images and with small homogeneous tolerance values. To tackle this limitation to some extent, this chapter presents an efficient GANMM (EGANMM) (see Appendix A and Appendix B).

In the next phase, morphological profiles (MPs) are constructed based on the aforementioned GANMM method. Morphological profiles (MPs), proposed by Pesaresi and Benediktsson (2001), are representations of the size and shape information of objects in the image. The profiles are composed of multiple



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

FIGURE 2.3: The comparison of classical MM and GANMM. (a) Original image. (b) Classical opening. (c) Classical closing. (d) Adaptive opening. (e) Adaptive closing. In this example, the homogeneity tolerance value is 30.

morphological processes using increasing sizes of a structuring element (SE). In this chapter, the adaptive MPs (using GANMM) at a pixel x of an image $f(I)$ is computed as,

$$\mathrm{MP}(x) = \{\mathrm{C}_{m_n}(x), ..., I(x), ..., \mathrm{O}_{m_n}(x)\} \tag{2.5}$$

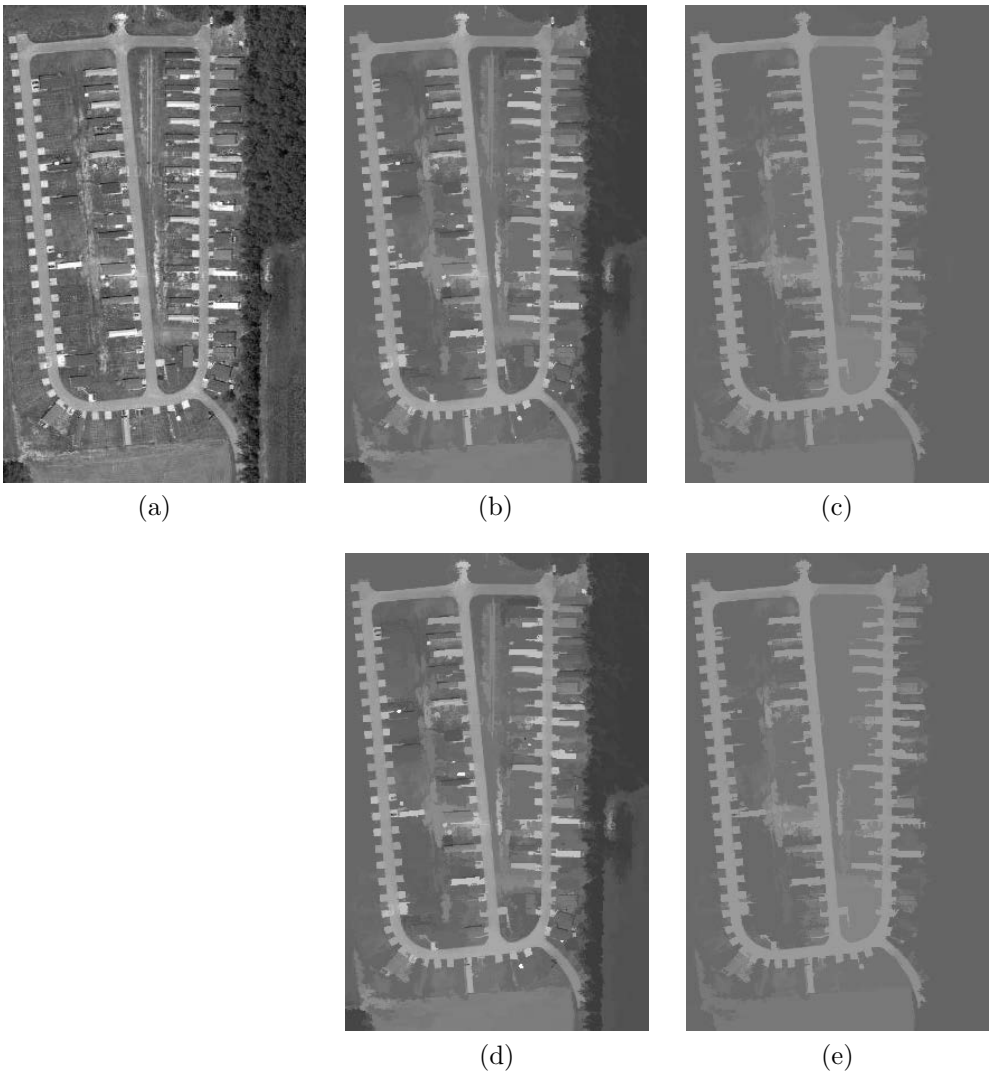
where m_n denotes the nth homogeneous tolerance.

For the multispectral imagery with B-bands, Equation 2.5 will be rewritten as:

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

As can be seen from the Equation 2.5, MPs contain spectral information as

well as the image spatial information. Figure 2.4 shows an example of MPs obtained by applying a series of homogeneity tolerance. After the construction of MPs, the pixel-wise classifier, such as Bayes classifier, Support Vector Machine (SVM) (Cortes and Vapnik, 1995) and Neural Networks, is then selected to classify the imagery. In this chapter, SVM is selected for two main reasons. First, previous studies show that the SVM performance is as good as or significantly better than other competing methods in most cases (Burges, 1998). Second, SVM is a supervised classification method which can still work when road materials and intensity change. Here, the SVM classification result



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

FIGURE 2.4: The construction of MPs based on adaptive openings and closings. (b)-(c) show adaptive opening results, while (d)-(e) adaptive closing results. In this example, homogeneity tolerances are 30 and 50.

is expressed as,

$$S(x) = \begin{cases} 1 & \text{, if } x \text{ is classfied as road} \\ 0 & \text{, otherwise} \end{cases} \tag{2.6}$$

2.2.2 Local homogeneous property

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

Roads are continuous features. Based on this property, the inference is that roads should be located in homogeneous regions. In this chapter, Local Geary's C (Anselin, 1995) is used to measure image local homogeneity properties. Local Geary's C, one of the local spatial statistics indicators, measures the autocorrelation between a pixel and its neighbouring pixels.

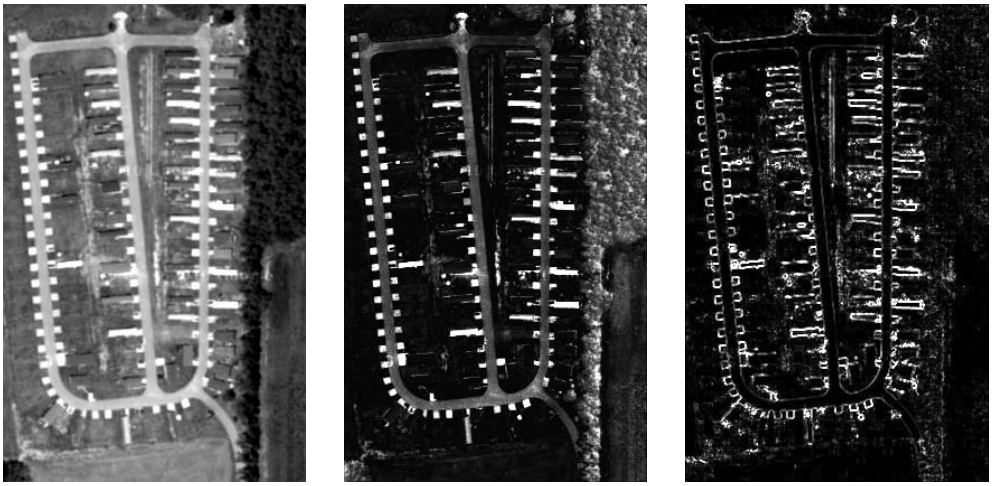
Local Geary's C is defined as:

$$\begin{aligned} c_i &= \frac{1}{\frac{1}{n}\sum_{j=1}^{n}(y_j - \bar{y})^2} \sum_{j=1}^{n} w_{ij}[(y_i - \bar{y}) - (y_j - \bar{y})]^2 \\ &= \frac{1}{\frac{1}{n}\sum_{j=1}^{n}(y_j - \bar{y})^2} \sum_{j=1}^{n} w_{ij}(y_i - y_j)^2 \end{aligned} \tag{2.7}$$

where n is the number of georeferenced observations, y_i is the value of the observation at the i^{th} location, $\bar{y}$ is the mean of the observations, w_{ij} is the weight of spatial relations between points i and j. For ease of computation, w_{ij} is set as 1 in this study.

The Local Geary's C index identifies areas of high variability between a pixel value and its neighbouring pixels. It is useful for detecting edge areas between clusters and other areas with dissimilar neighbouring values. Figure 2.5 gives

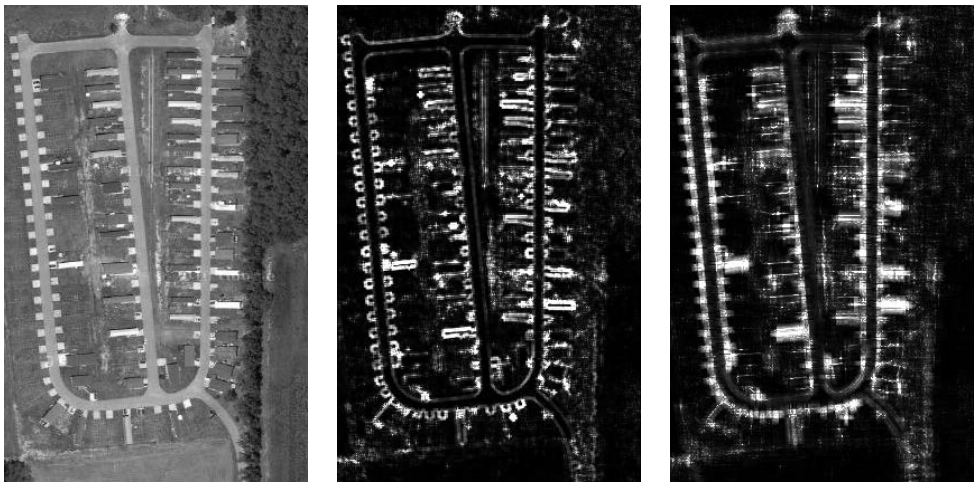
a comparison result of three local spatial statistics. As can be seen, Local Geary's C is more suitable to measure the local homogeneity of the grey values. In this study, the window size is 2 pixels as suggested by experimental results.



(a) (b) (c)

FIGURE 2.5: (a)-(c) show the Getis-Ord Gi map, Local Moran's I map, and Local Geary's C map respectively.

Figure 2.6 shows the results of Local Geary's C on an image with different window sizes. As shown in Figure 2.6(c), the Local Geary's C computed from large size windows damages the homogeneous property of small areas. After



(a) (b) (c)

FIGURE 2.6: (a) Original image. (b) Local Geary's C result (window size = 2 pixels). (c) Local Geary's C result (window size = 6 pixels).

the computation of the Local Geary's C of every spectral band, a binary SVM

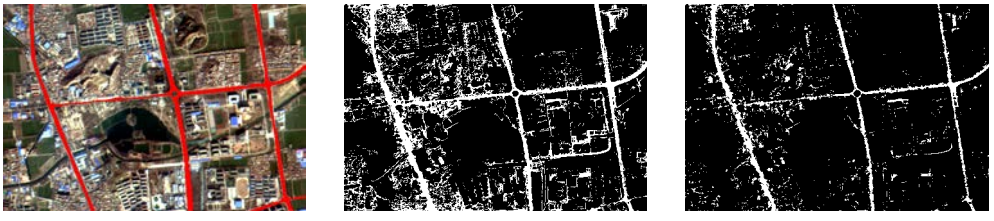
classification is then applied to extract homogeneous regions. The binary classification result is defined as,

$$L(x) = \begin{cases} 1 & \text{, if } x \text{ is in the homogeneous region} \\ 0 & \text{, otherwise} \end{cases} \tag{2.8}$$

Once the spectral-spatial classification and Local Geary's C results are ready, these two results can be combined to improve the robustness and the accuracy of extracted roads by incorporating different road features. After Step A and Step B, the classification and Local Geary's C binary result is combined using the following information fusion rule,

$$F(x) = \begin{cases} 1 & \text{, if } S(x) = 1 \textit{ and } L(x) = 1 \\ 0 & \text{, otherwise} \end{cases} \tag{2.9}$$

A comparison of classification and information fusion results is presented in Figure 2.7. As can be seen, the information fusion not only removes misclassified roads to some extent, but also disconnects most misclassified and true roads. After disconnecting misclassified and true roads, it can be seen from Figure 2.7(c) that misclassified and true roads have different geometrical properties. This opens a door to filter information fusion result using shape features to further improve road extraction accuracy.



(a) (b) (c)

FIGURE 2.7: (a) Original image. The reference road is shown in red. (b) Classification result. (c) Information fusion result.

2.2.3 Filtering misclassified road pixels

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

For an $m \times n$ image, its $(p+q)$ order moment (Flusser, 2006) is defined as,

$$m_{pq} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x)^p \cdot (y)^q I(x,y) \tag{2.10}$$

where $I(x,y)$ is the grey value at the point (x,y), p and q are non-negative integers, $(p+q)$ is called the order of the moment.

Corresponding central moment μ_{pq} is then given as,

$$\mu_{pq} = \sum_{x} \sum_{y} (x - x_c)^p (y - y_c)^q I(x,y) \tag{2.11}$$

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

Based on Equation 2.11, the normalized moment η_{pq} is described as,

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{2.12}$$

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

The sum of second order moments is computed as:

$$M_1 = \eta_{02} + \eta_{20} \tag{2.13}$$

Previous studies (Lu et al., 2009) show that roads have large values of M_1 as they are complex and irregular, and hence, segments with values of M_1 less than a threshold T_{M_1} are removed. In this chapter, the T_{M_1} threshold is automatically defined with Otsu's adaptive method (Otsu, 1975). Connected component analysis (CCA) is used to label each disjoint segment and compute its value of M_1. The detailed filtering steps by image moments are given in Algorithm 1, as follows:

Algorithm 1 Removing misclassified road pixels

Input
The original classified image
Output
The filtered image.

Label the connected segments.
Obtain second moment (M_1) value of each connected segment.
Select the T_{M_1} threshold by using Otsu's method.
for e **do**ach component
    if (**then**$M_1 \leq T_{M_1}$)
        delete that segment
    end if
end for

2.2.4 Road centreline extraction

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

Local linear kernel regression (Hastie et al., 2009) solves a separate weighted least squares problem at each target point x_0

$$\min_{\alpha(x_0),\beta(x_0)} \sum_{i=1}^{N} K_\lambda\left(x_0, x_i\right)\left[y_i - \alpha\left(x_0\right) - \beta\left(x_0\right) x_i\right]^2 \tag{2.14}$$

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

After solving Equation 2.14, the estimate is then expressed as,

$$\hat{f}(x_0) = \hat{\alpha}(x_0) + \hat{\beta}(x_0)\, x_0 \tag{2.15}$$

The matrix form of Equation 2.15 is given as

$$\begin{aligned} \hat{f}(x_0) &= b(x_0)^T \left(\mathbf{B}^T \mathbf{W}(x_0) \mathbf{B}\right)^{-1} \mathbf{B}^T \mathbf{W}(x_0)\, y \\ &= \sum_{i=1}^{N} l_i(x_0)\, y_i \end{aligned} \tag{2.16}$$

where $\beta = \begin{bmatrix} b(x_1)^T \\ b(x_2)^T \\ \vdots \\ b(x_N)^T \end{bmatrix}_{N\times 2}$, $\mathbf{W} = \begin{bmatrix} K_\lambda(x_0, x_1) & & \\ & \ddots & \\ & & K_\lambda(x_0, x_N) \end{bmatrix}_{N\times N}$, $b(x)^T = (1, x)$.

In real world applications, road networks are complex and generally have many junctions. Hence, road centrelines cannot be directly extracted by regression. To overcome this problem, a road network decomposition method is proposed in this study. Figure 2.8 gives a summary of the centreline extraction from classified roads using a locally weighted regression method. Firstly, junction areas of the road network are detected by the tensor voting method (Medioni et al., 2000, Mordohai and Medioni, 2006), which will be introduced in the next section. Secondly, the road network is decomposed to unconnected parts using the tensor voting method. Thirdly, centrelines are extracted by applying the locally weighted regression to every branch. Finally, all isolated centrelines are combined with junctions.

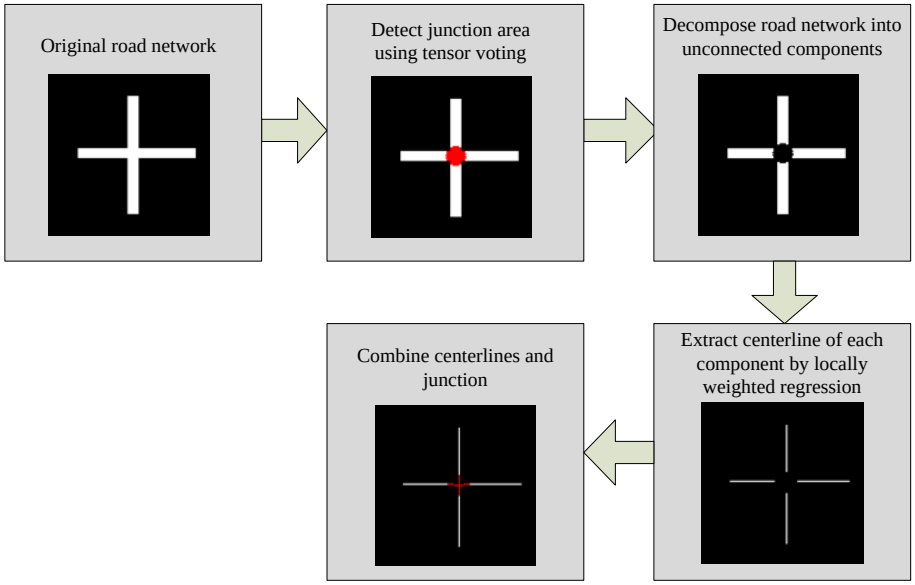


FIGURE 2.8: Road centreline extraction using local linear kernel smoothing regression.

2.2.5 Road network generation

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

Tensor voting is a perceptual grouping and segmentation framework introduced by Medioni *et al.* (Medioni et al., 2000, Mordohai and Medioni, 2006). In 2-D, a second-order symmetric tensor T is defined as,

$$\begin{aligned} T &= \begin{bmatrix} \overrightarrow{e_1} & \overrightarrow{e_2} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \overrightarrow{e_1} \\ \overrightarrow{e_2} \end{bmatrix} \\ &= \lambda_1 \overrightarrow{e_1}\overrightarrow{e_1}^T + \lambda_2 \overrightarrow{e_2}\overrightarrow{e_2}^T \end{aligned} \tag{2.17}$$

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

$$T = (\lambda_1 - \lambda_2)\, \overrightarrow{e_1}\overrightarrow{e_1}^T + \lambda_2 \left(\overrightarrow{e_1}\overrightarrow{e_1}^T + \overrightarrow{e_2}\overrightarrow{e_2}^T \right) \tag{2.18}$$

where $\overrightarrow{e_1}\overrightarrow{e_1}^T$ is a stick tensor with associated saliency $(\lambda_1 - \lambda_2)$ which indicates an elementary curve, $(\overrightarrow{e_1}\overrightarrow{e_1}^T + \overrightarrow{e_2}\overrightarrow{e_2}^T)$ with associated saliency λ_2 describes a structure which has no orientation preference to a location where multiple orientations coexist. The saliency decay function has the following form:

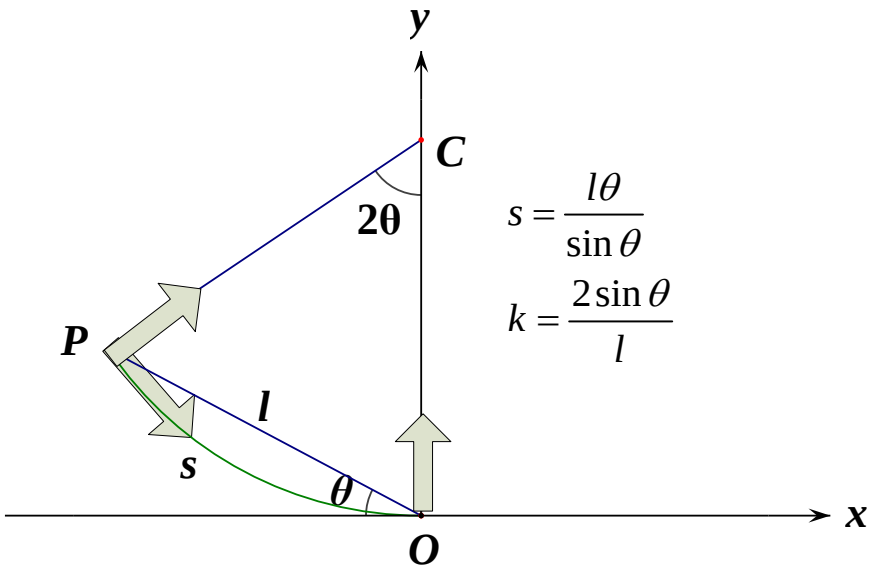


FIGURE 2.9: Votes cast by a stick tensor located at the origin O. C is the centre of the osculating circle passing through points P and O.

$$DF(s, \kappa, \sigma) = e^{-\left(\frac{s^2 + c\kappa^2}{\sigma^2}\right)} \tag{2.19}$$

where s is the arc length OP (see Figure 2.9), κ is the curvature, c controls the degree of decay with curvature, and σ is the scale factor which determines the effective neighbourhood size.

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

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

An important advantage of tensor voting is that it can recognise global context refinement automatically and the connection hypothesis does not need to be set in advance. Another advantage of tensor voting is that it has only one parameter, scale factor σ, to be set by users. These advantages of tensor voting make it an ideal choice to eliminate road discontinuities. Figure 2.10 shows two examples of road connection results. It can be seen that discontinuities are successfully removed.

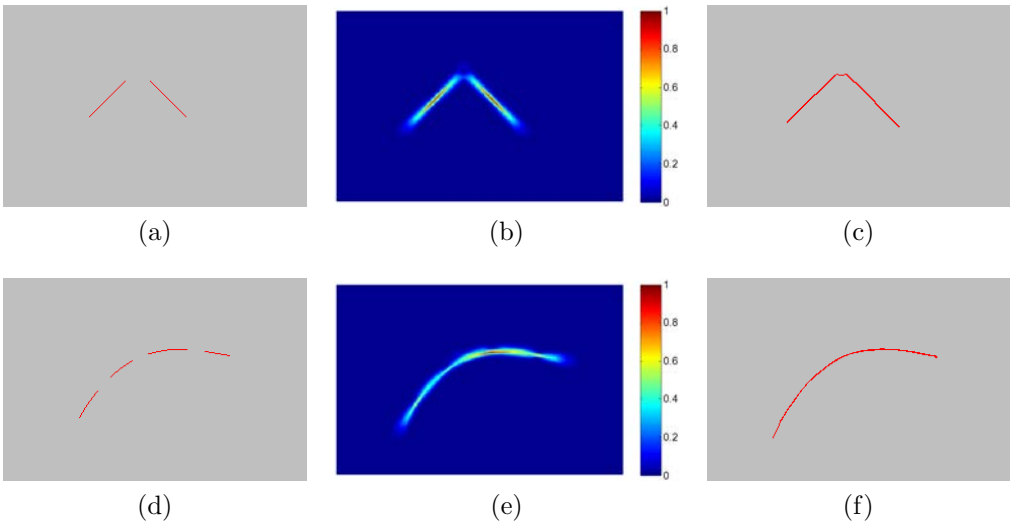


FIGURE 2.10: Road connections using tensor voting. The first to third columns show the input images, stick saliency maps, and connection results, respectively.

2.3 Experimental evaluation

This chapter focuses on main road extraction from urban remote sensing images. In this chapter, the definition of "main road" depends on road width. The road whose width is 4-6 pixels is defined as "main road". The corresponding road reference map is generated by hand drawing method.

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

2.3.1 Efficiency comparison of GANMM and EGANMM

The computational efficiency of GANMM and EGANMM is firstly compared on an gray image with a spatial size of 301 $\times$ 301 pixels. To this end, the homogeneity tolerance value is changed from 3 to 33 with an increment step of 5 and we observe its influence on the computational efficiency on GANMM and EGANMM, respectively. The comparison results are reported in Figure 2.11.

From Figure 2.11, it is clear that the computational load of GANMM fluctuates significantly over the change of homogeneity value. At the smallest homogeneity value (i.e. 3) GANMM stands at the highest the computational load. With the increment of homogeneity value, the computational load drops rapidly. The figures remain stable when the homogeneity value is becoming larger than a certain value. For instance, in Figure 2.11, when the homogeneity value is becoming larger than 28, the computational load is gradually close to 30s. Moreover, the image spatial size also has a factor on the GANMM efficiency. For the same homogeneity tolerance value, the larger spatial size, the higher computational load of GANMM is. For instance, when the homogeneity tolerance value is 3, the computational load is approximately 350s for an image with a spatial size of 301 $pixels \times 301\ pixels$. Therefore, both the homogeneity tolerance value and the image spatial size influence the GANMM efficiency. The small homogeneity tolerance value or large spatial size leads to high computational load of GANMM.

On the contrary, EGANMM is insensitive to the image spatial size and homogeneity value. With the increase of image spatial size and homogeneity value, the computational load of EGANMM is steadily changed. Figure 2.11 indicates that EGANMM is more efficient than GANS, particularly when the homogeneity value is small or the image spatial size is larger. This demonstrates that the computational efficiency of EGANMM is less sensitive to the homogeneity tolerance value than that of GANMM. In other words, EGANMM is more robust to the homogeneity tolerance value than GANMM.

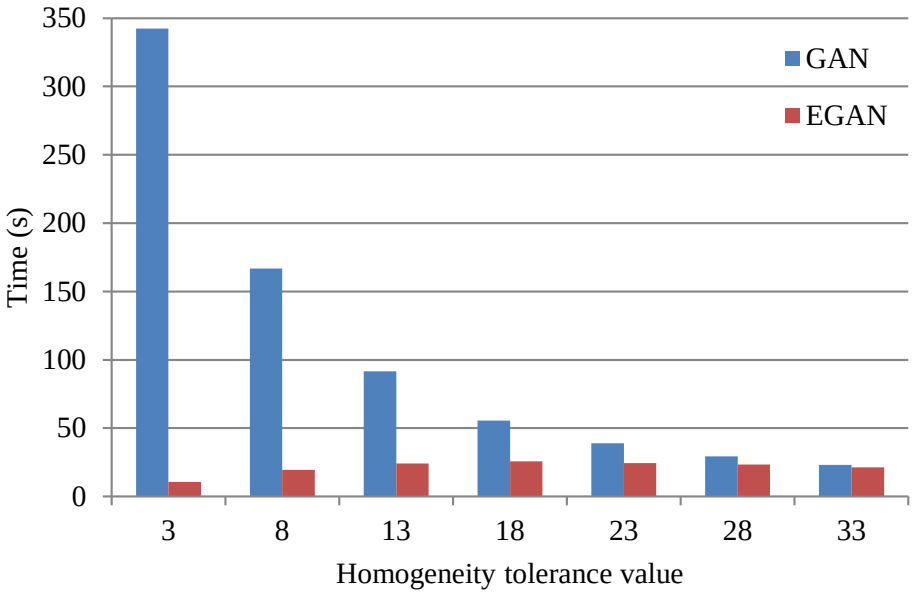


FIGURE 2.11: The efficiency comparison of classical GANMM and the proposed efficient GANMM.

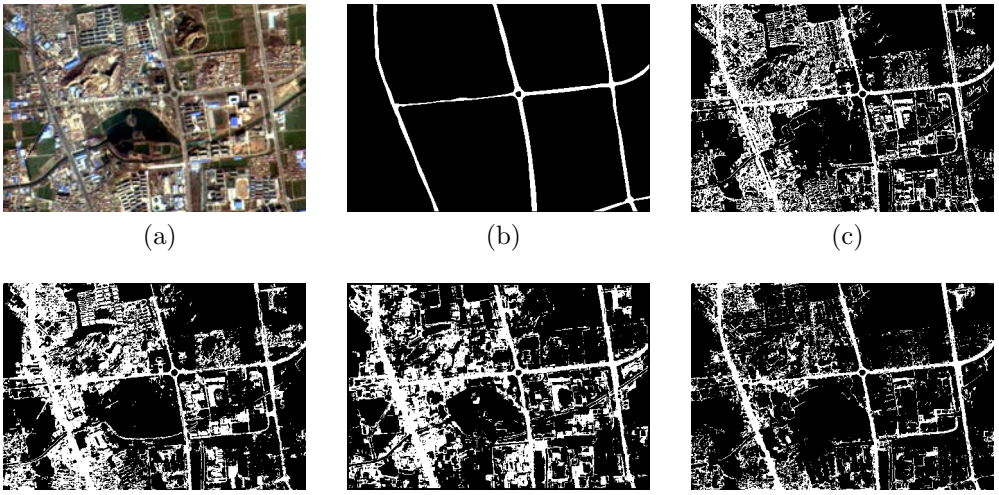
2.3.2 Tests on different spectral-spatial classification methods

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

The experimental area is a part of the Xuzhou City image recorded by an infrared multispectral spectrometer carried by the Ziyuan-3 satellite, a Chinese Earth Observation satellite. The Ziyuan-3 satellite technical data is given in Table 2.1. Figure 2.12(a) gives the image of the dataset. The hand-drawn reference road map is given in Figure 2.12(b). Figure 2.12(c) shows the road extraction result by pixel-wise SVM classification (Cortes and Vapnik, 1995), Figure 2.12(d) shows the road extraction result by the algorithm proposed by Tarabalka *et al.* (Tarabalka et al., 2009), using the EM clustering approach. The result by the algorithm proposed by Fauvel *et al.* (Fauvel et al., 2008) is illustrated in Figure 2.12(e) and Figure 2.12(f) shows the roadmap detected by the method proposed in this study.

TABLE 2.1: Ziyuan-3 satellite technical data

General features	
Launch date	11-Jan-12
Launch vehicle	Long March 4B
Nominal lifetime	5 years
Spectral bands and resolution	Panchromatic (front-facing and rear-facing): 3.5m
	Panchromatic (ground-facing):2.1m
	Multi-spectral (ground-facing):6.0m
	Panchromatic (front-facing and rear-facing): 52.3m
Swatch	Panchromatic (ground-facing): 51.1m
	Multi-spectral (ground-facing): 51.0m
Wavelength range	Panchromatic: 450nm-800nm
	Multi-spectral
	Blue: 450nm-520nm
	Green: 520nm-590nm
	Red: 630nm-690nm
	Near infrared: 770nm-890nm



(a) (b) (c)

(d) (e) (f)

FIGURE 2.12: (a) Original image. (b) Reference road map. (c) Road extraction result by pixel-wise SVM (Cortes and Vapnik, 1995). (d) Road extraction by Tarabalka's method (Tarabalka et al., 2009). (e) Road extraction by Fauvel's method (Fauvel et al., 2008). (f) Road extraction result by the proposed method.

The classification accuracy and error rate (Sokolova and Lapalme, 2009) are used to evaluate different spectral-spatial classification methods. The results are presented in Table 2.2. As can be seen, Tarabalka's method (Tarabalka et al., 2009) and Fauvel's method (Fauvel et al., 2008) achieve lower classification accuracy compared to pixel-wise SVM which uses only spectral information. This is because many non-road pixels are misclassified as roads by using these two methods. The results indicate that GAN integration together with the spectral classification results substantially improves road classification accuracy. Compared to classification using only spectral information, there is a 10% improvement in the proposed classification method. A vision comparison reveals that the completeness and correctness of the proposed method are superior to those of the other three methods. Hence, spectral-spatial classification based on GAN is used in the proposed method.

TABLE 2.2: The comparison of detected road accuracy by different spectral-spatial classification methods

Method	Classification accuracy (%)	Error rate (%)
Pixel-wise SVM (Cortes and Vapnik, 1995)	74.21	25.79
Tarabalka's method (Tarabalka et al., 2009)	70.87	79.13
Fauvel's method (Fauvel et al., 2008)	72.71	27.29
Proposed method	84.83	15.17

2.3.3 Experiment 1

In this experiment, an image recorded by the Ziyuan-3 satellite is used to verify the performance of the proposed method. The image is 322 by 472 pixels. Figure 2.13(a) shows the study area in image form. The manually formed ground truth data is shown in Figure 2.13(b). The remote sensing image used in this study has a spatial resolution of 6m per pixel. At this resolution, smaller roads are unclear, hence only "main roads" with widths of about 5-6 pixels are able to be extracted. Firstly, the morphological profile (MP)

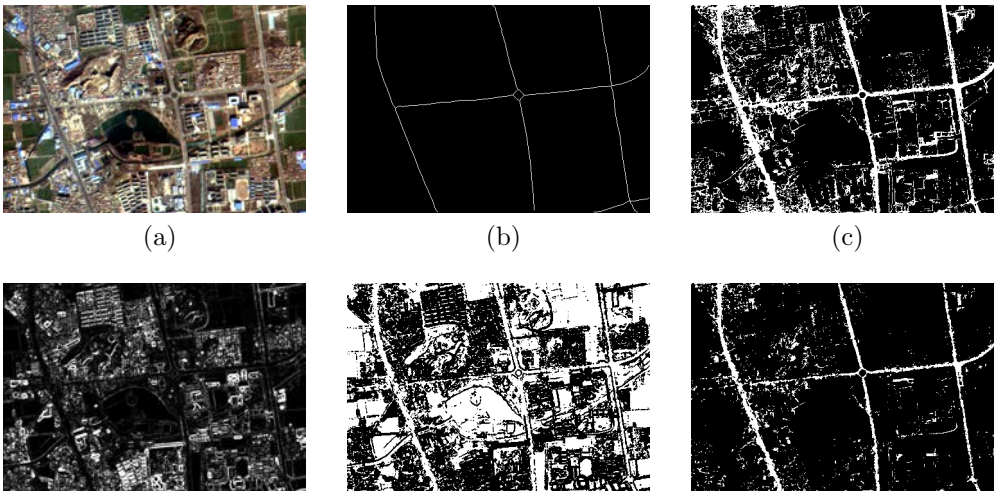


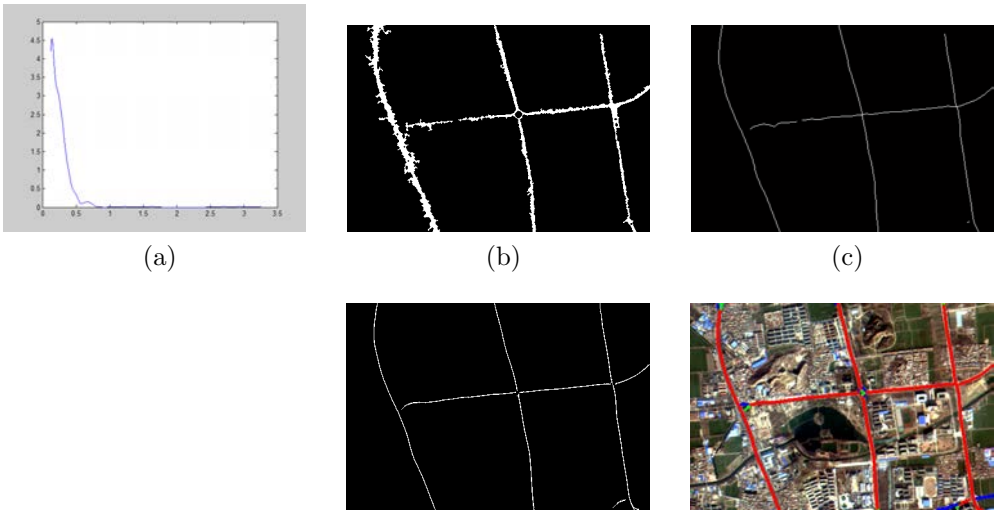
FIGURE 2.13: (a) The study area image. (b) A hand-drawn reference road map. (c) Classification result using SVM. (d) Example of Local Geary's C results. (e) Binary classification of Local Geary's C map. (f) Fusion result.

was constructed by applying a series of GAN openings and with increasing homogeneity tolerance. In this study, the homogeneity tolerance is set to 10, 20 ... 40. SVM was then trained to classify the whole image. In this chapter, Libsvm library (Chang and Lin, 2011) was used to train SVM classifier. The optimal parameters were selected as $C = 128$ and $\gamma = 0.00313$. The SVM classification result is shown in Figure 2.13(c). The road features are shown in white and the non-road features in black. As can be seen from this Figure, road features are extracted well by spectral-spatial classification. However, due to spectral similarity, false roads such as parks, buildings and bare soil were misclassified as roads and further processing was needed.

After spectral-spatial classification, Local Geary's C of each band was computed. Figure 2.13(d) shows an example of the Local Geary's C result. The result shows that the homogeneous areas have low intensity and edge pixels have large intensity. By applying supervised binary classification, the Local Geary's C map was segmented into two parts: homogeneous regions and edge pixels, as shown in Figure 2.13(e) where homogeneous regions are shown in white and edge pixels in black.

The classification result and Local Geary's C binary result were fused by applying logical 'and' operation (see Figure 2.13(f)). As can be seen, fusion of classification and Local Geary's C result reduced misclassified roads to some extent. Some misclassified roads which connected with roads were also disconnected after applying information fusion which is convenient to remove by using shape feature in the following step. Based on these results, it is concluded that information fusion leads to improve the robustness of road extracted.

After fusion processing, the second moment value of each object in road imagery was computed. Figure 2.14(a) shows the second moment distribution. As seen in Figure 2.14(a), the second moment difference between road feature and non-road feature is obvious. In this experiment, the second moment threshold is set as 0.33. The objects whose second moment values smaller than 0.33 were removed. The filtering result is shown in Figure 2.14(b). The locally weighted regression was then performed to extract road centreline and is shown in Figure 2.14(b). The road centreline map is shown in Figure 2.14(c). To eliminate discontinuity, tensor voting was performed. In this study, the scale parameter was set as 20. The road network generation result is shown in Figure 2.14(d). The superposition result of the original image and extracted road centreline is shown in Figure 2.14(e).
Figure 2.15 shows comparison results produced by the thinning algorithm and the proposed method. It is clear from results that the road centreline extracted by the thinning algorithm has many spurs which reduces the smoothness and accuracy of the result. As seen from the results, the centreline extracted by



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

FIGURE 2.14: (a) Second order moment distribution of information result. (b) Filtering result. (c) Road centreline extraction result. (d) Road network generation result by tensor voting. (e) Road network and original image superposition. True positives are shown in red, false positives in green and false negatives in blue.

(a) (b)

FIGURE 2.15: (a) The thinning algorithm extraction result. (b) The proposed method extraction result.

the proposed method does not produce spur and retains the road's smoothness. However, the result also indicates that the proposed method fails to handle the complicated junctions, such as the crossing circle.

2.3.4 Experiment 2

In the second experiment, the developed method was tested on another image of an urban area in Xuzhou. The image has a spatial dimension of 469×477. An image of the dataset and the hand-drawn reference map are presented in Figure 2.16(a) and 2.16(b), respectively. In this experiment, the optimal SVM

parameters were selected as: $C = 512$ and $\gamma = 0.5$. The spectral-spatial classification result is illustrated in Figure 2.16(c). The Local Geary's C map and its binary map are given in Figure 2.16(d) and Figure 2.16(e), respectively. The fusion result of spectral-spatial classification and Local Geary's C result is given in Figure 2.16(f).

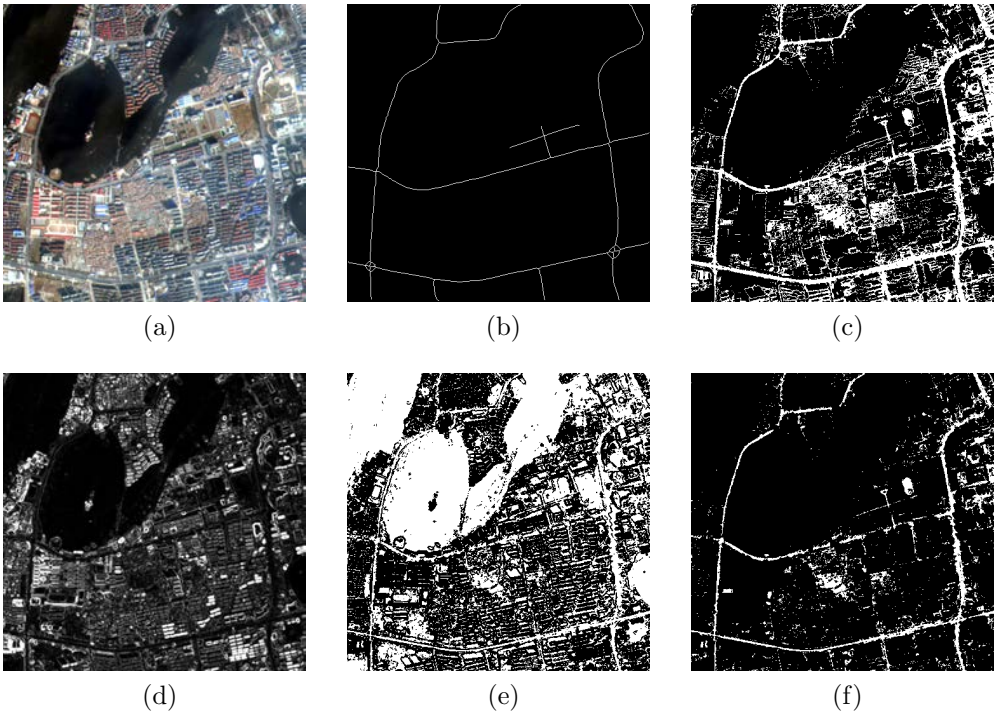


FIGURE 2.16: (a) The image of the study area. (b) Hand-drawn reference road map. (c) Classification result from SVM. (d) Example of Local Geary's C results. (e) Binary classification of Local Geary's C map. (f) Fusion result.

As for the previous experiment, the second order moment of each object was computed after performing information fusion. The second order moment distribution is shown in Figure 2.17(a). From Figure 2.17(a), it is seen that the second order moment value of road and non-road features is easily separated. In this experiment, the second order moment threshold is 0.45. The filtering algorithm with threshold 0.45 gave the result shown in Figure 2.17(b). The locally weighted regression algorithm was performed to extract the road centreline, resulting in Figure 2.17(c). The road network was then connected by the tensor voting method to eliminate gaps between the road segments. Figure 2.17(d) shows the road network generation result. The superposition

result of the original image and extracted road centreline is given in Figure 2.17(e). To evaluate the proposed method, the following three accuracy mea-

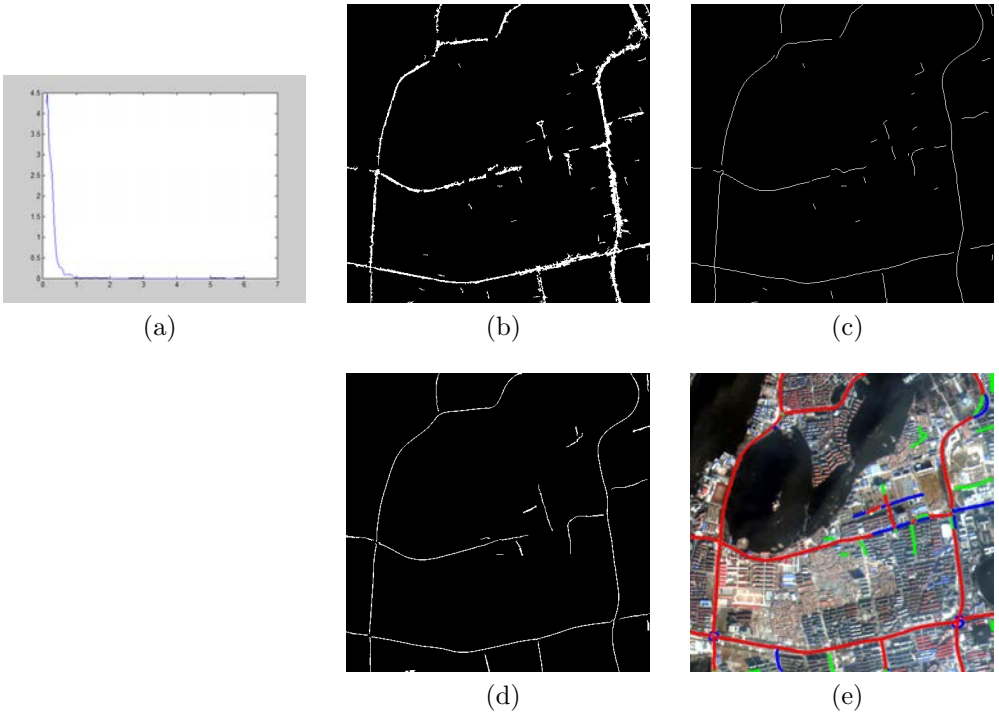


FIGURE 2.17: (a) Second order moment distribution. (b) Filtering result. (c) Road centreline extraction result. (d) Road network generation result by tensor voting. (e) Road network and original image superposition. True positives are shown in red, false positives in green and false negatives in blue.

sures proposed by Wiedemmann *et al.* (1998) were used in this study.

$$Completeness = \frac{TP}{TP + FN} \tag{2.20}$$

$$Correctness = \frac{TP}{TP + FP} \tag{2.21}$$

$$Quality = \frac{TP}{TP + FP + FN} \tag{2.22}$$

where TP, FN and FP represent true positive, false negative and false positive, respectively.

Table 2.3 shows the proposed method performance. Mayer *et al.* (2006) claim that the achievement of a completeness of at least 60% and a correctness of at least 75% are the absolute minimum for road extraction results, before

they can be considered useful in practice. From Table 2.3, it is seen that the proposed method in this study achieves these goals.

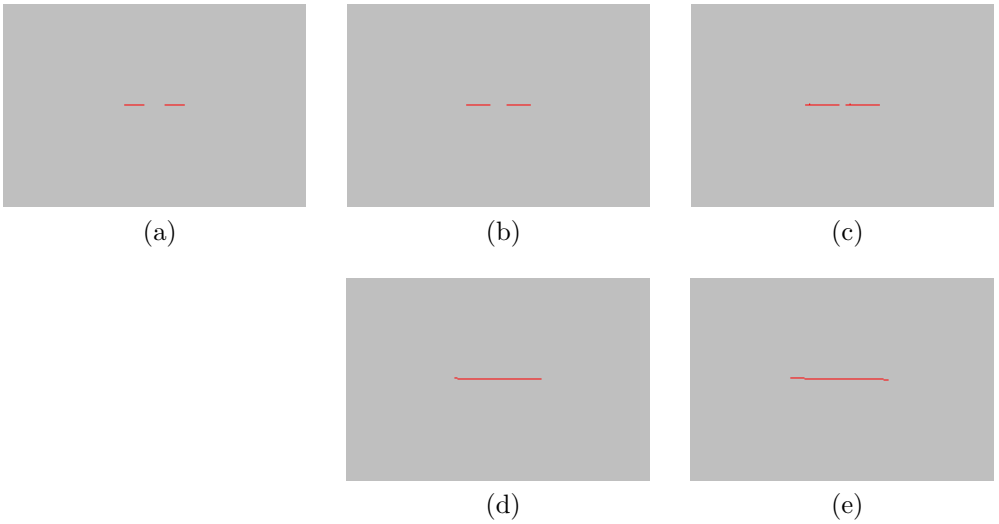
TABLE 2.3: The performance of the proposed method

Experiment	Completeness (%)	Correctness (%)	Quality (%)
1	94.03	98.1	92.34
2	85.24	83.38	72.86

2.3.5 Tests on scale parameter

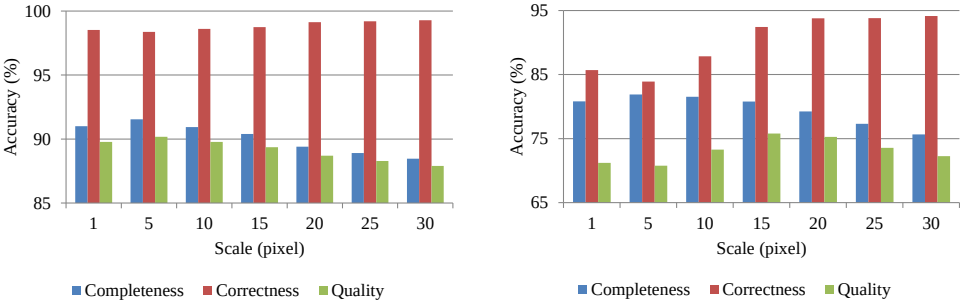
Tensor voting is used in the proposed method to eliminate the road discontinuity caused by image noise, such as shadows, trees and other obstructions. Only one parameter needs to be set by the users: scale parameter σ. In this section, the sensitivity of the proposed method to variations of the scale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

First, a simulation image was used to test the influence of the scale parameter setting on the connection result. Figure 2.18(a) shows an image with two disconnected line segments with a 20 pixels gap. Figure 2.18(b)-Figure 2.18(e) show the connection results using different scale parameters. As seen from the results, the gap fails to connect when σ is too small (e.g. $\sigma = 1$). When σ is progressively larger, the gap becomes increasingly smaller (e.g. $\sigma = 5$) until it ceases to exist (e.g. $\sigma = 10$). However, it should be noted that some undesired segments are produced when σ is very large (e.g. $\sigma = 20$). This is mainly because input tokens begin to heavily "cross talk" (Medioni et al., 2000). Second, two previous experimental road centrelines, extracted by locally weighted regression were chosen to test the scale parameter setting's influence on connection result. Figure 2.19 gives a quantitative evaluation result. As can be seen, the scale parameter is linearly related to correctness. This is because, as indicated above, the large scale parameter will eliminate most gaps, and hence produce a more correct road network. It is also seen that completeness and quality are firstly improved when σ is becomes larger



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

FIGURE 2.18: Tests on scale parameter selection. (a) Original image: the gap in the image is 20 pixels. (b) Connection result with $\sigma = 1$. (c) Connection result with $\sigma = 5$. (d) Connection result with $\sigma = 10$. (e) Connection result with $\sigma = 20$.



(a) (b)

FIGURE 2.19: Scale parameter influences on road network generation accuracy test. (a) Experiment 1. (b) Experiment 2.

(e.g. $1 \leq \sigma \leq 10$), and then decreases (e.g. $10 \leq \sigma \leq 30$). Again as indicated above, this is because many undesired road segments are produced when σ becomes larger. This phenomenon leads to a decrease of completeness and quality. Thus, in conclusion, a moderate scale parameter σ should be carefully selected to achieve a balance between correctness and completeness and quality.

2.3.6 Comparisons with the existing methods

In this section, the proposed method is compared with four existing road extraction methods from the literature. These four methods are introduced by Shi and Zhu (2002), Song and Civco (2004), Miao et al. (2013), and Huang and Zhang (2009). All methods were programmed and performed in MATLAB®[1]. Figure 2.20 gives the comparison results of different road extraction methods. To evaluate these methods quantitatively, completeness, correctness and quality of each method are computed. The results are presented in Table 2.4.

As is seen in Table 2.4, in terms of the completeness criteria, the proposed method gives the second best result after Song's method. Of all these methods, the proposed method achieves the best performance in terms of correctness and quality indices. From Figure 2.4, it is seen that the thinning algorithm used by Song and Huang's methods leads to low values of correctness and quality. Shi's method depends on a binary map obtained by simple thresholding. However, as urban imagery's complexity, such as image noise, material change, shadow, it is difficult to set a suitable threshold value. This leads to low accuracy in Shi's method.

The proposed method has also been tested on several Very High Resolution (VHR) remotely sensed imagery such as WorldView-2, QuickBird and Ikonos. The QuickBird dataset can be downloaded from VPLab (2015). The comparison results of the proposed method and several existing methods in the literature are presented in Figures 2.21-2.23. Figure 2.21 and Figure 2.22 show the results obtained using the road extraction methods on satellite images of developed urban areas, whereas Figure shows the results for urban/suburban areas. The quantitative evaluation results are presented in Table 2.5 – Table 2.7. Clearly, it can be claimed that the proposed method has some advantages compared to the existing methods in this field such as Shi and Zhu's method

[1]MATLAB (matrix laboratory), developed by MathWorks, is a multi-paradigm numerical computing environment and fourth-generation programming language.

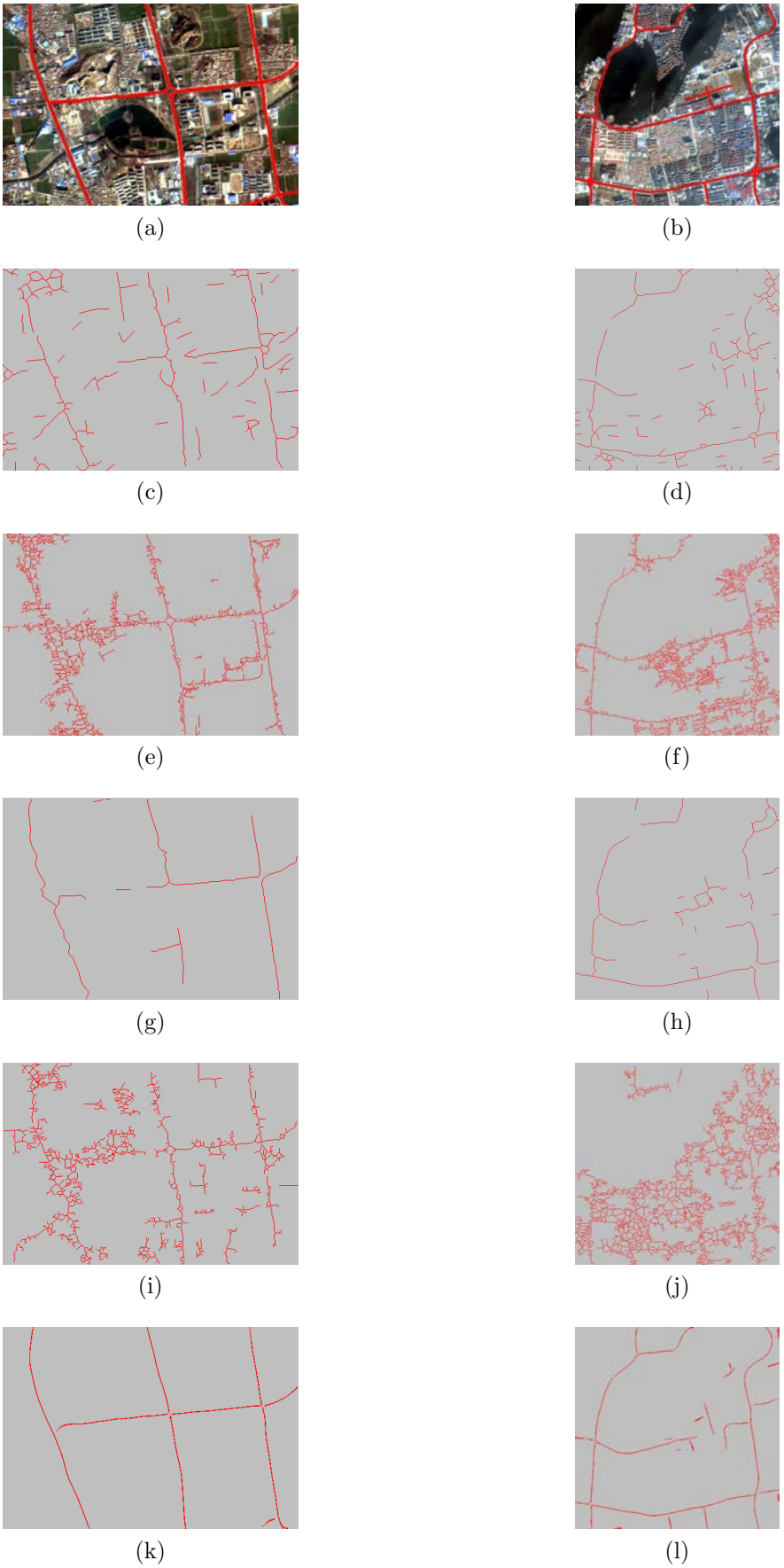


FIGURE 2.20: A comparison of different road extraction methods on Ziyuan-3 images. Row 1: Input images. The hand-drawn road reference map is shown in red. Row 2-6 show the results of Shi and Zhu's method, Song's method, Miao et al., Huang and Zhang's method, and the proposed method, respectively.

TABLE 2.4: The performance of different road extraction methods on Ziyuan-3 datasets

Sensors: Ziyuan-3						
	Study Area I			Study Area II		
Method	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)
Shi	79.55	37.56	34.26	67.01	45.58	37.23
Song	97.3	32.57	32.28	94.78	24.47	24.14
Miao	76.08	86.97	68.3	63.01	67.39	48.28
Huang	88.8	27.54	26.61	81.02	16.8	16.16
Proposed	94.03	98.1	92.34	85.24	83.38	72.86

(Shi and Zhu, 2002), Song's method (Song and Civco, 2004) and Huang's method (Huang and Zhang, 2009). The proposed method was also compared with other approaches shown in Mayer et al. (Mayer et al., 2006). The quantitative evaluation results are presented in Table 2.8. As can be seen from Table 2.8, the proposed method yields the moderate accuracy and the accuracy of the proposed method is close to the best.

2.4 Summary

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

Firstly, morphological profiles obtained from general adaptive neighbourhood mathematical morphology (GANMM) opening and closing have been used to implement spectral-spatial classification to extract initial road network. The experiments indicate that GANMM based spectral-spatial classification extracts initial road network with higher accuracy compared with other spectral-spatial classification methods. The classification and Local Geary's C results are then fused in such a way that it removes some misclassified roads and disconnects some misclassified roads from true roads. The fusion approach improves the robustness and accuracy of the roads extracted.

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

Thirdly, locally weighted regression is used to extract road centrelines from classified imagery. A major advantage of using regression is that it does not

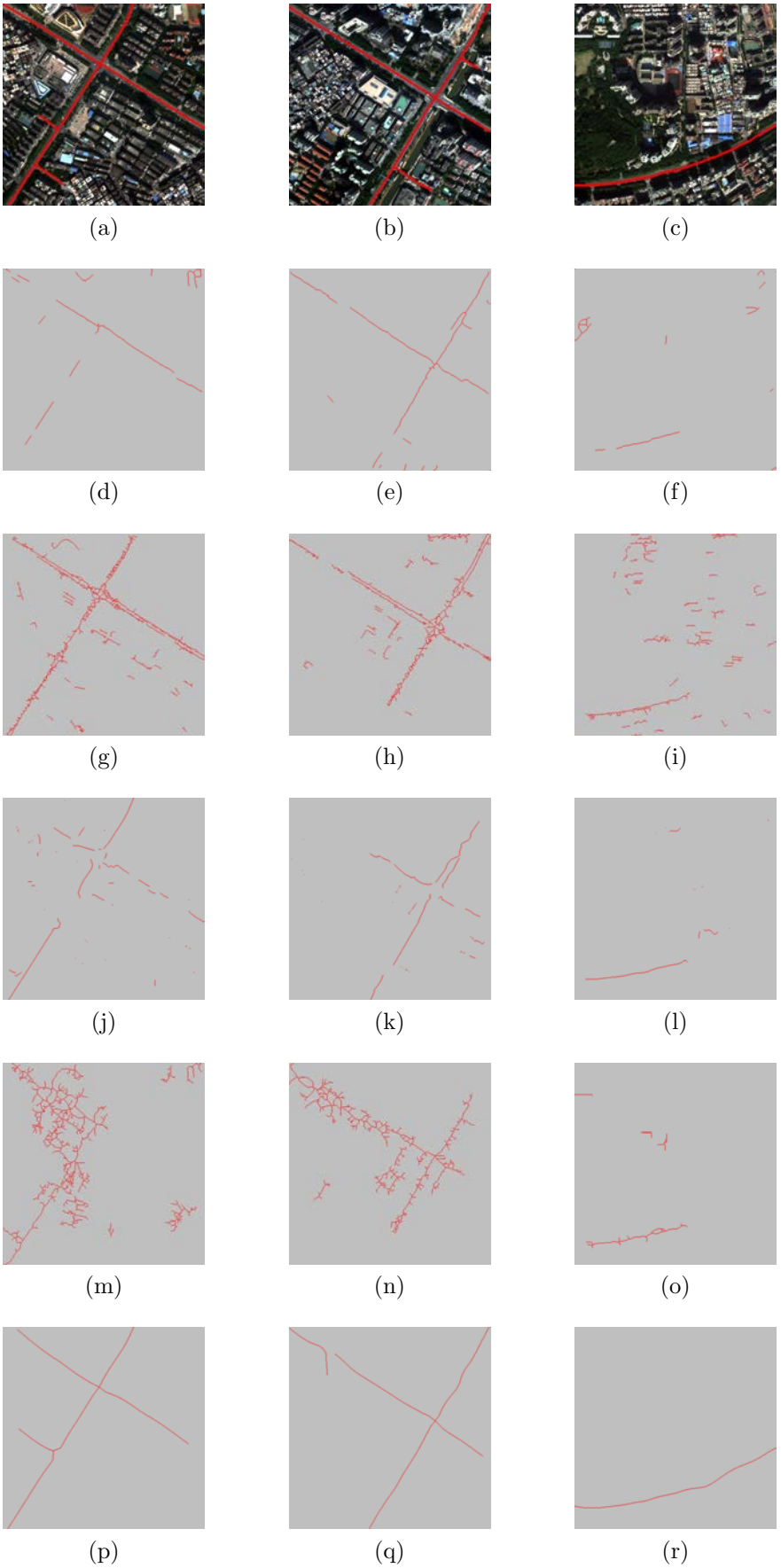


FIGURE 2.21: A comparison of different road extraction methods on Worldview-2 images. Row 1: Three input images of size (512×512). The road reference map is shown in red. Row 2-6 show the results of Shi and Zhu's method, Song's method, Miao et al., Huang and Zhang's method, and the proposed method, respectively.

TABLE 2.5: The performance of different road extraction methods on Worldview-2 datasets

Sensors: WorldView-2									
	Study Area I			Study Area II			Study Area III		
Method	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)
Shi	39.47	54.02	29.54	80.05	81.79	67.94	30.26	31.22	18.15
Song	93.34	48.6	46.97	81.95	41.88	38.34	59.01	18.05	16.04
Miao	47.52	55.84	34.54	40.72	55.91	30.82	51.01	73.46	43.07
Huang	52	14.79	13.01	68.92	30.42	26.75	50	48.71	32.75
Proposed	70	76.54	57.63	69.51	77.22	57.68	81.45	81.25	68.57

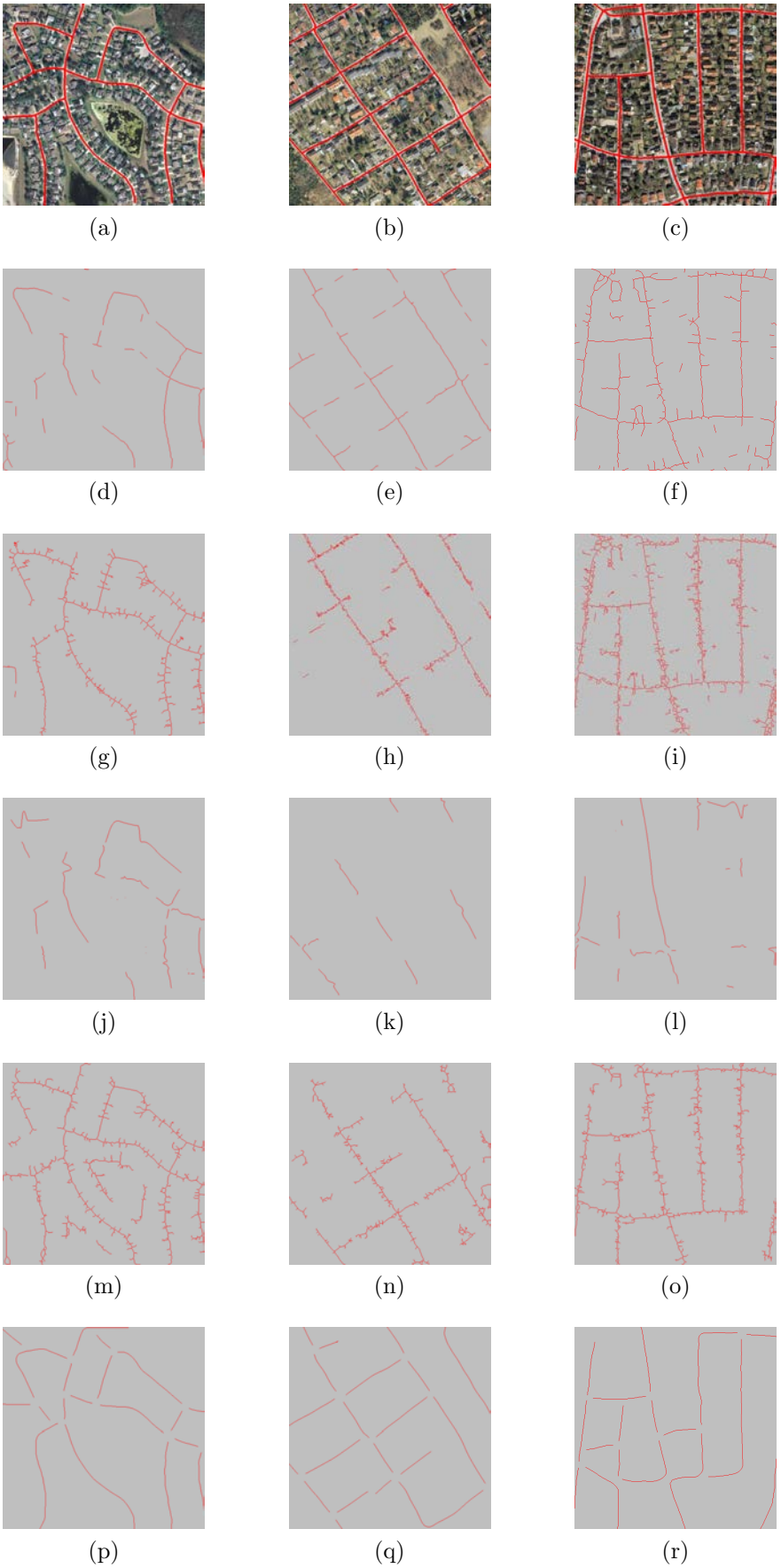


FIGURE 2.22: A comparison of different road extraction methods on Quickbird images. Row 1: Three input images. The road reference map is shown in red. Row 2-6 show the results of Shi and Zhu's method, Song's method, Miao et al., Huang and Zhang's method, and the proposed method, respectively.

TABLE 2.6: The performance of different road extraction methods on Quickbird datasets

Sensors: QuickBird									
	Study Area I			Study Area II			Study Area III		
Method	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)
Shi	64.84	84.11	57.77	70.65	93.79	67.5	84.42	63.36	56.73
Song	90.73	57.75	54.54	72.64	71.91	56.59	91.3	50.62	48.29
Miao	55.51	84.36	50.33	22.06	96.91	21.9	33.67	87.4	32.11
Huang	94.94	50.07	48.77	70.04	64.01	50.25	83.98	61.89	55.35
Proposed	75.76	75.04	60.51	55.45	68.31	44.1	67.7	83.03	59.47

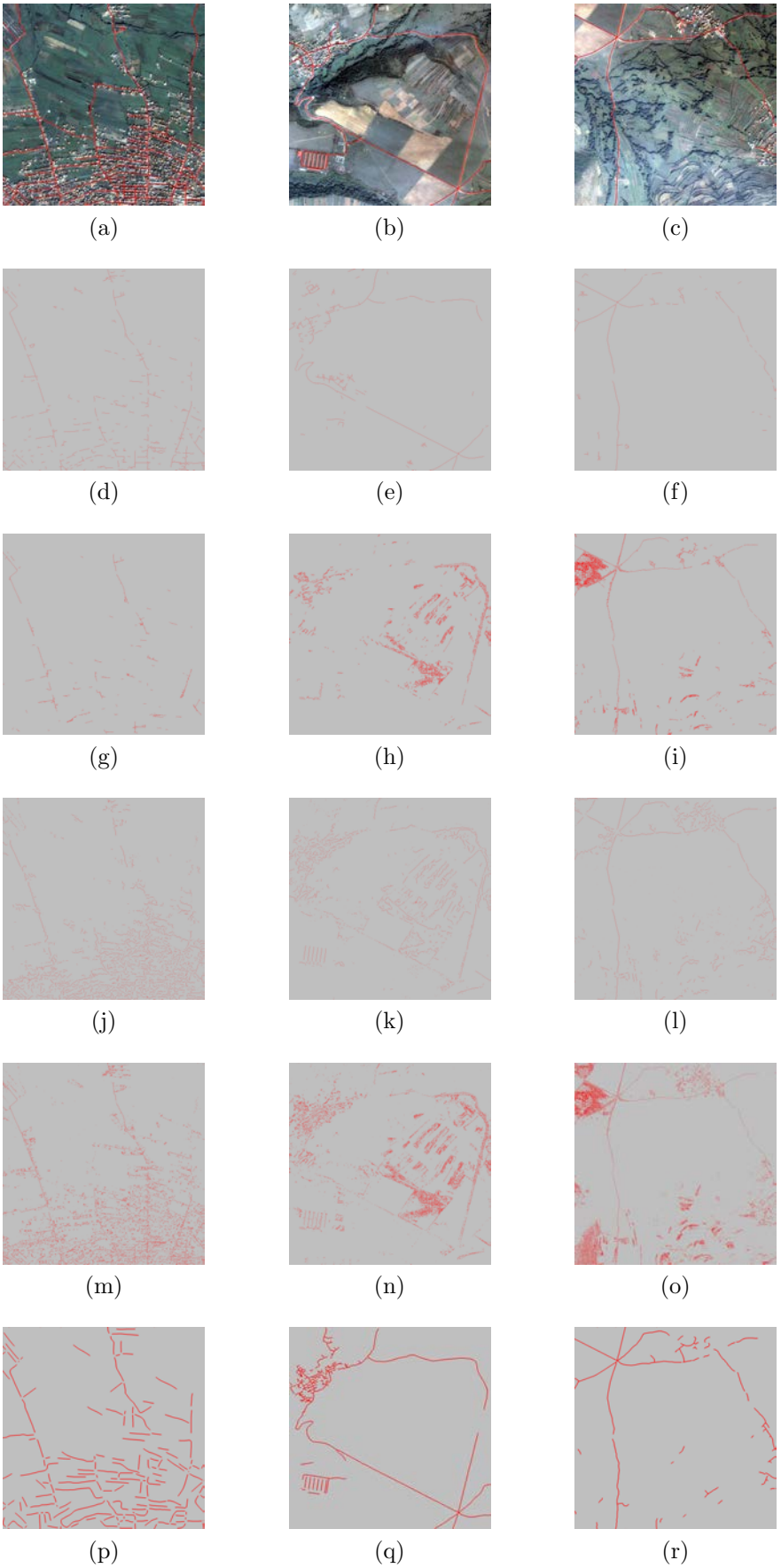


FIGURE 2.23: A comparison of different road extraction methods on Ikonos images. Row 1: Three input images. The road reference map is shown in red. Row 2-6 show the results of Shi and Zhu's method, Song's method, Miao et al., Huang and Zhang's method, and the proposed method, respectively.

TABLE 2.7: The performance of different road extraction methods on Ikonos datasets

Sensors: Ikonos									
	Ikonos1-Sub1			Ikonos3-Sub1			Ikonos3-Sub2		
Method	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)
Shi	32.79	46.67	23.86	46.76	56.98	34.56	48.49	74.96	41.73
Song	17.92	68.19	16.54	11.62	3.87	2.99	81.23	16.68	16.06
Miao	30.22	23.91	15.41	11.73	7.11	4.63	56.99	30.47	24.77
Huang	68.84	23.55	21.28	21.47	3.69	3.25	90.37	8.83	8.75
Proposed	34.29	63.39	28.63	76.87	64.68	54.14	93.65	61.54	59.08

TABLE 2.8: Results of the quantitative evaluation

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

produce spurs like thinning algorithm, and hence it retains the smoothness of road centrelines.

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

The proposed method has two parameters which need to be set by users: 1) homogeneity tolerance value and 2) scale parameter. The experimental results have been presented on two urban images to evaluate the proposed method which shows that the proposed method achieves a good performance for urban main road network extraction. However, experimental results show that the proposed method fails to handle complicated road junctions (i.e. circle junction). This is mainly because the simplification of road centreline extraction model based on locally weighted regression. A possible method to overcome this problem would be to use snake method to handle junctions. Another limitation of the proposed method is that it is not suitable for low-resolution imagery (less than 8-m resolution). This is because main roads just have 2-3 pixels and the local homogeneity of the grey values cannot be obtained. The third limitation of the proposed method is that it currently needs to train the SVM for each input image which limits the applicability of the method in practice. A possible solution to this limitation is to collect massive amounts of road training samples to perform the SVM training offline (Das et al., 2011) which does not require the interaction from users.

Chapter 3

The Object-Based Method for Road Extraction

3.1 Introduction

This chapter tries and goes further in the approach to road extraction from multispectral images by adopting the object-based image analysis (OBIA) paradigm (Blaschke, 2010). Compared to traditional pixel-wise methods, OBIA offers several advantages. For instance, pixel-wise methods for road extraction rely on information by a single pixel or its neighborhood. By contrast, OBIA not only uses spectral information, but also other spatial and spectral features derived from objects, such as shape and texture characteristics (Huang and Zhang, 2009, Shi et al., 2014a, Song and Civco, 2004). In general, OBIA achieves superior performances than pixel-wise methods when dealing with VHR satellite images, hence its attractiveness also in the field of road delineation (Blaschke et al., 2008). In line with this new trend, this work proposes an object-based method for automatic road delineation from VHR satellite images. To this end, two new filters have been designed to enhance road shape features and detect road networks in VHR satellite images, and

the recognized objects are then connected into a network using a novel active contour technique.

3.2 The proposed methodology

To extract road centerlines from VHR satellite images, the approach proposed in this chapter is based on the use of two object-based spatial filters and tensor voting. Figure 3.1 summarizes the main processing steps of the proposed method. As shown in this figure, the method mainly consists of the following

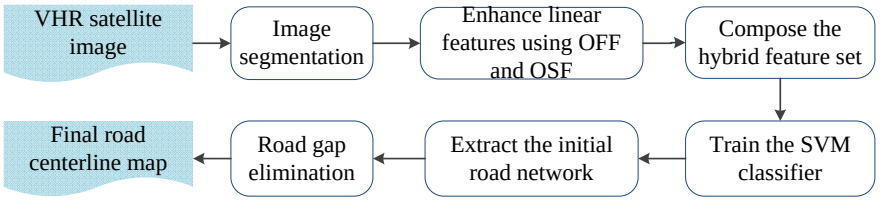


FIGURE 3.1: Flowchart of the proposed methodology.

five steps:

1. the original VHR satellite images is segmented into objects;
2. based on the extracted objects, object-based Frangi's filter (OFF) and object-based shape filter (OSF) are applied to enhance road features;
3. training samples are generated to train a support vector machine (SVM) classifier;
4. the initial road network is extracted by the trained SVM classifier;
5. tensor voting, active contour and the geometry information are integrated to connect road gaps and remove spurs to improve the road smoothness at the same time.

Details of each step are described in the following sections.

(a) (b) (c)

FIGURE 3.2: (a) The test image. (b) The ground reference data set. The road, grass, building, and bare soil are shown in red, green, blue, and brown, respectively. (c) Image segmentation result using ISODATA.

3.2.1 Image segmentation

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze (Gonzalez et al., 2009, Senthilnath et al., 2012). In Figure 3.2(a), we provide an example of a Pleiades-1a image with a spatial size of 400×400 pixels. The test image has four multispectral bands with 2m spatial resolution (the panchromatic band of Pleiades-1a was not used) and it will be used throughout the chapter to visualize the computation steps. The corresponding ground truth data set contains four classes, including road, grass, building, and bare soil (see Figure 3.2(b)). As visible in Figure 3.2(c), after segmentation obtained by a simple ISODATA algorithm (Ball and Hall, 1965), the original multispectral image is divided into many objects with different shapes. Starting from this preliminary result, in this study two filters are considered, aimed at enhancing road features in the VHR satellite images. These two filters are introduced in the later sections.

3.2.2 The object-based Frangi's filter

The central idea of the object-based Frangi's filter (OFF) is to compute the Frangi's filter (Frangi et al., 1998) response value in the object space, and not for individual pixels. Let Ω to be the image domain, $\mathbb{R}$ real value, and $I : \Omega \rightarrow \mathbb{R}, \Omega \subseteq \mathbb{R}^2$ an image. For a 2D image I, let $H(x, y)$ denote the Hessian matrix

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 I}{\partial x \partial x} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y \partial y} \end{bmatrix} \tag{3.1}$$

whose eigenvalues are $\lambda_i, i \in \{1, 2\}$ ordered such as $\lambda_1 \geq \lambda_2$, and $\vec{e_i}$ represent the corresponding eigenvectors. Frangi et al. (1998) proposed a multi-scale linearness measure that assigns a value between 0 and 1 to each point in the image. This value describes the confidence of a point being inside a linear region, so that 0 means that the pixel is definitely not located in a linear region while 1 has the opposite meaning. For a single scale σ_s, the filter response at a point $\boldsymbol{x}$ is calculated as

$$p_E(x, y) = \begin{cases} 0 & , \text{if } \lambda_1 \geq 0 \\ e^{-\frac{R^2}{2\beta^2}} \left(1 - 2e^{-\frac{S^2}{2c^2}}\right) & , \text{otherwise} \end{cases} \tag{3.2}$$

where $R = \left|\frac{\lambda_2}{\lambda_1}\right|$, $S = (\lambda_1^2 + \lambda_2^2)^{\frac{1}{2}}$, and β and c are constant normalization factors. R_b is essential to distinguish between plate-like and line-like structures since only in the line-like case it becomes zero. S is a measure of the "second order linearness". Note that value of S becomes low for homogeneous backgrounds, where eigenvalues tend to be small. The filter parameters, as suggested by Frangi, are set to $\beta = 0.5$ and $c = \frac{1}{2}\max\{S(x, y) \,|\, (x, y) \in \Omega\}$. The f_{Frangi} filter can be extended to the multi-scale processing. Given a set of scales Ω_{σ_S}, the multi-scale filter responses are combined as

$$E(\boldsymbol{x}) = \max\left\{f_{Frangi}^{\sigma_s} | \sigma_s \in \Omega_{\sigma_S}\right\} \tag{3.3}$$

where, for instance, $\Omega_{\sigma_S} = \{3, 5, 7, 9, 11\}$. For multispectral satellite images, responses are defined as

$$E(\boldsymbol{x}) = \frac{1}{B} \sum_{i=1}^{B} E^i(\boldsymbol{x}) \tag{3.4}$$

where B is the band number, and $E^i(\boldsymbol{x})$ is the $f_{Frangi}(\boldsymbol{x})$ result of the ith band.

Despite its popularity (Yuan et al., 2011), the f_{Frangi} filter suffers from the following two major drawbacks:

1. a locations offset due to edges and other non-linear structures can still have large response values (see for instance area A in Figure 3.3(a)); and
2. f_{Frangi} can not well enhance areas of junctions among linear features (see area B in Figure 3.3(a)), because they do not have a single specific direction.

To tackle these two drawbacks, in this chapter an object-based Frangi's filter (OFF) is introduced. The central idea of OFF is to compute the Frangi's filter in the object space, by assigning to each precomputed segment the average value of the output of the Frangi's filter for all the pixels which belongs to it. Figure 3.3(a) and 3.3(b) compare the results of the original f_{Frangi} filter and the new object-based version. It can be seen that the Frangi's filter fails in two cases: it produces low response values at road junction areas, whereas at land cover class boundary it produces high response values. By contrast, the OFF has amended these two limitations to some extent, and hence the new version is more effective to detect linear features.

3.2.3 The object-based shape filter

The object-based shape filter (OSF) is meant to measure the shape of each segment using its second order moment (Flusser, 2006). For an $m \times n$ image, its $(p+q)$ th order moment is defined as

$$m_{pq} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} x^p y^q I(\boldsymbol{x}), \tag{3.5}$$

where $I(\boldsymbol{x})$ is the grey value at the pixel $\boldsymbol{x}$, and p and q are non-negative integers. The corresponding central moment μ_{pq} is then given as

$$\mu_{pq} = \sum_{x} \sum_{y} (x - x^c)^p (y - y^c)^q I(\boldsymbol{x}), \tag{3.6}$$

where the coordinates (x^c, y^c) denote the centroid of $I(\boldsymbol{x})$, $x^c = \frac{m_{10}}{m_{00}}$ and $y^c = \frac{m_{01}}{m_{00}}$. Based on Equation 3.6, the normalized moment is described as: $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$, where $\gamma = \frac{p+q+2}{2}$. The sum of second order moments is computed as: $M = \eta_{02} + \eta_{20}$. The OSF computes the average M value for each connected component or segment.

An example of the output of OSF is given in Figure 3.3(c). As can be seen, the road and bare soil in Figure 3.2(a) have similar spectral values in the original data and it is quite challenging to distinguish between these two features using spectral information alone. By contrast, OSF results in Figure 3.3(c) is able to greatly differentiate the two scene elements. Indeed, roads have higher second order moment values than bare soil.

3.2.4 Road network extraction

After applying OFF and OSF, three features can be computed for each segment, grouped into the following vector:

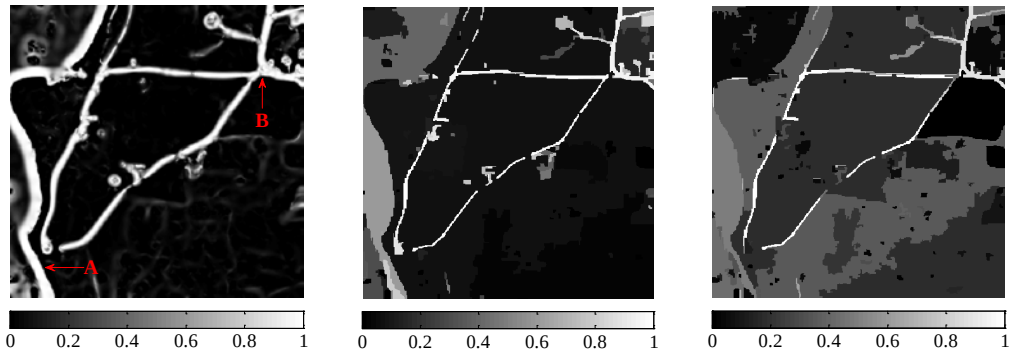


FIGURE 3.3: (a) The original filter f_{Frangi} result; (b) OFF result; and (c) OSF result. The value in this figure denotes the probability of road feature: the larger the value, the brighter the corresponding object, and the higher the probability that that object belongs to a road.

$$\mathrm{U}\left(\boldsymbol{x}^i\right) = \left\{\mathrm{SV}\left(\boldsymbol{x}^i\right), \mathrm{OFF}\left(\boldsymbol{x}^i\right), \mathrm{OSF}\left(\boldsymbol{x}^i\right)\right\} \tag{3.7}$$

where $\mathrm{SV}\left(\boldsymbol{x}^i\right) = \left\{\mathrm{SV}^b\left(\boldsymbol{x}^i\right)\right\}_{b=1}^{B}$ is the mean spectral value of all pixels $\boldsymbol{x}^i$ within the object $o^i (1 \leq N_o$ and is the number of objects), and B is the number of spectral bands. Here, $\mathrm{SV}\left(\boldsymbol{x}^i\right) = \frac{1}{N_{o^i}} \sum_{\boldsymbol{x} \in o^i} \mathrm{SV}^b\left(\boldsymbol{x}\right)$ where N_{o^i} is the pixel number of the object o^i.

After the construction of the hybrid feature vector $\mathrm{U}\left(\boldsymbol{x}\right)$, a support vector machine (SVM) (Hastie et al., 2009) classifier is subsequently applied. The rationale for selecting SVM is based on two considerations. First, SVM is a supervised classification method very robust to intensity changes caused, for instance, by similar yet different road materials. Second, previous studies showed that the SVM performance is as good as or significantly better than other competing methods in most cases (Shi et al., 2014a). Algorithm 2 gives the implementation detail of the SVM classification procedure. An example of the road classification using SVM is presented in Figure 5(a). Despite the SVM shows already an excellent performance, there are still some misclassified pixels, as well as discontinuities that need to be further processed.

Algorithm 2 The SVM classification algorithm

1: **Input**
2: The original satellite image
3: **Output**
4: The classified image.

1: For each pixel, compute the hybrid feature vector U ($\boldsymbol{x}$).
2: Randomly select 5% of ground truth data for each class as training samples.
3: Determine the parameters of the SVM Gaussian kernel, such as the kernel parameter γ and penalty parameter C, by five-fold cross validation.
4: Use the parameters selected at Step 3 to train the SVM classifier.
5: Classify the whole test image using the SVM classifier trained as in Step 4.

3.2.5 Reduction of the road gaps

To fill road discontinuities, in this work we exploit tensor voting, a powerful tool in the field of computer vision (Medioni et al., 2000, Mordohai and Medioni, 2006), as introduced in Chapter 2. After completing the tensor voting procedure, each pixel in the image is provided with a stickness probability computed by the tensor decomposition in Equation 2.18. An example of the vote result using tensor voting is illustrated in Figures 3.4(a)-3.4(b). It can be seen that pixels located on roads have large stickness values while pixels outside roads have small stickness values (see Figure 3.4(b)). It can be seen that tensor voting can be applied to the scene without any need to build road segments connection hypotheses in advance. Moreover, it depends on one parameter only (the scale parameter σ), the only one to be set by the user.

After the stickness computation using tensor voting, road features may be extracted via the local maxima algorithm. Specifically, a pixel is judged as a road pixel if the centroid of a searching window around it achieves the maximum value along the minimum axis $\vec{e_2}$. Unfortunately, it is observed that this method fails to extract junction areas. To avoid this drawback, this study relies on active contour (Lankton and Tannenbaum, 2008, Lankton, 2009) to extract road features from the stickness saliency map. Let B denote the characteristic function that identifies a local region.

$$B(x,y) = \begin{cases} 1 & ,\|x-y\| \leq r \\ 0 & ,\text{otherwise} \end{cases} \tag{3.8}$$

where r is the radius parameter. Let's define an energy function as follows:

$$E(\phi) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} B(x,y) F(I,\phi,x,y)\,\mathrm{d}y\mathrm{d}x \tag{3.9}$$

where $F = -\left(\mu_{\text{in}}(x) - \mu_{\text{out}}(x)\right)^2$ is the "internal" energy measure, and μ_{in} and μ_{out} are the intensity mean values of the foreground and background, respectively. Active contours are lines that evolve according to the following equation:

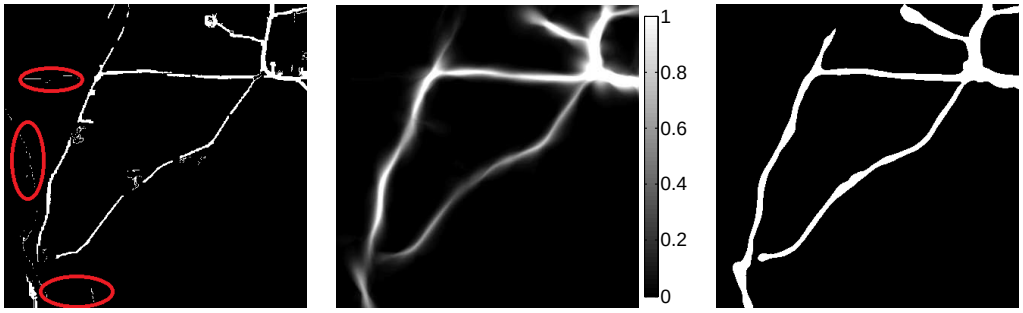
$$\frac{\partial\phi}{\partial t}(x) = \delta\phi(x) \int_{\Omega_y} G\mathrm{d}y + \lambda\delta\phi(x)\,\mathrm{div}\left(\frac{\nabla\phi(x)}{|\nabla\phi(y)|}\right)\|\nabla\phi(x)\| \tag{3.10}$$

where $G = B(x,y)\,\delta\phi(y)\left(\mu_{\text{in}}(x) - \mu_{\text{out}}(x)\right)\left(\frac{(I(y)-\mu_{\text{in}}(x))^2}{A_{\text{in}}(x)} - \frac{(I(y)-\mu_{\text{in}}(x))^2}{A_{\text{out}}(x)}\right)$, λ serves as the penalty term to keep the curve smooth, and $A_{\text{in}}(x)$ and $A_{\text{out}}(x)$ are the areas of the local interior and local exterior regions, respectively given by

$$A_{\text{in}} = \int_{\Omega_y} B(x,y) H\phi(y)\,\mathrm{d}y \tag{3.11}$$

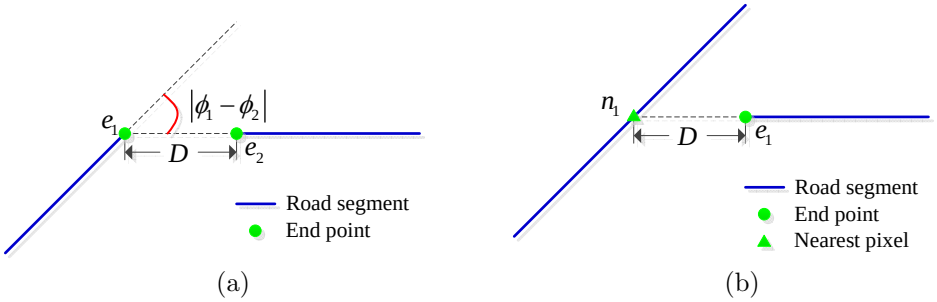
$$A_{\text{out}} = \int_{\Omega_y} B(x,y)\left(1 - H\phi(y)\right)\mathrm{d}y \tag{3.12}$$

Although road networks extracted by the local maxima algorithm are unsatisfactory, they can be used as the starting positions for the subsequent active contour algorithm. An example of road extraction using active contours is illustrated by Figure 3.4(c). It can be seen that 1) all discontinuities are eliminated; and 2) most spurs are also removed, thus improving the road smoothness. By means of this approach, it is also worth mentioning that some misclassified road pixels are removed. This is visible in Figure 3.4(c),



(a) (b) (c)

FIGURE 3.4: (a) Road segments extracted using SVM, where pixels in the ellipsoid regions are misclassified as roads; (b) stickness saliency map; and (c) using active contours.



(a) (b)

FIGURE 3.5: Gap filling using geometry information: (a) Two end points fulfilling the conditions to be connected; (b) an end point and an internal segment point fulfilling the conditions to be connected.

where misclassified road pixels within the areas highlighted by ellipses in Figure 3.4(a) are removed by using the active contours algorithm based on the saliency map. Unfortunately, Figure 3.4(c) shows that some correct road pixels are removed, too, because of their weak stickness saliency values. Despite this limitation, active contours represent a major advance with respect to road network extraction from the stickness saliency map. Figure 3.4(b) indicates that the tensor voting procedure tends to blur the road boundary, hence reducing the accuracy of the extracted road objects. To tackle this problem, this chapter jointly uses the road geometry information to solve the connectivity problem rather than using the tensor voting alone. However, it should be pointed out that the tensor voting result is more accurate than the original result (see Figure 3.4(a) and Figure 3.4(c)) and thus makes it possible to

use the road geometry information to further eliminate road gap to improve road extraction accuracy. In this work, we exploit two geometric properties to capture the relations between two parts of the road segment, as illustrated by Figure 3.5. They are:

1. Edge direction similarity. To obtain smooth reconstructions, we consider the direction of each pair of consecutive edges. The angular difference between their direction should satisfy the following condition:

$$|\phi_1 - \phi_2| \leq \frac{\pi}{2} \tag{3.13}$$

2. Gap length. The distance between any pair of end points should not exceed a given threshold σ.

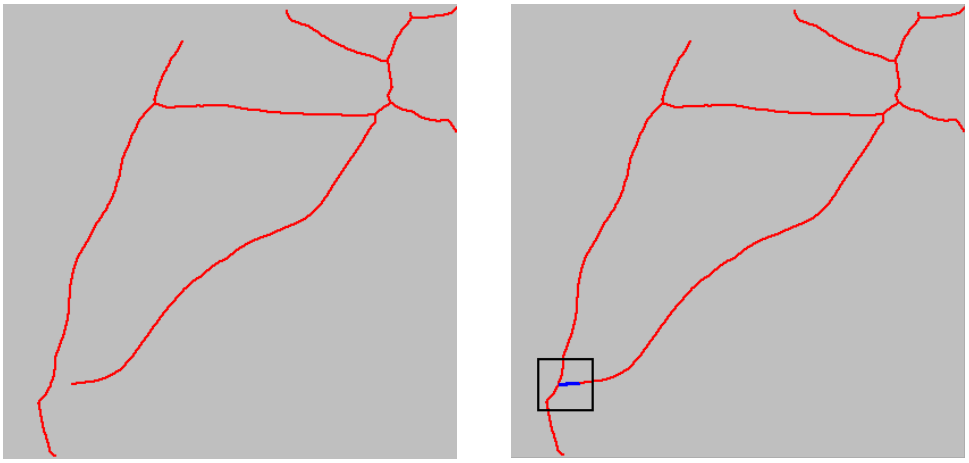
$$D\left(e_1, e_2\right) \leq \sigma \tag{3.14}$$

Algorithm 3 Gap reduction using the geometry information

Input
The original road centerline map
Output
The result of road gap elimination.

Extract the road centerline from the active contour result according to the method proposed in (Miao et al., 2014a).
Detect end points from the road centerline result. For any pair of end points, if they satisfy conditions of edge direction similarity and gap length, then connect these two points using the geodesic method (Miao et al., 2014a). If an end point cannot find another end point to be connected, then find whether an internal segment point that satisfies the condition of gap length. Otherwise, go to the next end point.
Repeat Step 2 until all end points are analyzed.

In this chapter, we consider two road gap cases, as shown in Figure 3.5. The first case (see Figure 3.5(a)) is that an end point can find another end point to which it can be connected, and which satisfies conditions of edge direction similarity and gap length. Another case is that an end point cannot find a second end point that satisfies two aforementioned conditions. However, there



(a) (b)

FIGURE 3.6: An example of road gap detection and filling using geometry information. (a) The centerline result of road segments presented in Figure 3.4(c). (b) The road gap connection result using geometry information.

is a point near to another segment that satisfies the condition of gap length. Point pairs that satisfy these conditions will be connected by the geodesic method (Miao and Shi, 2014), that can adaptively link two points without any predefined connection hypothesis. Algorithm 3 presents the details of the gap elimination procedure using the above mentioned geometry information. Figure 3.6 shows an example of the road gap detection and filling using this approach. It can be seen that the geometry information, as a complement of tensor voting, can further improve the accuracy of the tensor voting and active contour result (see the rectangle area in Figure 3.6(b)).

3.3 Experimental results

In this section, several experiments are described in order to test the proposed method. In the last test, our method is also compared with other methods available in literature to show its advantages and limitations.

3.3.1 Tests on different segmentation methods

The proposed OFF and OSF filters depend on objects features, and thus on the image segmentation result. Hence, the sensitivity of the proposed approach to the selected image segmentation method has been checked first. To this aim, this study compares five segmentation methods, including 1) Expectation Maximization (EM) (Dempster et al., 1977), 2) Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA) (Ball and Hall, 1965), 3) Statistical Region Merging (SRM) (Nock and Nielsen, 2004), 4) mean shift (Comaniciu and Meer, 2002), and 5) fuzzy c-clustering (Nock and Nielsen, 2006). For quantitative comparison, three accuracy measures are used, including: 1) Completeness, 2) Correctness, and 3) Quality (Wiedemann et al., 1998). Please note that the ground truth datasets are obtained by visually digitizing the roads from original satellite images. The comparison of the results for different segmentation methods are reported in Table 3.1. The proposed road extraction technique shows a similar performance using different segmentation methods, resulting into a good insensitiveness to the selection of image segmentation method. However, the results in Table 3.1 suggest that SRM achieves the best balance between Completeness and Correctness, making it the method of choice from now on.

TABLE 3.1: Results of different segmentation methods

Segmentation method	Completeness	Correctness	Quality
EM	0.9	0.83	0.76
ISODATA	0.92	0.82	0.76
SRM	0.9	0.85	0.78
Mean shift	0.93	0.81	0.77
Fuzzy c-clustering	0.86	0.86	0.75

3.3.2 Tests on system parameters

The image segmentation step is a fundamental and core step of the proposed method. Any segmentation error caused by under segmentation or over segmentation will have an impact on OFF and OSF, which in turn affects the final extraction accuracy. To examine the segmentation influence, the region number in SRM is changed from 32 to 2048 with a factor of 2. The experimental results show that the road extraction accuracy decreases with the region number (see Figure 3.7(a)). In other words, over segmentation negatively affects the performance of the proposed method. The steady decline in the road extraction accuracy can be due to an increasing difficulty to measure the linear features using OFF and OSF under the over segmentation case. Thus, this study sets the region number to 32 to achieve the best performance.

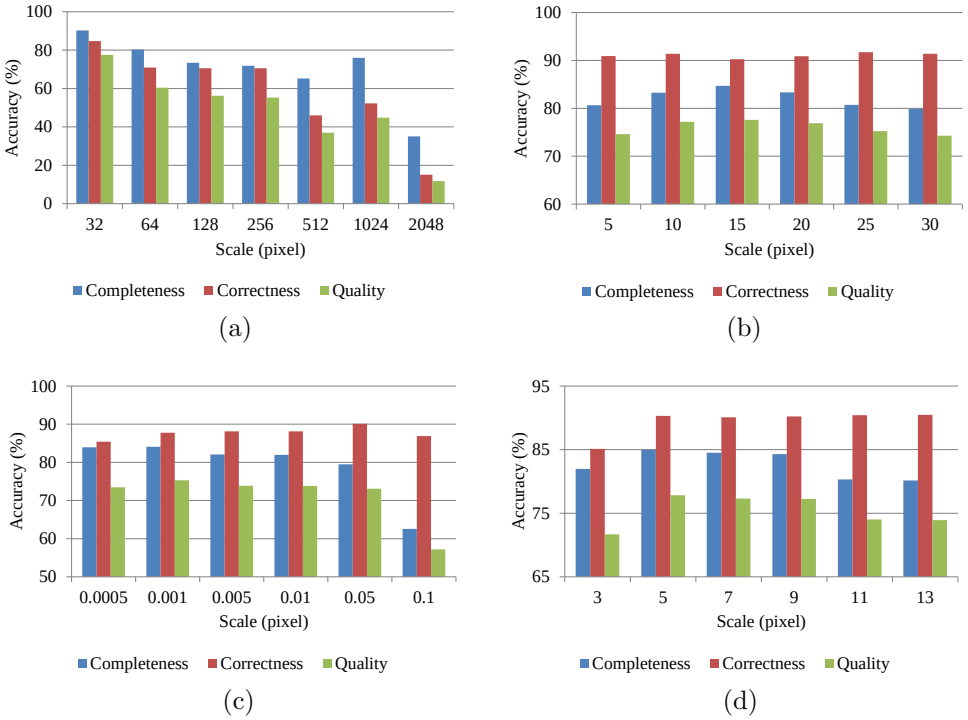


FIGURE 3.7: The performance curves for each of the tested parameters.

Secondly, to test the influence of the scale parameter in the tensor voting processing step, its value was adjusted from 5 to 30 pixels with an increase of 5 pixels each step. The test results are presented in Figure 3.7(b). It can be

seen that, with an increase in the scale parameter σ, there is a slight increase in the Correctness value, which indicates that σ has a weak but positive effect on Correctness. Instead, Completeness and Quality values increase first when σ increases from 5 to 15, but then decrease when σ exceeds 15. The reason for this phenomenon is that a large scale parameter makes pixels 'cross-talk' during tensor voting and in turn produces un-desired branches (i.e. spurs) leading to a decrease in Completeness and Quality. Therefore, one must find a compromise between Correctness and Completeness (i.e. the best Quality). In our experience, tensor voting (TV) produces the best Quality result with a σ value around a value of 15 pixels, and this number will be used throughout the following tests.

The third set of tests to select parameter values involved the length term of active contour. This term value was changed from 0.0005 to 0.1, and its effects on the final road network extraction performance were observed. As it can be seen from Figure 3.7(c), a change in the length term has frail influence on the Correctness. When the length term is increased from 0.0005 to 0.05, the Completeness and Quality values changed slightly. However, when the length term is greater than 0.05, the Completeness and Quality values fell drastically. This is because small length term produces results with spurs, while large length term produces smooth results that remove spurs as well as some actual roads. This shows that a length term smaller than 0.05 has weak influence on the Completeness and Quality, whereas a length term greater than 0.05 has significant and negative impact. Therefore, to achieve the best performance, the length term is selected 0.001 in this study.

Finally, the fourth test examines the impact of the active contour radius parameter. This parameter was changed from 3 to 13 pixels, and the quantitative evaluation result is reported in Figure 3.7(d). Completeness values gradually decrease when the radius is increased. Instead, Correctness and Quality increase first, when the radius is changed from 3 to 5, but when the value is between 5 and 9, Correctness and Quality gradually decrease. Eventually,

when the radius length exceeds 9, Correctness and Quality experience a notable decrease. This is largely due to the fact that with an increase in the radius value some weak stickness saliency areas are regarded as the background. Therefore, the radius parameter was set to 5 pixels to achieve the best balance between the Completeness and Correctness metrics.

3.3.3 Test areas

To test the proposed method, it was tested first on three samples of an Ikonos image of the urban area of Hobart (Australia). The size of these three images, shown in Figure 3.8(a)-3.8c, is 400 × 400 pixels. The test image has four multispectral bands with the spatial resolution of 2m per pixel and the proposed method is validated on the original multispectral bands without pan-sharpening. The manually extracted ground truth data sets are shown in red. The SVM classifier was firstly trained based on the training samples selected and then used to classify the three test images. In this chapter, the LIBSVM library (Chang and Lin, 2011) was used, and the optimal parameters selected were $C = 64$ and $\gamma = 0.0625$. To assess the impact of the proposed method on the results of optical satellite image road extraction, a comparison was carried out with three state-of-the-art methods: 1) the multivariate adaptive regression splines based road centerline extraction method (MARS_RCE) (Miao et al., 2013), 2) the geodesic method based road centerline extraction method (G_RCE) (Miao and Shi, 2014), and 3) the principal curve based road centerline extraction method (PC_RCE) (Miao et al., 2014a). Figure 3.8 gives the comparison results of different road extraction methods. As it can be seen, road features are extracted well by the proposed method. This performance is also confirmed quantitatively by road extraction accuracy in terms of three measures (i.e. Completeness, Correctness, and Quality). Table 3.2 lists the quantitative evaluation results for the three test images. As shown in Table 3.2, the highest Quality values yielded by the four methods are equal to 0.98, 0.97, and 0.94, respectively.

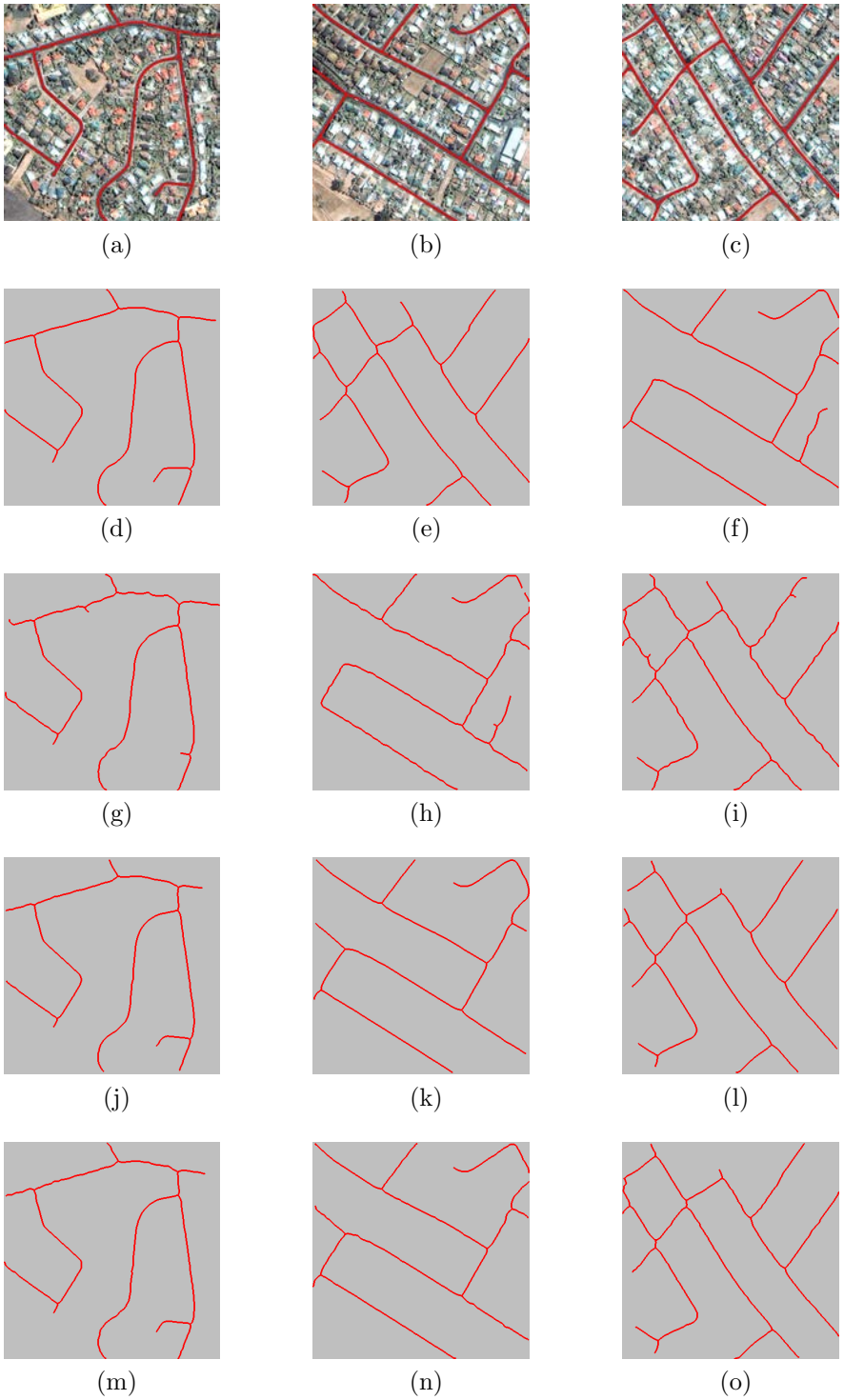


FIGURE 3.8: Road extraction results on three Ikonos images depicting portions of Hobart (Australia): the first row shows three test areas, while the second to fifth row show the corresponding road extraction results produced by MARS_RCE, G_RCE, PC_RCE, and the proposed method, respectively.

The proposed method outperforms MARS_RCE, G_RCE, and PC_RCE in the first and third cases. The reason why the accuracy of the second test area is

lower than the ones for the other two areas is mainly the imprecision in the classification. Indeed, some correct roads are discarded and some pixels are misclassified as road, as shown in Figure 3.8(e).

In a second experiment, the developed method was tested on three image of an urban area recorded by the QuickBird satellite. The test datasets were downloaded from a freely available image database assembled for road extraction purposes (VPLab, 2015). The three sample images have a spatial size of 512×512 pixels and the test image has four multispectral bands without pan-sharpening, and its spatial resolution is 2.4 m. Figure 3.9 presents the test images, while the corresponding ground truth vectors are shown in red. To assess the results of the presented method, in this area, a comparison similar to the previous one was carried out, using the three above mentioned methods (MARS_RCE, G_RCE, PC_RCE). Experimental results show that road extraction maps obtained by the proposed method show fewer road gaps, confirming the necessity of incorporating tensor voting and the geometrical information about road features to discard discontinuities. Table 3.3 reports the quantitative evaluation results, reporting Quality measures for of 0.74, 0.91, and 0.90, respectively. The reason for this unsatisfactory result for the first area is due to the shadows, jeopardizing road spatial features. The quantitative analysis however indicates that the proposed method outperforms MARS_RCE, G_RCE, and PC_RCE, which confirms its reliability.

Finally, a third experiment was conducted using the dataset acquired by Ikonos over the urban area of Moonah (Australia). The scenes (each one of 1000 pixels by 1000 pixels) were selected as test areas. Figure 3.10 shows these images and their corresponding ground truth datasets generated by visual interpretation. As for previous experiments, the proposed approach was compared with MARS_RCE, G_RCE, and PC_RCE, and the results are reported in Table 3.4. Once again, the proposed method outperforms the other three methods in term of the Quality metric. This shows that the proposed method achieves the best balance between Completeness and Correctness.

TABLE 3.2: Quantitative results for three Ikonos samples depicting portions of Hobart (Australia)

Method	MARS_RCE			G_RCE		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.97	0.99	0.96	0.82	0.85	0.72
II	0.98	0.99	0.97	0.89	0.89	0.8
III	0.95	0.98	0.93	0.94	0.94	0.89
Method	PC_RCE			The proposed method		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.94	0.99	0.94	0.99	0.99	0.98
II	0.96	0.99	0.95	0.94	0.93	0.88
III	0.91	0.98	0.9	0.96	0.98	0.94

TABLE 3.3: Quantitative results for three Quickbird test images VPLab (2015).

Method	MARS_RCE			G_RCE		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.77	0.93	0.72	0.61	0.91	0.57
II	0.97	0.91	0.89	0.86	0.95	0.82
III	0.94	0.83	0.79	0.81	0.88	0.73
Method	PC_RCE			The proposed method		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.61	0.92	0.58	0.79	0.92	0.74
II	0.84	0.96	0.82	0.93	0.98	0.91
III	0.82	0.92	0.76	0.95	0.94	0.9

TABLE 3.4: Quantitative results for three Ikonos samples of Moonah (Australia).

Method	MARS_RCE			G_RCE		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.87	0.72	0.65	0.86	0.86	0.75
II	0.84	0.75	0.66	0.88	0.91	0.8
III	0.56	0.91	0.54	0.99	0.58	0.57
Method	PC_RCE			The proposed method		
Study Area	Completeness	Correctness	Quality	Completeness	Correctness	Quality
I	0.93	0.83	0.79	0.94	0.92	0.87
II	0.82	0.95	0.79	0.82	0.97	0.81
III	0.99	0.62	0.62	0.86	0.83	0.73

(a) (b) (c)

FIGURE 3.9: Road extraction results on three QuickBird test samples (VPLab, 2015): (a)-(c) show the three test areas, respectively.

(a) (b) (c)

FIGURE 3.10: Road extraction results for three test areas extracted from an Ikonos image of Moonah (Australia).

3.4 Summary

In this chapter an automatic road extraction method has been presented to be applied to extract road network from Very High Resolution (VHR) satellite images. The proposed method is based on objects derived from image segmentation. Two filters, namely the object-based Frangi's filter (OFF) and the object-based shape filter (OSF), were designed to enhance road features. Once road candidates are selected, a tensor voting method followed by an active contour approach exploiting geometry information help to connect road by filling their gaps without any road connection hypothesis.

Experimental results and comparisons show that the proposed method is able to extract road networks with a Completeness and Quality values comparable

to those by state-of-the-art techniques. However, a few road segments remain disconnected in the final results. This is largely because these segments have large road gaps. The next step would be to iteratively apply tensor voting with a small scale parameter. Also, object-based filters using more road features, such as road width and edge information, will be considered in future steps of this work. The third limitation of the proposed method is that it currently needs to train the SVM case by case which limits its practical applicability. A possible solution to automate SVM parameter training is to collect massive amounts of road training samples (Das et al., 2011) to perform the SVM training offline that does not require the interaction from users.

Chapter 4

Fusion of Multiple Road Extraction Results Using Geometric Characteristics

4.1 Introduction

Previous chapters are mostly focusing on road extraction in rural/suburb areas. However, road network extraction in urban environment is a much more challenging task because 1) road network in urban environment is generally dense, 2) many non-road pixels, such as car parking, 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. Thus, this chapter 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; 2) dense urban environment. This chapter also examines 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.

4.2 The proposed methodology

The objective of this chapter is to design a computationally efficient approach to extract accurate urban road networks from VHR optical satellite images. Figure 4.1 graphically summarizes the main steps of the proposed method. The basic elements are a clustering method and a linearness filtering, implemented in a multi-scale 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 sub-sections.

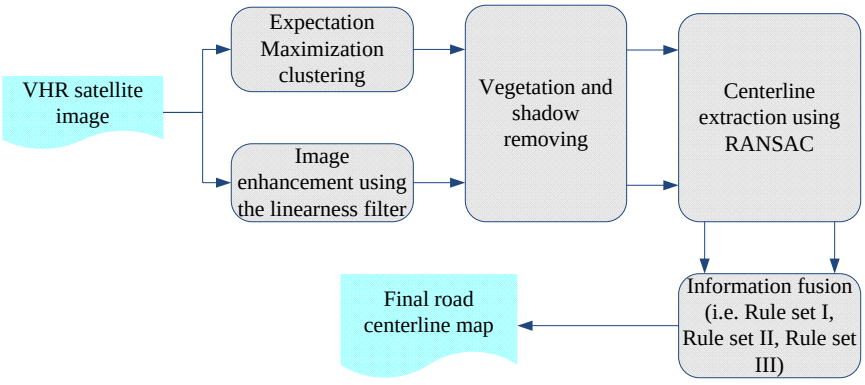


FIGURE 4.1: Schematic flowchart of the proposed method.

4.2.1 Expectation maximization clustering

With the aid of the segmentation technology, the input image can be subdivided into non-overlapping regions. To each of these regions a set of parameters is assigned, e.g. the average gray and texture levels, its shape parameters,

its size, etc. 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 chapter, Expectation Maximization (EM) clustering, one of the most powerful segmentation algorithms, proposed by Dempster *et al.* (1977), 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.

$$p\left(x; \mu_c, \Sigma_c\right) = \sum_{c=1}^{C} w_c p_c\left(x\right) \tag{4.1}$$

$$p_c\left(x\right) = \frac{1}{\left(2\pi\right)^{\frac{d}{2}} \left|\Sigma_c\right|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\left(x-\mu_c\right)^T \Sigma_c^{-1}\left(x-\mu_c\right)\right\} \tag{4.2}$$

where C is the cluster number, p_c is the Gaussian distribution density with mean μ_c and covariance matrix Σ_c, 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:

$$\mu_c^i = \frac{1}{w_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} \left(x_{j.c}^{i-1}\right) \tag{4.3}$$

$$\Sigma_c^i = \frac{1}{w_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} \left(x_{j.c}^{i-1} - \mu_c^i\right)\left(x_{j.c}^{i-1} - \mu_c^i\right)^T \tag{4.4}$$

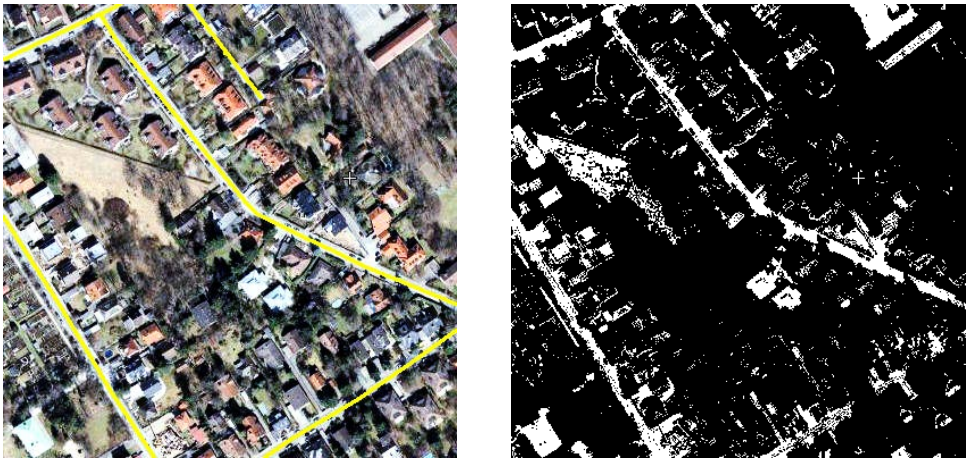
$$w_c^i = \frac{m_c^{i-1}}{N} \tag{4.5}$$

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

$$x_j \in Q_c^i :| \Pr\left(c \mid x_j\right) = \max_l \Pr\left(l \mid x_j\right) \tag{4.6}$$

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 Figure



(a) (b)

FIGURE 4.2: (a) True color image of QuickBird image. The ground truth dataset, obtained by visual interpretation, is shown in yellow; (b) EM clustering result.

4.2(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 (VPLab, 2015). The spatial size of the test image is 400 pixels by 400 pixels and the test image has four multispectral bands with the spatial resolution of 2m per pixel. This image will be used throughout the chapter to visualize intermediate results. Image segmentation results are thus presented in Figure 4.2(b).

4.2.2 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 chapter presents a new linearness filter that combines the Hessian matrix, the local standard variance, and geometrical features.

4.2.2.1 Hessian matrix based filter

In the field of image processing, Hessian matrix (Gonzalez et al., 2009) is an important tool that has various applications, because it is able to capture the

local image structure (Lindeberg, 1998). Let us define $I : \Omega \to \mathbb{R}, \Omega \subseteq \mathbb{R}^2$ a 2D gray 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:

$$H(x,y) = \begin{bmatrix} G_\sigma * \frac{\partial^2 I}{\partial x \partial x} & G_\sigma * \frac{\partial^2 I}{\partial x \partial y} \\ G_\sigma * \frac{\partial^2 I}{\partial y \partial x} & G_\sigma * \frac{\partial^2 I}{\partial y \partial y} \end{bmatrix} \tag{4.7}$$

where G_σ is a Gaussian kernel centered at (x, y) with band width equal to σ. Eigen-decomposition (Press, 2007) is subsequently applied to the Hessian matrix H, resulting in:

$$\begin{aligned} H &= \begin{bmatrix} \overrightarrow{e_1} & \overrightarrow{e_2} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \overrightarrow{e_1} \\ \overrightarrow{e_2} \end{bmatrix} \\ &= \lambda_1 \overrightarrow{e_1} \overrightarrow{e_1}^T + \lambda_2 \overrightarrow{e_2} \overrightarrow{e_2}^T \end{aligned} \tag{4.8}$$

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:

$$\begin{cases} R &= \left| \frac{\lambda_2}{\lambda_1} \right| \\ S &= (\lambda_1^2 + \lambda_2^2)^{\frac{1}{2}} \end{cases} \tag{4.9}$$

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 linearness". 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:

$$p_E(x,y) = \begin{cases} 0 & , \text{if } \lambda_1 \geq 0 \\ e^{-\frac{R^2}{2\beta^2}} \left(1 - e^{-\frac{S^2}{2c^2}} \right) & , \text{otherwise} \end{cases} \tag{4.10}$$

where $\beta = 0.5$ and $c = (1/2) \max \{ S(x) | x \in \Omega \}$ as suggested by Frangi *et al.* (1998). Please 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 (Yuan et al., 2011): 1) edges and similar high-contrast quasi-linear structures can still have large likelihood values, resulting in false positives; and 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 in order 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.

4.2.2.2 The local standard variance and the circularity index

By regarding a local window with the size σ, equivalent to the Gaussian kernel bandwidth in Equation 4.7, the mean radiance $\bar{r}_i(x, y, \sigma, \theta)$ of the ith spectral band along the direction identified by an angle θ can be expressed as:

$$\bar{r}_i(x, y, \sigma, \theta) = \frac{1}{2\sigma} \sum_{k=-\sigma}^{\sigma} V_I \tag{4.11}$$

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:

$$s_i(x, y, \sigma, \theta) = \left(\frac{1}{2\sigma - 1} \sum_{k=-\sigma}^{\sigma} (V_I - \bar{r}_i(x, y, \sigma, \theta))^2 \right)^{\frac{1}{2}} \tag{4.12}$$

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

$$tSD\,(x, y, \sigma, \theta) = \sum_{i=1}^{B} s_i\,(x, y, \sigma, \theta) \tag{4.13}$$

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:

$$\begin{cases} tSD_o\,(x, y, \sigma, \theta) & = \arg\max\limits_{tSD} tSD\,(x, y, \sigma, \theta) \\ \theta_o\,(x, y, \sigma, \theta) & = \arg\max\limits_{\theta} tSD\,(x, y, \sigma, \theta) \end{cases} \tag{4.14}$$

From Equation 4.14, 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 (Gonzalez et al., 2009), originally defined for a 2D binary X, as:

$$C\,(X) = \frac{4S\,(X)}{\pi L^2\,(X)} \tag{4.15}$$

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 grey level images, as presented in (Morard et al., 2013). 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

$$s_p = |I - C| \tag{4.16}$$

where I represents a 2D gray 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:

$$p_E(x,y) = \begin{cases} 0, \text{if } \lambda_1 \geq 0 \text{ or } tSD_o \geq 0.1 \text{ or } s_p \leq 10 \\ e^{-\frac{tSD_o^2}{3\sigma}} e^{-\frac{R^2}{2\beta^2}} \left(1 - e^{-\frac{S^2}{2c^2}}\right), \text{otherwise} \end{cases} \tag{4.17}$$

where $tSD_o \geq 0.1$ represents possible edge/boundary regions, $s_p \leq 10$ denotes low contrast or low Signal-to-noise ratio (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.

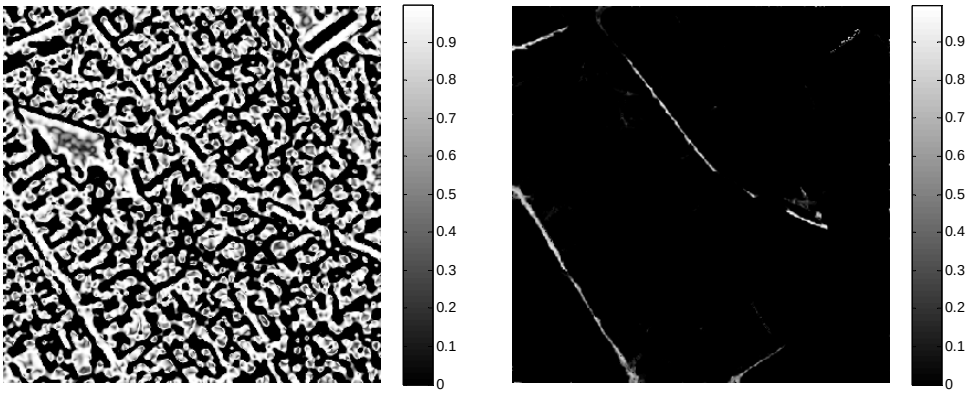
In practical applications, roads usually have different widths, and a multi-scale processing is required. Given a set of scales Ω_σ, linearness filter responses at multiple scales are combined using a max rule:

$$p_E(x,y) = \max\left\{ {p_E}^{\sigma_i}(x,y) \middle| \, \sigma_i \in \Omega_\sigma \right\} \tag{4.18}$$

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. Figure 4.3(a) and 4.3(b) compare 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 (Otsu, 1975) is applied to produce a binary mask.

To improve the extraction, in this chapter a pixel-wise ratio operator, called R_{vs} (Ok et al., 2013), 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:

$$R_{vs} = \frac{B_S - B_I}{B_S + B_I} \tag{4.19}$$



(a) (b)

FIGURE 4.3: (a) True color image of QuickBird image. The ground truth dataset, obtained by visual interpretation, is shown in yellow; (b) EM clustering result.

The advantage of this ratio is that it can simultaneously detect vegetation and shadow areas. To detect vegetation and shadow areas, Otsu's method (Otsu, 1975) 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.

4.2.3 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 by the different grey values in a segment due to image noise (e.g., branches and holes). Regarding these factors, traditional methods, such as the thinning algorithm (Gonzalez et al., 2009), 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 (Miao et al., 2013). 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

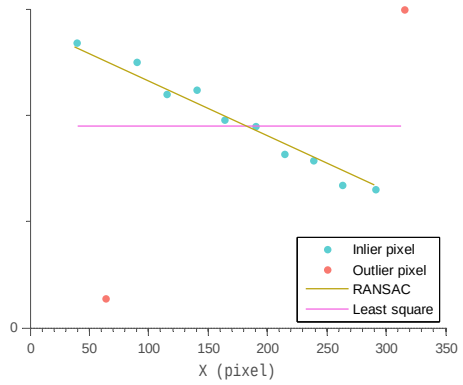
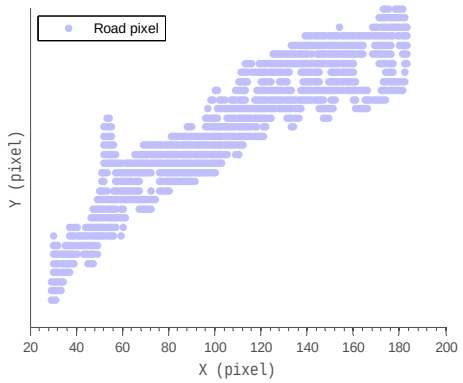
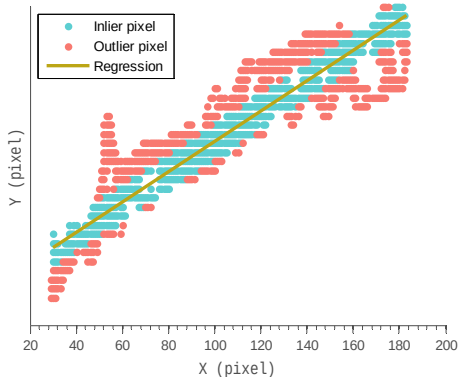


FIGURE 4.4: Difference between Least Square Regression and RANSAC.



(a)



(b)

FIGURE 4.5: An example of (a) road pixels identified in a candidate road segment; (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.

by noise and "false/outlier" data (i.e., pixels on edge and 'holes'). Accordingly, regression results may be strongly influenced by outliers (see Figure 4.4). A robust estimation is thus needed to obtain satisfactory centerlines.

To achieve stable and accurate results, this study relies on RANdom SAmple Consensus (RANSAC) (Fischler and Bolles, 1981) to extract centerlines from road segments, thanks to its efficiency and robustness to outliers. The RANSAC algorithm is a non-parametric 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 a) and b) are iteratively repeated and eventually the largest inlier set is selected, and the model re-estimated from it. Figure 4.5 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:

$$L_i = \begin{cases} \text{RANSAC}\left(A_i\right) & \text{, if Length}\left(\text{RANSAC}\left(A_i\right)\right) \geq T_L \\ \varnothing & \text{, otherwise} \end{cases} \tag{4.20}$$

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 4). The central idea of this

algorithm is to implement RANSAC in a recursive manner. For a complicated road segment (i.e., curved and with branches), Algorithm 4 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.

Algorithm 4 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 $\Omega^i_{outlier}$;
    $L\left\{i\right\} = l_i$; $\Omega_{inlier}\left\{i\right\} = \Omega^i_{inlier}$;
    $S \leftarrow LargestConnectedComponent\left(\Omega_S - \Omega_{inlier}\right)$;
    $\Omega_S \leftarrow \Omega\left(S\right)$;
    if $\frac{number(\Omega_{inlier})}{N} \geq 0.9$ **then**
        break;
    else
        continue.
    end if
end while

4.2.4 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.

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

$$R^I = \text{RANSAC}\left(\Omega_S \wedge \Omega_L\right) \tag{4.21}$$

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

$$R^{II} = \text{RANSAC}\left(\Omega_S \vee \Omega_L\right) \tag{4.22}$$

4.2.4.2 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:

$$R^{III} = P_4 \centerdot P_3 \centerdot P_2 \centerdot P_1 \left\{\text{RANSAC}\left(\Omega_S\right) \vee \text{RANSAC}\left(\Omega_L\right)\right\} \tag{4.23}$$

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 2D Euclidean space, two segments are said to be parallel if their extensions to infinitely long lines never cross each other, as illustrated in Figure 4.6(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 (L_i, L_j) are judged to be parallel if and only if they fall into the same category. In this study, K-Means algorithm (Hastie et al., 2009) is selected for the clustering task due to its simplicity and efficiency.

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 pre-defined threshold d_1, then the shorter segment is discarded. Hence, the spatial regularization rule P_1 for the parallel case is given by:

$$P_1\left(L_i, L_j\right) = \begin{cases} \arg\max_l \left(L_i, L_j\right) & \text{, if } D^1 \leq d_1 \\ \left(L_i, L_j\right) & \text{, otherwise} \end{cases} \tag{4.24}$$

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 Figure 4.6(b). In this case, if the smallest distance between their end points is smaller than pre-defined threshold d_2, then the two segements will be connected by extending the longer one to touch the other one.

$$P_2\left(L_i, L_j\right) = \begin{cases} LSF\left(L_i, L_j\right) & \text{, if } D^2 \leq d_2 \\ \left(L_i, L_j\right) & \text{, otherwise} \end{cases} \tag{4.25}$$

where $D^2 = \left\|M_i^2 M_j^1\right\|$, and LSF stands for least square fitting (LSF) (Hastie et al., 2009), 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 Figure 4.6(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 pre-defined threshold d_3, then these two lines will be connected. The spatial regularization strategy P_3 for this case is thus defined as:

$$P_3\left(L_i, L_j\right) = \begin{cases} \left(L_i, L_j, L_j^E\right) & \text{, if } D^3 \leq d_3 \\ \left(L_i, L_j\right) & \text{, otherwise} \end{cases} \tag{4.26}$$

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, let's 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,

$$a_1 x + b_1 y + c_1 = 0 \tag{4.27}$$

$$a_2 x + b_2 y + c_2 = 0 \tag{4.28}$$

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 pre-defined threshold d_4, the two lines will be extended and connected (see Figure 4.6(d)). The spatial regularization strategy for the intersection case can be expressed as:

$$P_4\left(L_i, L_j\right) = \begin{cases} \left(L_i, L_j, L_i^E, L_j^E\right) & \text{, if } D^4 \leq d_4 \\ \left(L_i, L_j\right) & \text{, otherwise} \end{cases} \tag{4.29}$$

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.

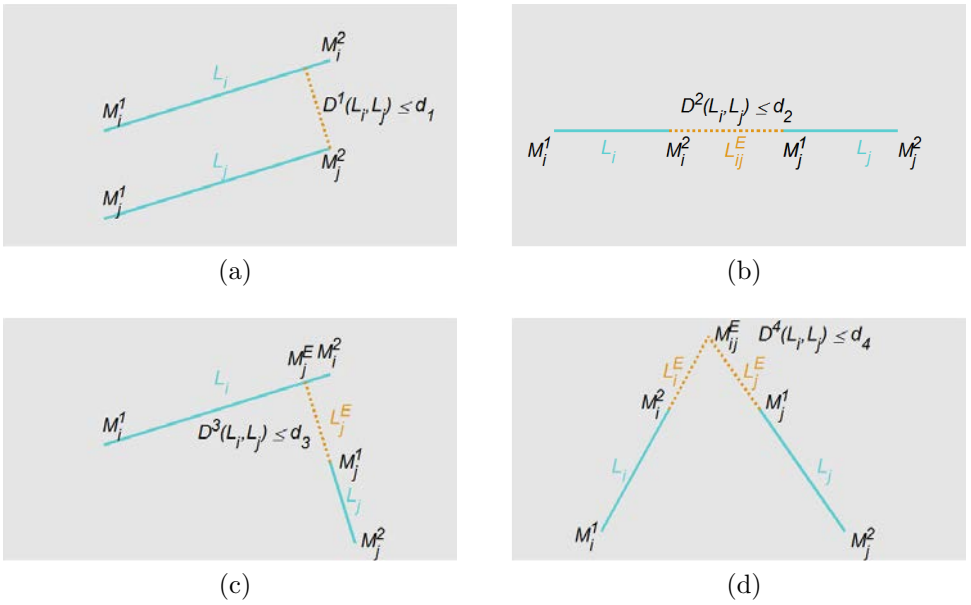


FIGURE 4.6: Four spatial regularization strategies considered in rule set III. (a) Parallel segments. (b) Collinear segments. (c) Extension of one line segment intersecting the extension of another. (d) Segments that intersect each other.

4.2.5 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 been formally proven in the No Free Lunch theorem (Wolpert and Macready, 1997) that it is impossible to tune a parametric algorithm such that it has optimal settings for all possible natural image scenes. Accordingly, in this chapter 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) $d_4 = \lfloor 2.5 * R_w$, where $\lfloor \bullet \rfloor$ rounds a number to the nearest integer less than or equal to it.

4.3 Experimental results

In this section, the proposed method is validated by using two multi-spectral satellite images. For quantitative evaluation, three measures are computed: 1) Completeness, 2) Correctness, and 3) Quality (Wiedemann et al., 1998). 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 4.1.

TABLE 4.1: Parameters used for each test site

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

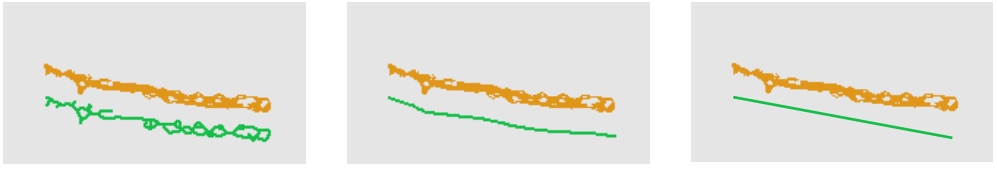
4.3.1 Comparison of different centerline extraction algorithms

The advantage of the proposed approach for road centerline extraction from the classified image has been checked first. To this end, the proposed method is compared with two methods existing in technical literature, that are 1) the thinning algorithm (Press, 2007), and 2) the Multivariate Adaptive Regression Splines (MARS) (Miao et al., 2013). The results of the comparison are reported in Figure 4.7.

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 Figure 4.7. 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 Figure 4.7(a)). Although the presence of spurious elements does not affect the interpretation of road centerline, it reduces its smoothness and causes inacuracies in road network patterns. By contrast, both MARS and RANSAC can provide results more accurate than the thinning algorithm, as illustrated in Figure 4.7(b) and 4.7(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.



(a) (b) (c)

FIGURE 4.7: Centerline results by (a) the thinning algorithm (Press, 2007), (b) MARS (Miao et al., 2013), 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.

4.3.2 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. Figure 4.8 presents the results of all of the information fusion rules. The comparison among results from different information fusion rules is reported in Table 4.2. Figure 4.8 shows that rule set I provides much cleaner extractions

(a) (b) (c)

FIGURE 4.8: (a), (b), and (c) show the information fusion results of rule set I, rule set II, and rule set III, respectively.

TABLE 4.2: Quantitative evaluation results of different information fusion rules

Fusion method	Completeness	Correctness	Quality
rule set I	0.6	1	0.6
rule set II	0.84	0.8	0.69
rule set III	0.82	0.92	0.77

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 4.2. The rationale for this results is that the logical *AND* operation aims at retaining as much correct information as possible. From Table 4.2, 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 Figure 4.8(a) and Figure 4.8(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 Figure 4.8(b) and Figure 4.8(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 linearness filter. The results in Table 4.2

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.

4.3.3 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 pixels by 1024 pixels and Figure 4.9(a) shows the study area. The test image has four multispectral bands with the spatial resolution of 2m per pixel and the proposed method is validated on the original multispectral bands with pan-sharpening. Figs. 4.9(b)-4.9(c) present the results of different road extraction strategies. From Figure 4.9, 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, Figure 4.9 also shows that the linearness filter cannot extract road pixels in some areas (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 overal performance improved (see Figure 4.9(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 post-processing road centerline segments by means of semi-automated methods.

Table 4.3 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 4.3 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 trade-off between completeness and correctness.

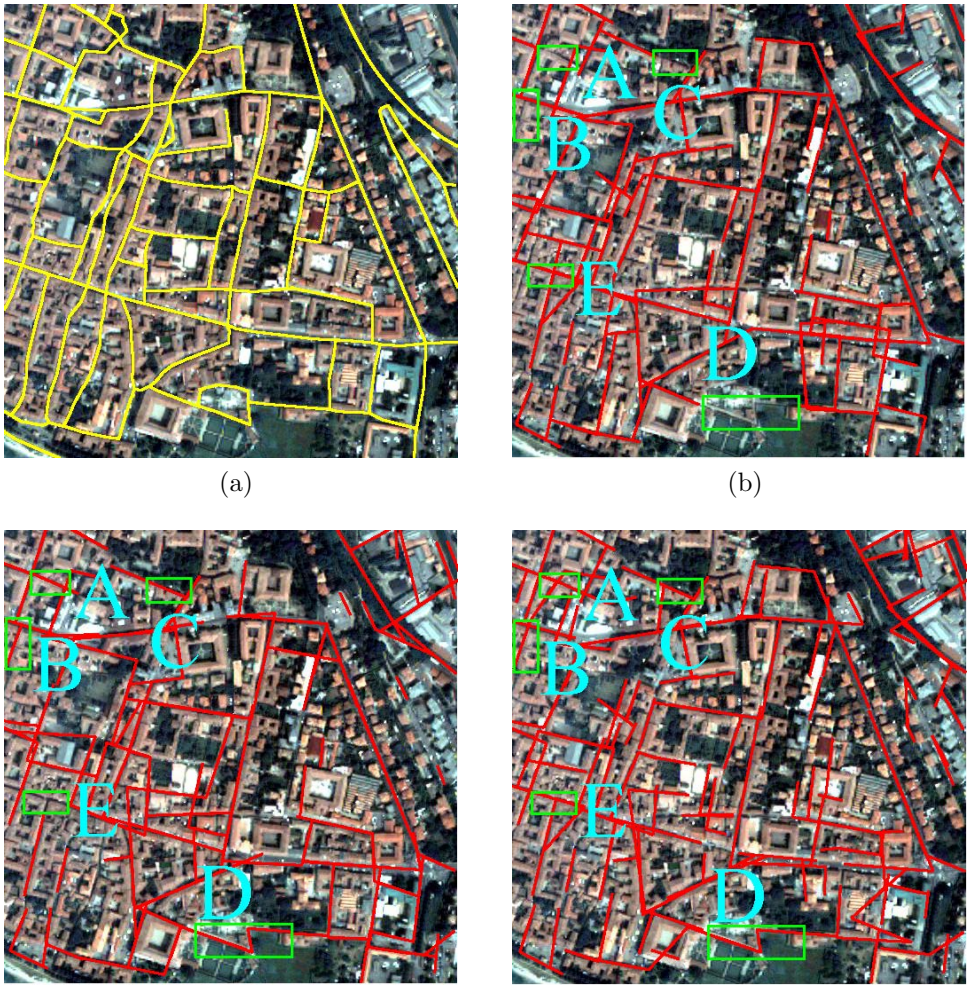
TABLE 4.3: Quantitative evaluation results of different road extraction strategies on Pavia test area.

Method	Completeness	Correctness	Quality
EM	0.5	0.66	0.4
The linearness filter	0.54	0.65	0.42
Fusion	0.69	0.63	0.49

4.3.4 Experiment 2

The second experiment was performed on a dataset acquired by QuickBird on Xuzhou City, People's Republic of China. A subset (1024 pixels by 1024 pixels) of the whole pan-sharpened image was selected as the experimental area, and presented in Figure 4.10(a). The road centerline extraction results using EM segment only, the linearness filter only, and the proposed method are shown in Figs. 4.10(b)-4.10(d), respectively. In Figure 4.10, the ground truth dataset is shown in yellow; while extraction results in red.

Table 4.4 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



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

FIGURE 4.9: 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.

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

(a) (b)

(c) (d)

FIGURE 4.10: 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.

improving road extraction accuracy.

TABLE 4.4: Quantitative evaluation results of different road extraction strategies on Xuzhou test area.

Method	Completeness	Correctness	Quality
EM	0.69	0.64	0.5
The linearness filter	0.66	0.52	0.41
Fusion	0.84	0.61	0.54

4.3.5 Comparison with state-of-the-art methods

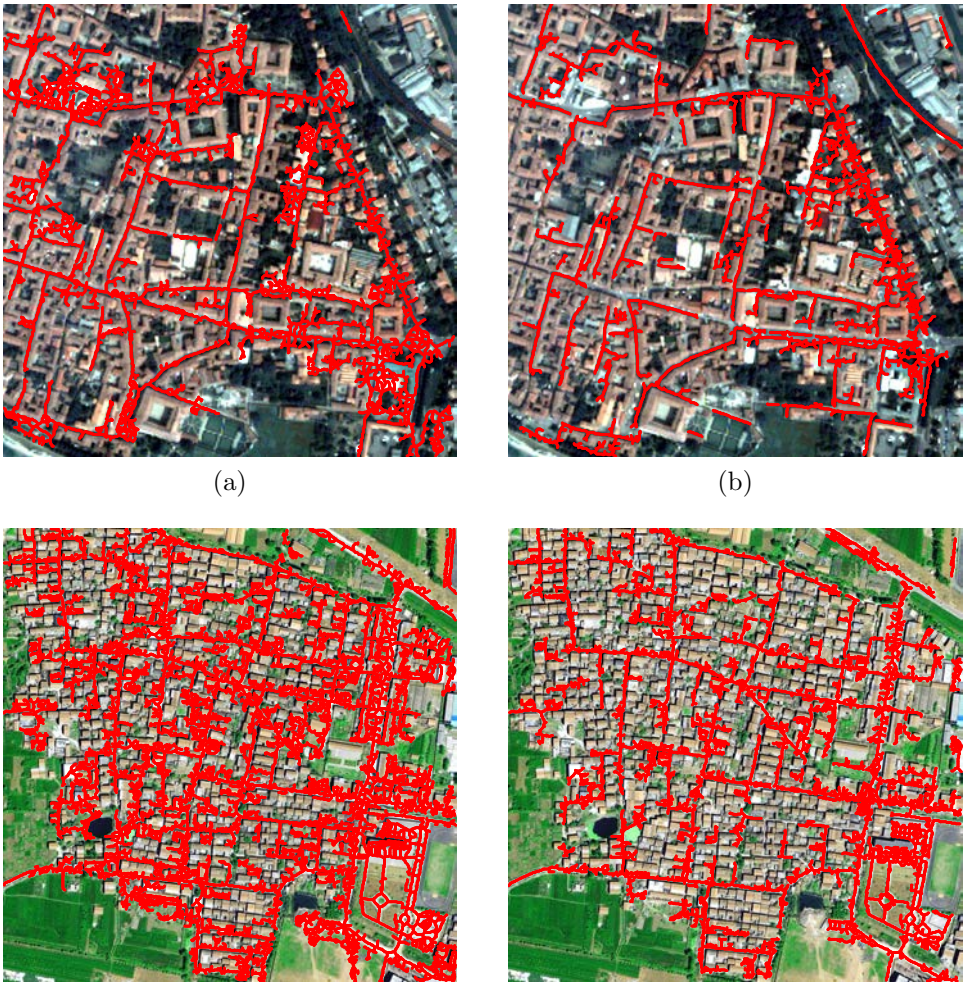
To assess the results of the presented method, in this area, a comparison was carried out with the two state-of-the-art methods: 1) Song's method (Song and Civco, 2004), and 2) Huang's method (Huang and Zhang, 2009). Figure 4.11 gives the comparison results of these two road extraction methods. As it can be seen, road features are extracted well by the proposed method. This performance is also confirmed quantitatively by road extraction accuracy in terms of three measures (i.e. Completeness, Correctness, and Quality). Table 4.5 lists the quantitative evaluation results for the three methods. As shown in Table 4.5, 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 post-classify 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.

4.4 Summary

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

TABLE 4.5: Comparison of different road extraction methods.

Study area	Method	Completeness	Correctness	Quality
Pavia	Song (Song and Civco, 2004)	0.72	0.45	0.38
	Huang (Huang and Zhang, 2009)	0.57	0.5	0.37
	Proposed	0.69	0.63	0.49
Xuzhou	Song (Song and Civco, 2004)	0.97	0.32	0.32
	Huang (Huang and Zhang, 2009)	0.94	0.48	0.46
	Proposed	0.84	0.61	0.54



(a) (b)

(c) (d)

FIGURE 4.11: Comparison results of different road extraction methods on two test areas: the first column show the results of Song's method (Song and Civco, 2004), while the second column the results of Huang's method (Huang and Zhang, 2009) on Pavia and Xuzhou, respectively.

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. In this chapter the proposed approach exploits information fusion rule sets, and is suited to Very High Resolution (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.

4.4.1 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 Frangi's filter (Frangi et al., 1998), 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 non-road 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 chapter 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 Signal-to-noise ratio (SNR) regions and at boundaries between land cover classes. We believe this approach provides far more cleaner result than the one proposed in (Frangi et al., 1998).

4.4.2 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 (Huang and Zhang, 2009). 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.

4.4.3 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.

Chapter 5

Accurate Road Centerline Extraction

5.1 Introduction

Many cutting-edge road extraction methods consist of two main steps: extracting road segments by applying a set of rules, followed by centerline delineation from road segments extracted in the first step. The commonly used method is the thinning algorithm (Gonzalez et al., 2009). Despite its computational efficiency, the result of the thinning algorithm always produces un-desired objects (i.e. spurs) that reduce the smoothness and accuracy of the centerline. To tackle this limitation to some extent, a self-organized clustering method (Doucette et al., 2001) was designed. This method firstly extracted initial points using K-medians clustering, followed by the linking algorithm to create the central line. Radon and Hough transform (Poullis and You, 2010, Zhang and Couloigner, 2007) were also introduced to extract the centerline from the classified image. Strong peak locations in the transform matrix are taken to be straight line pixels in the original image. Radon and Hough transform are suitable for straight line extraction and generally fail to process curvilinear cases. Recently, in the area of computer vision, a so-called subspace

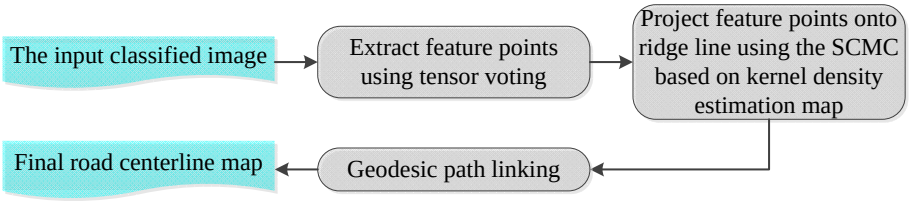


FIGURE 5.1: Flowchart of the F-SCMS method.

constrained mean shift (SCMS) method (Ozertem and Erdogmus, 2007, 2009, 2011) was proposed to extract accurate centerline from point clouds. First, the probability that a pixel is located on the centerline is computed using the kernel density estimation (KDE). Then, all discrete points are iteratively projected onto the centerline based on mean shift. Despite its good performance, when the number of points is large, SCMS becomes very slow, which limits its real applications. To improve the computational efficiency of SCMS, this chapter firstly presents a feature point based SCMS (F-SCMS). It should be noted that F-SCMS relies on the spatial connection condition. To reduce this spatial constraint, this chapter subsequently designs a Gaussian Mixture Model based SCMS, named GMM-SCMS.

5.2 Feature point based road centerline extraction

This section devises a feature-point-based approach to extract accurate centerlines from classified images. Figure 5.1 summarizes the proposed method. The proposed method consists of the three following steps.

1. Feature points (i.e., junction and end points) are first extracted from the classified image.

2. The probability of each pixel being located on the road centerline is computed using the kernel density estimation (KDE) method. Subsequently, the feature points are projected onto ridge lines using the subspace constrained mean shift (SCMS) method.

3. The projected feature points are linked by the geodesic method to create the central line to formulate the road network.

The details of each step are described in the following section.

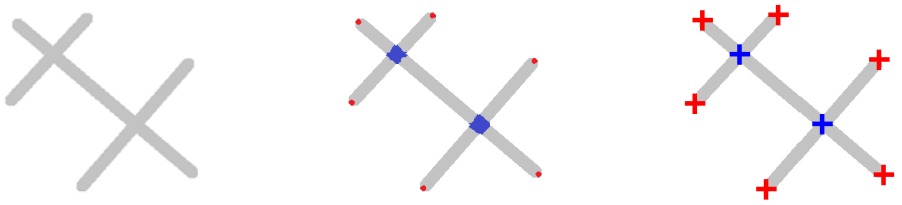
5.2.1 Extraction of feature points

From Chapter 2, we can know that, after tensor voting, vote analysis is performed to extract two types of feature points: 1) junction points and 2) end points. Feature points are classified based on the following two rules.

1. Points that satisfy $\lambda_1 \approx \lambda_2 > 0$ are classified as region or junction points.
2. The local maxima of polarity are detected as endpoints. The local maxima are the points within a given neighborhood at which the pixel takes the largest value. The image dilation (Gonzalez et al., 2009), one of mathematical morphology operations, can be applied to detect local maxima.

An example of feature point extraction using tensor voting is presented in Figure 5.2. As can be seen from Figure 5.2(b), feature points extracted by tensor voting are areas, not single points. In this study, centroids of these areas are taken as feature points, resulting in Figure 5.2(b).

After the extraction of feature points, the corresponding connectivity matrix is constructed using Algorithm 5. The connectivity matrix of each connected component (CP) can be defined as follows: a feature point is connected to any feature points except itself. In this means, if some junctions are missed, they can still be recovered when the algorithm tries to link endpoints. Another advantage of introducing this connectivity matrix is that, for each end point, there is not necessary to determine which junction point is connected to, thereby resulting in an efficient way of constructing such connectivity matrix. Meanwhile, Figure 5.3 shows an example of generating the connectivity matrix



(a) (b) (c)

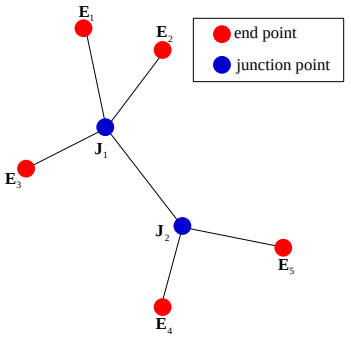
FIGURE 5.2: (a) Input image. (b) Feature points extraction results using tensor voting, where junction points are shown in blue and end points in red. (c) Centroids of the junction points are shown in blue crosses and end points in red crosses.

from foregoing extracted feature points. The feature points detected by the tensor voting are shown in Figure 5.3(a), and the corresponding connectivity matrix is shown in Figure 5.3(b). As can be seen, the connectivity matrix is symmetric, and hence only the lower triangular elements need to be processed by the following procedures.

Algorithm 5 The connectivity matrix construction algorithm

Input
The classified image.
Output
The connectivity matrix.

Perform connected component analysis (CCA) (Gonzalez et al., 2009) on the classified image.
Select one connected component $comp_i$ and then perform tensor voting to extract feature points. Suppose the component contains M_i feature points (i.e., junction points and end points).
for do $m_1 = 1, \ldots, M_i$
    for do $m_2 = 1, \ldots, M_i$ and $m_2 \neq m_1$
        $Conn\,(m_1 m_2) = 1$
    end for
end for
Repeat the above steps until all of the components are analysed.



(a)

	J1	J2	E1	E2	E3	E4	E5
J1	0	1	1	1	1	1	1
J2	1	0	1	1	1	1	1
E1	1	1	0	1	1	1	1
E2	1	1	1	0	1	1	1
E3	1	1	1	1	0	1	1
E4	1	1	1	1	1	0	1
E5	1	1	1	1	1	1	0

(b)

FIGURE 5.3: Example of constructing the connectivity matrix. (a) Feature points extracted by tensor voting. (b) The corresponding connectivity matrix.

5.2.2 Adjustment of initial road centerline using SCMS

The major axis produced by GMM in Section 5.2.1 can be regarded as the initial road centerline that is commonly coarse. For the purpose of accurate road centerline result, it is promising to obtain accurate road centerline using the adjustment techniques. To this end, subspace constrained mean shift (SCMS) is selected to align the major axis. However, SCMS is iteratively implemented on the whole classified road pixels, which leads to high computational load. On the other hand, the major axis can also be taken as the simplification case of the classified road map that only account for small size pixels. In other words, GMM finds the representative points of the classified road pixels to be used in principal curve projection; and therefore, GMM followed by SCMS, termed as GMM-SCMS, provides a new means to lower the computational load. The SCMS method is briefly introduce as follows.

Considering a d-variate random sample $X_1, X_2, \cdots, X_n$ drawn from a population with its density function f, the kernel density estimator (KDE) is defined as

$$\hat{f}\left(x; \sum\right) = n^{-1} \sum_{i=1}^{n} K_{\sum}\left(x - X_i\right) \tag{5.1}$$

where $x = (x_1, x_2, \cdots, x_d)^{\mathrm{T}}$ and $X_i = (X_{i1}, X_{i2}, \cdots, X_{id})^{\mathrm{T}}$, $i = 1, 2, \cdots, n$. In addition, the symmetric positive definite $d \times d$ matrix $\sum$ is the bandwidth

matrix, which critically determines the performance of the KDE function $\hat{f}$. K is a d-variate kernel function defined as

$$K_{\sum}(x) = \left|\sum\right|^{-1/2} K\left(\sum^{-1/2} x\right) \tag{5.2}$$

which satisfies $\int K(x)\,dx = 1$. Here, $|\cdot|$ indicates the determinant operation, and the common choice of Gaussian KDE functions is adopted, $K(x) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}x^T x\right)$.

By generalizing the Scott's rule of thumb, bandwidth matrix $\sum^{-1} = n^{\frac{2}{d+4}} \hat{\sum}^{-1}$ is chosen here (see Ahamada and Flachaire (2010) for detail). Consequently, the kernel density estimator is constructed above as

$$p\left(x; \sum\right) = \frac{1}{n(2\pi)^{d/2} |\sum|^{1/2}} \sum_{i=1}^{n} \exp\left(-\tfrac{1}{2}(x - X_i)^{\mathrm{T}} \sum^{-1} (x - X_i)\right) \tag{5.3}$$

In consideration of the numerical stability when solving the principle curve, the logarithm of the density function $p(x; \sum)$ is involved,

$$f(p(x)) = \log(p(x)) \tag{5.4}$$

The corresponding gradient and Hessian matrices for KDE are

$$\begin{aligned} g_f(x) &= f'(p(x)) \nabla p(x)^{\mathrm{T}} \\ &= p(x)^{-1} g(x) \end{aligned} \tag{5.5}$$

$$\begin{aligned} H_f(x) &= f''(p(x)) \nabla p(x)^{\mathrm{T}} \nabla p(x) + f'(p(x)) \nabla^2 p(x) \\ &= f''(p(x)) g(x) g(x)^{\mathrm{T}} + f'(p(x)) H(x) \\ &= -p(x)^{-2} g(x) g(x)^{\mathrm{T}} + p(x)^{-1} H(x) \end{aligned} \tag{5.6}$$

Thus, we get the fixed-point update rule for adjusting the coarse initial point,

$$m_{\sum}(x) = \frac{\sum_{i=1}^{n} \exp\left(-\frac{1}{2}(x - X_i)^{\mathrm{T}} \sum^{-1} (x - X_i)\right) X_i}{\sum_{i=1}^{n} \exp\left(-\frac{1}{2}(x - X_i)^{\mathrm{T}} \sum^{-1} (x - X_i)\right)} - x \tag{5.7}$$

On account of the similarity of the above iteration formula to the mean-shift algorithm and thus it has the name of subspace constrained mean shift (SCMS) method, which is presented in Algorithm 6. An example of adjusting coarse points to produce accurate centerline utilizing SCMS is illustrated in Figure 5.4.

Algorithm 6 Subspace Constrained Mean Shift

Input
The coarse road centerline derived from the major axis $X_1, X_2, \cdots, X_n$.
Output
The adjusted discrete points $X_1^{new}, X_2^{new}, \cdots, X_n^{new}$.

Provide the band width matrix $\sum^{-1} = n^{\frac{2}{d+4}} \hat{\sum}^{-1}$ (i.e., the kernel covariance).
for do $i = 1, \ldots, K$
    $P = I - qq^T$ and $s = Pm_{\sum}(x)$
    while $s < e_{thre}$ **do**
        $X_i^{t+1} = X_i^t + s$
    end while
    $X_i^{new} = X_i^{t+1} + s$
end for

5.2.3 The geodesic method

After the projection of feature points onto principal curves in Step B, the geodesic method (Peyré et al., 2010) is subsequently used to create the central line by linking the feature points. Consider a smooth curve $r : [0, 1] \to \Omega$ on an image, where Ω is the image domain that is defined as $\Omega = [0, 1]^2$. The smooth curve γ is generally constrained as

$$\{\gamma : [0, 1] \to \Omega \backslash \gamma(0) = x_s and \gamma(1) = x_e\} \tag{5.8}$$

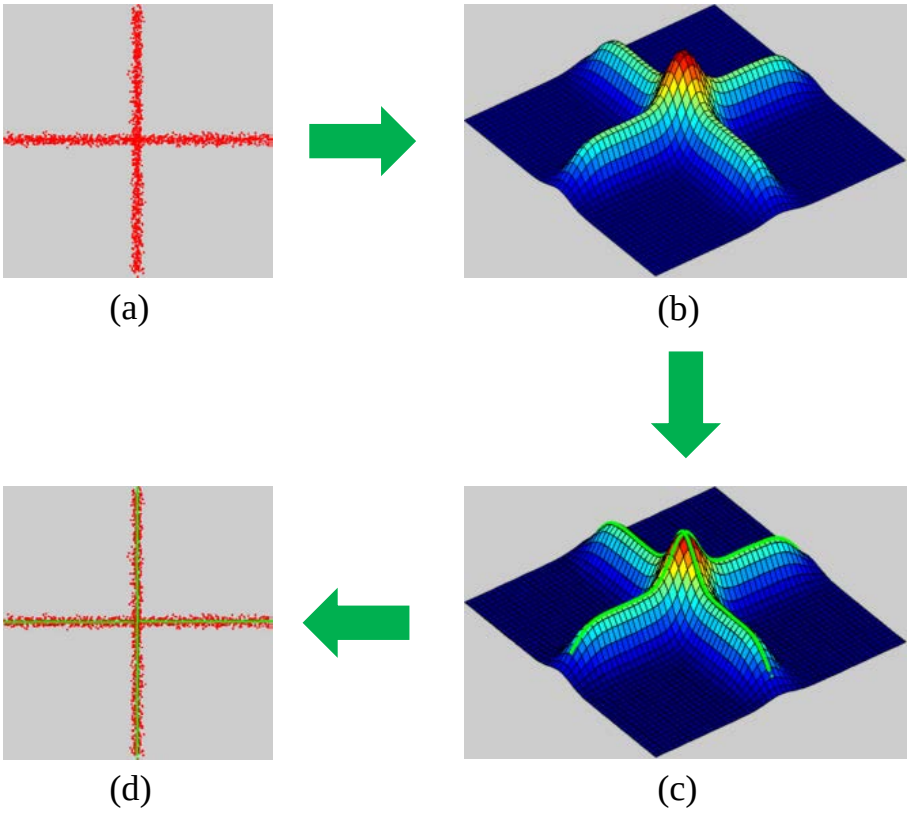


FIGURE 5.4: An example of a principal curve. (a) Noisy input points. (b) The estimated probability density map. (c) The principal curve over the probability density map where the principal curve is shown in green. (d) The principal curve projected back onto the plane of original points.

where x_s and x_e respectively denote the starting point and ending point.

Let $L(\gamma)$ denote the weighted length of the smooth curve γ. The weighted length $L(\gamma)$ can be computed as follows

$$L(\gamma) = \int_0^1 W(\gamma(t))||r'(t)||dt \tag{5.9}$$

where $\gamma'(t) \in \mathbb{R}^2$ is the derivative of γ, and t denotes the parameter of the curve.

Let $f : \Omega \to \mathbb{R}$ denote the probability density estimation map that is generated by KDE. For the purpose of road network formulating, a road can be approximately defined as a smooth curve that has a constant kernel density value $c \in \mathbb{R}$. Based on this definition of a road model, a weight function $W(x)$

can be defined as

$$W(x) = |f(x) - c| + \varepsilon \tag{5.10}$$

where ε (i.e., $\varepsilon = 0.01$) is a small value that prevents $W(x)$ from vanishing. In this study, the constant value c is fixed to $c = f(x_e)$. As can be seen from Equation 5.10, this is the area where the principal curve (i.e., the road centerline) is passing through the area that has low weights.

The minimal path that links two feature points is a global minimizer of the length that should satisfy the following equation

$$\gamma^* = \underset{\gamma \in (x_s, x_e)}{\arg\min}\, L(\gamma) \tag{5.11}$$

The steps of the feature points linking procedure based on the geodesic method are presented in Algorithm 7.

Algorithm 7 Feature points linking using the geodesic method

Input
The classified image.
Output
The road centerline map.

Compute the kernel density estimation map of the classified image.
Extract road feature points using tensor voting.
Construct the weight matrix using Equation 6.1.
Construct the geodesic distance map and find the geodesic path.
Link feature points using the geodesic path.

5.3 Centerline extraction without spatial connection constraints

The objective of this study is to devise a computationally efficient approach (i.e. GMM-SCMS) with high general ability to extract accurate centerlines from classified images. In contrast to F-SCMS, GMM-SCMS has no spatial connection constraint. Figure 5.5 summarizes the proposed method. The

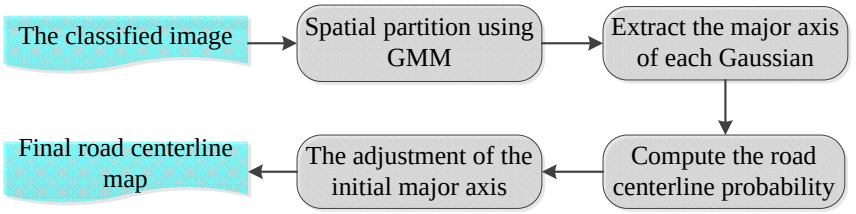


FIGURE 5.5: Flowchart of the GMM-SCMS method.

proposed method consists of the three following steps.

1. The classified image is partitioned using the Gaussian Mixture Model (GMM) method.

2. The major axis is subsequently extracted from the ellipsoid of each Gaussian.

3. The major axis is adjusted based on SCMS to produce accurate road centerline.

In the following sub-sections, details of each step are described.

5.3.1 Image partition using GMM

Road pixels in the classified map can be represented as 2D discrete joint random variables. We can partition the classified image into multiple segments by taking road pixels as observations. To do so, the Gaussian Mixture Model (GMM) is used to implement image clustering. The reason of selecting GMM is that the Gaussian models used by the expectation-maximization (EM) algorithm are flexible and the EM result is able to accommodate clusters of variable size. This is because that EM benefits from the Gaussian distribution present in the data set. Therefore, GMM is able to capture the complex spatial structure. In other words, GMM is able to simplify the classified road map as well as retain its spatial topology. The concepts of GMM are briefly introduced as follows.

GMM can be written as a linear superposition of K Gaussians that leads to a probability density in the form

$$s_{K,V,\Sigma,w}(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_k \Phi_{V_k,\Sigma_k}(\boldsymbol{x}) \tag{5.12}$$

where $\{\pi_k\}_{k=1}^{K}$ are mixture weights subject to: $0 \leq \pi_k \leq 1$, $\sum_{k=1}^{K} \pi_k = 1$, Φ is a multivariate Gaussian probability distribution of mean ν_k and covariance matrix $\sum_k$.

$$\Phi_{\nu,\sum}(x) = \frac{1}{2\pi\sqrt{|\sum|}} e^{\frac{-1}{2}(\mathbf{x}-\nu)^T \sum^{-1}(\mathbf{x}-\nu)} \tag{5.13}$$

The expected value of the indicator variables under the posterior distribution is then given by

$$\gamma(z_{ij}) = \frac{\pi_j \Phi_{\nu_j,\sum_j}(\mathbf{x}_i)}{\sum_{k=1}^{K} \pi_k \Phi_{\nu_k,\sum_k}(\mathbf{x}_i)} \tag{5.14}$$

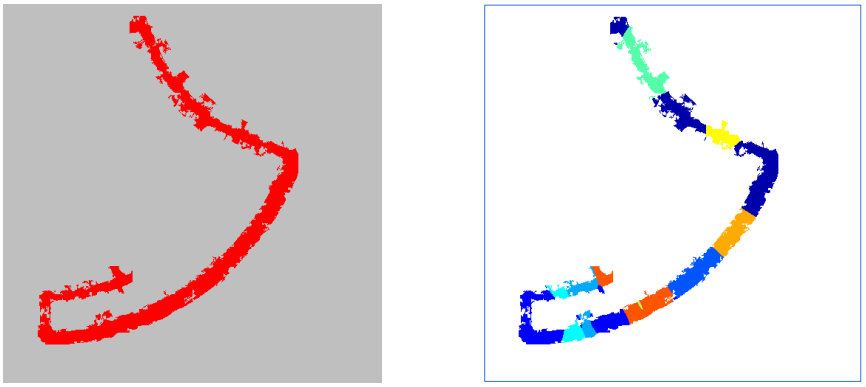
The log-likelihood function is defined as:

$$\ln p(x, z\,|\Phi) = \sum_{i=1}^{N}\sum_{k=1}^{K} \gamma(z_{ij}) \ln\left\{\pi_k \Phi_{\nu_k,\sum_k}(\mathbf{x}_i)\right\} \tag{5.15}$$

The parameters of each component are updated using the Expectation-Maximization (EM) algorithm as follows:

$$\begin{aligned} \pi_k &= \frac{1}{N}\sum_{i=1}^{N} \gamma(z_{ij}) \\ \mu_j &= \frac{1}{\sum_{i=1}^{N} \gamma(z_{ij})} \sum_{i=1}^{N} \gamma(z_{ij}) \mathbf{x}_i \\ \sum\nolimits_j &= \frac{1}{\sum_{i=1}^{N} \gamma(z_{ij})} \sum_{i=1}^{N} \gamma(z_{ij}) (\mathbf{x}_i - \mu_j)(\mathbf{x}_i - \mu_j)^T \end{aligned} \tag{5.16}$$

The EM step is iteratively repeated until convergence. It is worthy of pointing out that initial estimates for the parameters (Φ) are necessary for the first EM iteration. To this end, the k-Means algorithm is selected to initialize such parameters involves running. The spatial cluster number for k-Means is



(a) (b)

FIGURE 5.6: (a) The classified road map. (b) The spatial partition result using GMM.

estimated by

$$K = \left\lceil \frac{A}{R_W \times M_L} \right\rceil \tag{5.17}$$

where $\lceil \bullet \rceil$ means to round the number to the nearest integer greater than or equal to this number, A is the area of the classified image, R_W is the average road width of the classified image, and M_L is the minimum area of a spatial cluster. Details of the image partition using GMM are illustrated in Algorithm 8. Figure 5.6 gives a conceptual example. Figure 5.6(a) is the classified image and its corresponding spatial partition result using GMM is shown in Figure 5.6(b).

Algorithm 8 Image Partition Using GMM

Input
The classified image.
Output
The mean μ and covariance matrix $\sum$.

Compute the spatial cluster number K using Equation 5.17.
Initial the parameters of EM using k-Means.
Run the EM algorithm to segment the classified image until convergence.
for do $i = 1, \ldots, K$
    compute the mean μ_i and covariance matrix $\sum_i$ of ith Gaussian.
end for

5.3.2 Extraction of initial road centerlines

After the spatial partition using GMM in Section 2.1, the mean μ and covariance matrix $\sum$ for each mixture are produced. The ellipsoids of each Gaussian can be defined using μ and $\sum$. In this study, the major axis of each ellipsoid is taken as the approximate initial centerline. Σ is a square matrix that can be factorized as

$$\Sigma = U \Lambda U^{\mathrm{T}} \tag{5.18}$$

where U is the eigenvector matrix, and Λ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. The inverse matrix of $\sum$ is given by

$$\sum^{-1} = U^{-\mathrm{T}} \Lambda^{-1} U^{-1} = U \Lambda^{-1} U = \sum_{i=1}^{p} \frac{1}{\lambda_i} \vec{u}_i \vec{u}_i^{\mathrm{T}} \tag{5.19}$$

The inverse matrix $\sum^{-1}$ can be rewritten as

$$\begin{aligned} (\vec{x} - \vec{u})^{\mathrm{T}} \sum^{-1} (\vec{x} - \vec{u}) &= (\vec{x} - \vec{u})^{\mathrm{T}} \left(\sum_{i=1}^{p} \frac{1}{\lambda_i} \vec{u}_i \vec{u}_i^{\mathrm{T}} \right) (\vec{x} - \vec{u}) \\ &= \sum_{i=1}^{p} \frac{1}{\lambda_i} (\vec{x} - \vec{u})^{\mathrm{T}} \vec{u}_i \vec{u}_i^{\mathrm{T}} (\vec{x} - \vec{u}) \\ &= \sum_{i=1}^{p} \frac{y_i^2}{\lambda_i} \end{aligned} \tag{5.20}$$

where $y_i \triangleq \vec{u}_i^{\mathrm{T}} (\vec{x} - \vec{u})$. The y variables define a new coordinate system that is shifted (by $\vec{u}$) and rotated (by U) with respect to the original coordinates: $\vec{y} = U (\vec{x} - \vec{u})$. Recall that the equation for an ellipse in 2D is

$$\frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} = 1 \tag{5.21}$$

Equation 5.21 indicates that the contours of equal probability density of a Gaussian lie along ellipse. Figure 5.7 demonstrates an example of the ellipsoid generated by the mean and covariance matrix. For each Gaussian mixture, the major axis of the ellipse is estimated. After that, all major axis will compose

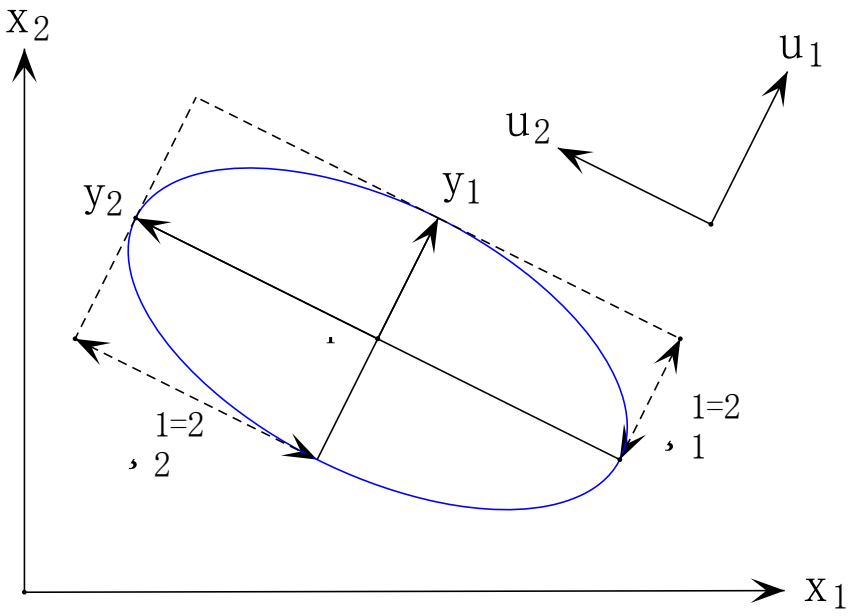
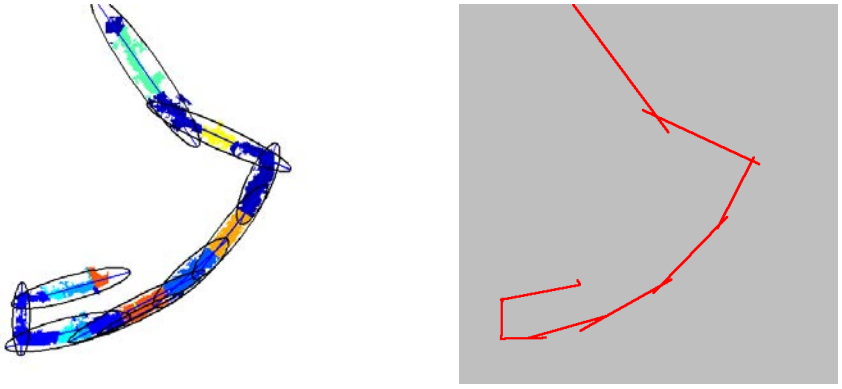


FIGURE 5.7: An example of the ellipsoid derived by the mean and covariance matrix.



(a) (b)

FIGURE 5.8: (a) Ellipse of each Gaussian. (b) The major axis which is shown in red.

of a coarse road centerline network that has a similar spatial topology with the true one, as shown in Figure 5.8.

5.3.3 Adjustment of initial road centerline using SCMS

The major axis produced by GMM in Section 5.3.2 can be regarded as the initial road centerline that is commonly coarse. For the purpose of accurate road centerline result, it is promising to obtain accurate road centerline using the

adjustment techniques. To this end, subspace constrained mean shift (SCMS) is selected to align the major axis. However, SCMS is iteratively implemented on the whole classified road pixels, which leads to high computational load. On the other hand, the major axis can also be taken as the simplification case of the classified road map that only account for small size pixels. In other words, GMM finds the representative points of the classified road pixels to be used in principal curve projection; and therefore, GMM followed by SCMS, termed as GMM-SCMS, provides a new means to lower the computational load. The SCMS method is briefly introduce as follows.

On account of the similarity of the above iteration formula to the mean-shift algorithm and thus it has the name of subspace constrained mean shift (SCMS) method, which is presented in Algorithm 6. An example of adjusting coarse points to produce accurate centerline utilizing SCMS is illustrated in Figure 5.4.

5.4 Experiment and results

In this section, several experiments that test the proposed method are described. The proposed method is also compared with other methods in the literature to show the advantages and disadvantages of the proposed method. In this study, MATLAB®[1] was used as the coding environment on a PC that has an Intel Core2Quad processor with 2.83-GHz clock speed.

5.4.1 Tests of parameters

To test the minimum area M_L influence on the road centerline extraction accuracy, we adjusted its value automatically from 5 to 30 with an increment of 5 for each step. The test results are presented in Figure 5.9. The performance

[1]MATLAB (matrix laboratory), developed by MathWorks, is a multi-paradigm numerical computing environment and fourth-generation programming language.

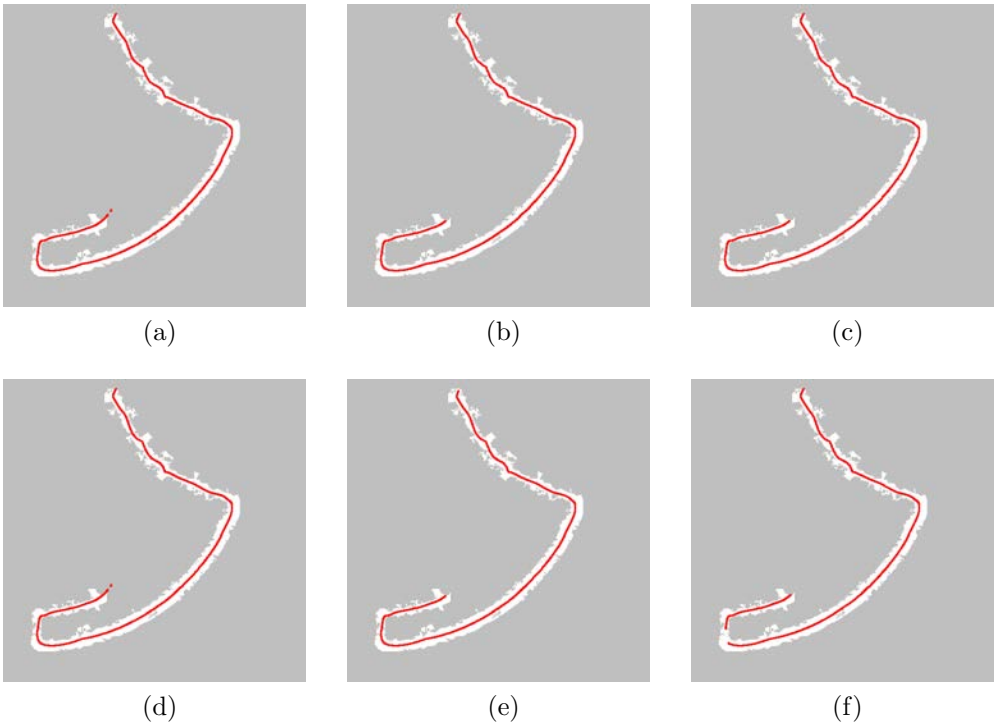
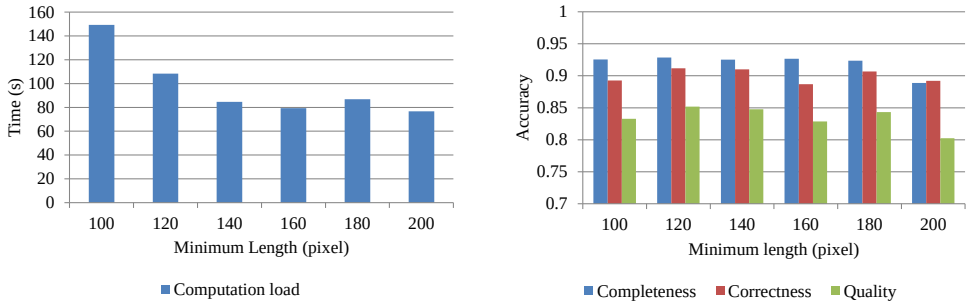


FIGURE 5.9: Road centerline extraction results with different the minimum area values. (a) 100. (b) 120. (c) 140. (d) 160. (e) 180. (f) 200. The road centerline extraction results are shown in red.

of the M_L influence is quantitatively evaluated in terms of the computational time and accuracy. Here, three accuracy measures, 1) Completeness, 2) Correctness, and 3) Quality (Wiedemann et al., 1998), were used. In this study, the ground truth dataset is produced by hand drawing method and the buffer width is fixed to 8 pixels.

Figure 5.10 reports quantitative evaluation results with different minimum area values. Figure 5.10(a) indicates that the computational efficiency steadily improved with the increase of the minimum area value. This is because that large M_L will result in initial road centreline with small number points, which in turn reduce the computational load of the adjustment using SCMS. Figure 5.10(b) illustrates the influence of M_L on the accuracy. It can be seen that, when M_L is increased from 100 to 180, there is a slight change of the accuracy, which indicates that M_L has a weak effect on the accuracy in this range of values. However, the accuracy is quick drop when M_L is exceeding 180. The reason for this phenomenon is that the large M_L cannot produce the initial



(a) (b)

FIGURE 5.10: The evaluation results of the minimum area influence. (a) Computational load. (b) Accuracy.

road centerline with the correct spatial topology as the classified image, as shown in Figure 5.9(f). To achieve the best balance between the computational efficiency and the accuracy, this study fixed M_L to 180 throughout the tests.

5.4.2 Experiments on simulated images

The proposed method is tested on three simulated images, including Circle image, Spiral image, and Snow image, as shown in the first column of Figure 5.11. The intermediate results using GMM are shown in the second column of Figure 5.11 while the corresponding precise centerline results using SCMS in the third column of Figure 5.11.

The propose F-SCMS and GMM-SCMS are compared with the cutting-edge method (i.e., SCMS). The comparison results are illustrated by Figure 5.12. From Figure 5.12, it can be seen that F-SCMS well works in Spiral and Snow cases, but fails to process Circle case. By contrast, both SCMS and GMM-SCMS are capable to extract centerline from three test images, which indicates SCMS and GMM-SCMS are more general than F-SCMS.

The comparison results are quantitatively evaluated in terms of computation time and three accuracy measures. Table 5.1 reports the computational time of three methods on test images. It can be seen that F-SCMS and GMM-SCMS generally achieve higher computational efficiency than SCMS. F-SCMS

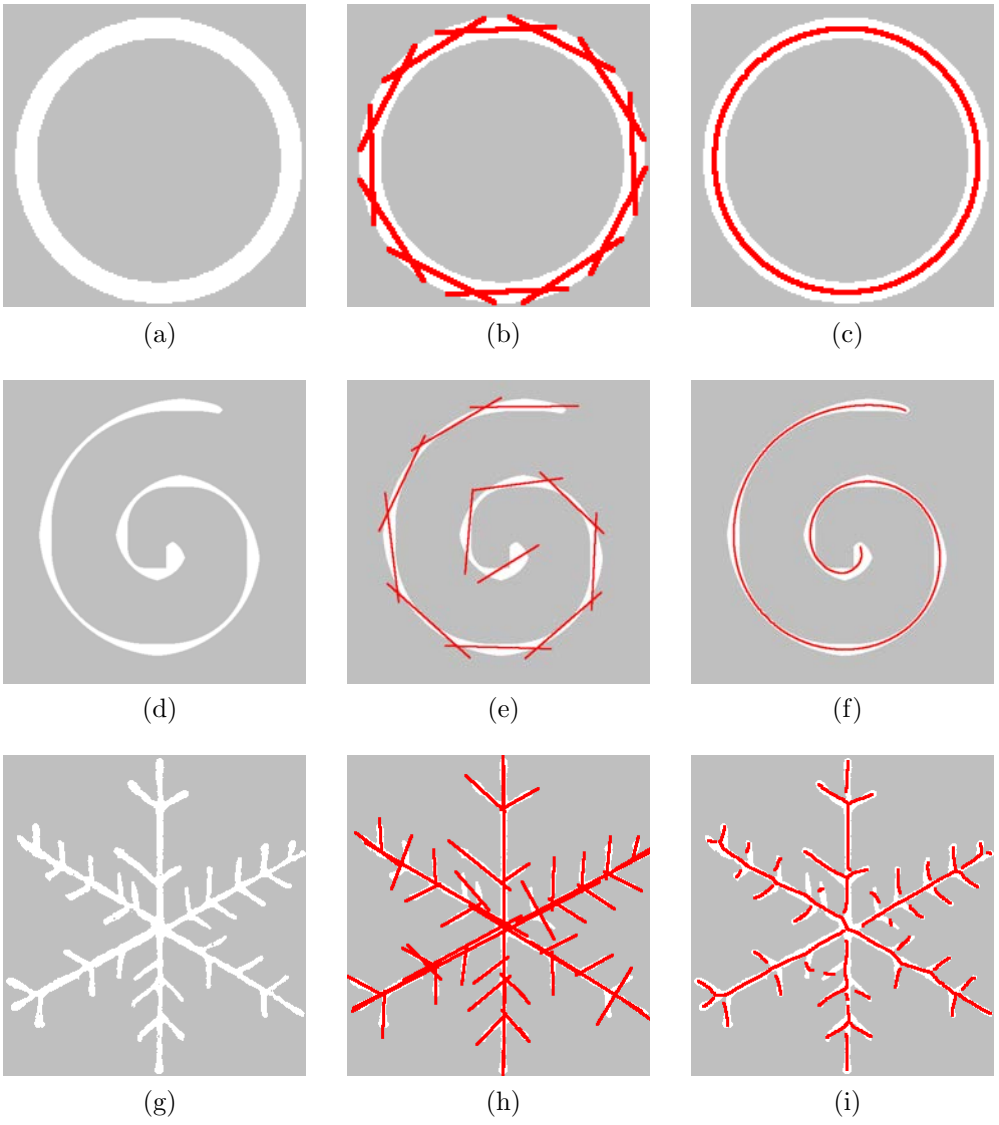


FIGURE 5.11: (a) The test images. (b) The major axis of GMM partitioning result. (c) The adjustment results using SCMS.

is more efficient than GMM-SCMS in Snow case. In spiral case, GMM-SCMS, however, is more efficient than F-SCMS. It seems that the efficiency of GMM-SCMS and F-SCMS is influenced by points number. F-SCMS is more efficient with the large points number while GMM-SCMS in small points number. The accuracy results are given in Table 5.2. From the practical prospective, Circle and Spiral cases are too easy and the Snow case is more similar to the real-world issue. Therefore, the accuracy is just compared in Snow case. It can be seen that F-SCMS produces the highest accuracy among these three methods and SCMS and GMM-SCMS achieve the similar accuracy.

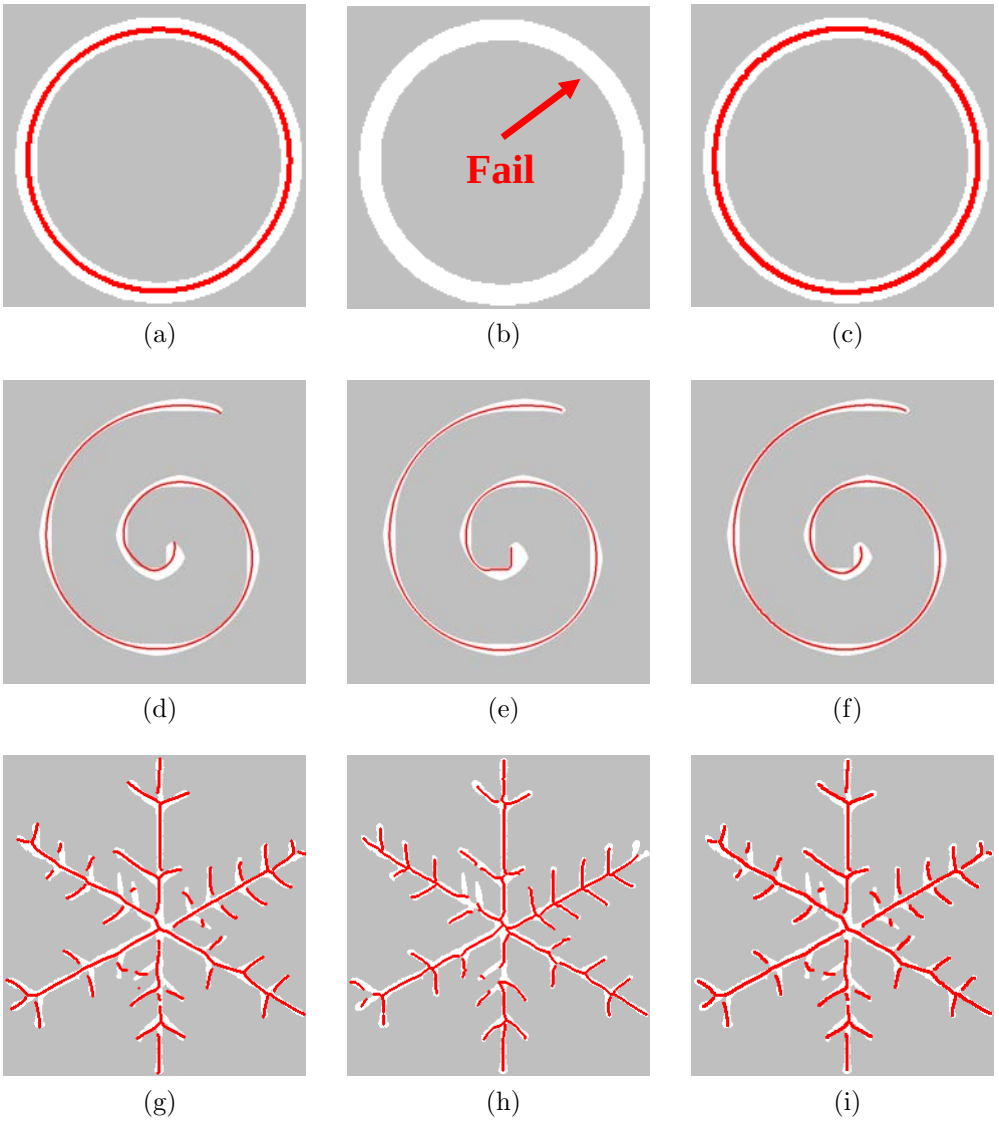


FIGURE 5.12: First to third columns show the results of SCMS, F-SCMS, and GMM-SCMS, respectively. The centerline extracted are shown in red.

5.4.3 Experiments on real satellite images

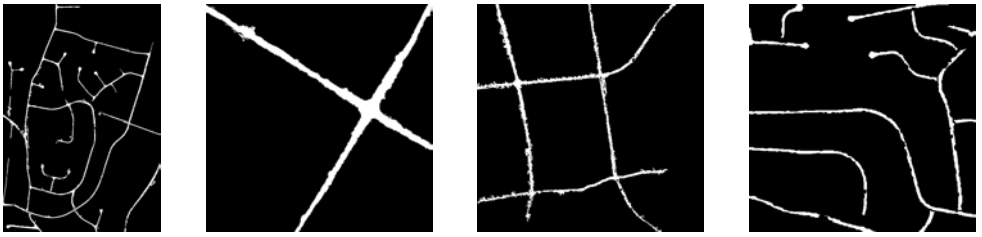
In this section, the proposed method F-SCMS and GMM-SCMS were tested on four multispectral satellite images. The test images are shown in Figure 5.13, where the ground truth datasets are shown in red. The corresponding classified images are presented in Figure 5.14. The proposed GMM-SCMS method is compared with two existing methods in literature: SCMS and F-SCMS. The comparison results are reported in Figure 5.15. The performances of three methods are quantitatively evaluated in terms of the computational

TABLE 5.1: Comparison of computation time for different centerline extraction methods

Experiment	Number of points	Computation time (s)		
		SCMS	F-SCMS	GMM-SCMS
Circle case	5953	20.38	Fail	1.76
Spiral case	18828	114.52	11.24	8.53
Snow case	11807	146.78	8.36	16.42

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

FIGURE 5.13: (a), (b), (c), and (d) show four test images, respectively. The ground truth datasets are shown in red.



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

FIGURE 5.14: (a), (b), (c), and (d) show classified images of the test images in Figure 5.13, respectively.

efficiency and the accuracy, as shown in Table 5.3 and Figure 5.16.

From Table 5.3, it can be seen that the computational load of SCMS is highest among these three methods. By contrast, both F-SCMS and GMM-SCMS produce higher computation efficiency than SCMS. In general, F-SCMS achieves better computational performance than GMM-SCMS, such as Cases 1-3. However, Table 5.3 also shows an interesting phenomenon that GMM-SCMS shows higher computational efficiency than F-SCMS (Miao et al. forthcoming) in Case 4. The phenomenon indicates that the number of connected components (CCs) influences the F-SCMS performance. For instance, the numbers

TABLE 5.2: Quantitative evaluation of different centerline extraction methods

Method	Circle case			Spiral case			Snow case		
	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)	Completeness (%)	Correctness (%)	Quality (%)
SCMS	100	100	100	84.13	85.61	73.71	81.03	90.53	74.70
F-SCMS	Fail	Fail	Fail	93.56	94.75	88.95	90.58	98.84	86.92
GMM-SCMS	100	100	100	96.94	97.74	94.82	78.77	92.36	73.96

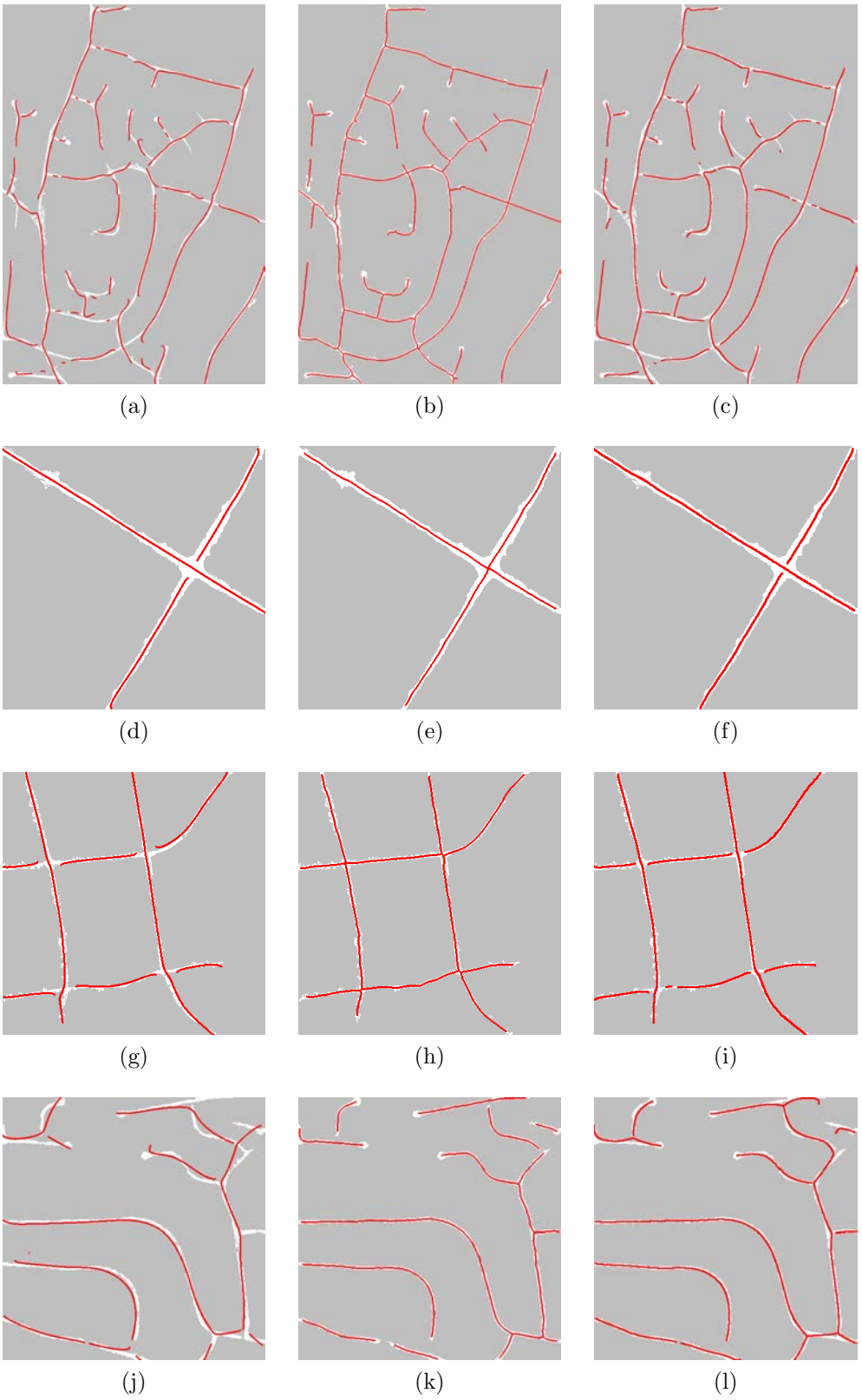


FIGURE 5.15: Comparison results of different road centerline extraction methods. (a) SCMS results. (b) F-SCMS results. (c) GMM-SCMS results. The centerlines are shown in red for display.

TABLE 5.3: Comparison of computation time for different centerline extraction methods

Experiment	Number of points	Computation time (s)		
		SCMS	F-SCMS	GMM-SCMS
1	35177	1142.95	73.48	110.68
2	11578	58.77	2.18	6.51
3	8009	35.73	1.97	8.84
4	16503	172.99	67.93	15.36

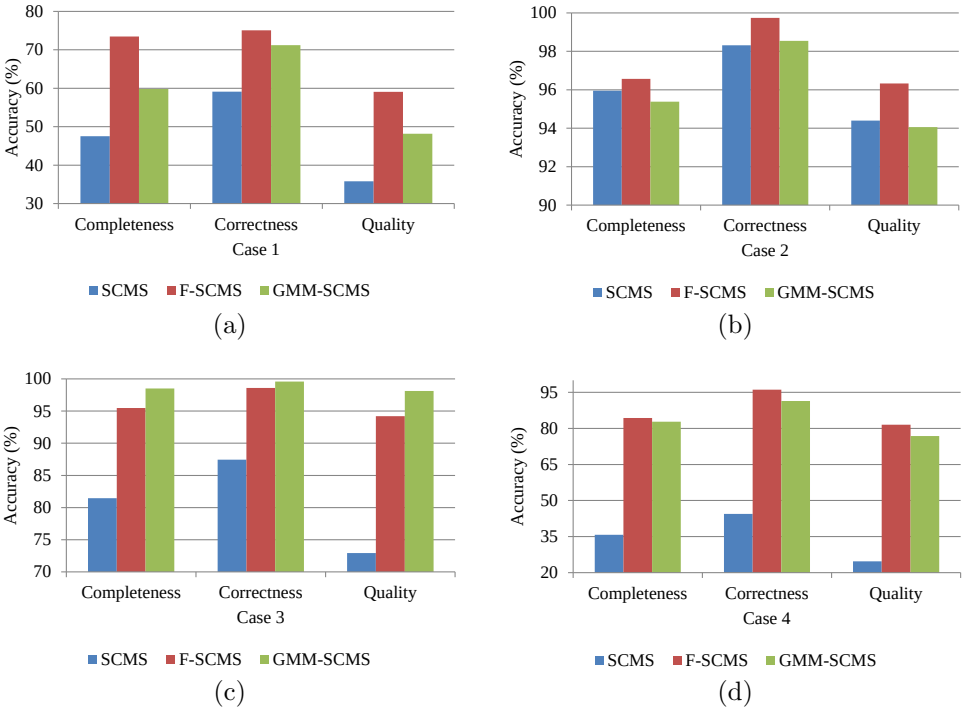


FIGURE 5.16: (a), (b), (c), and (d) show quantitative evaluation results of different centerline extraction methods for Case 1, 2, 3, 4, respectively.

of CCs in Cases 1-4 are 3, 1, 1, and 8, respectively. It seems that when the number of connected components is high, GMM-SCMS is more efficient than F-SCMS. The main reason is that F-SCMS detects feature points (i.e. end and junction points) and then links these points iteratively for each connected component. Detection of feature points from the classified image costs a lot of time. Therefore, high number of CCs decreases the performance of F-SCMS.

Figure 5.16 demonstrates the accuracy of three methods. As can be seen, SCMS produces the lowest accuracy while F-SCMS the highest. In general, GMM-SCMS achieves the modest accuracy among these three methods.

By comparing the proposed method with two existing methods on simulated and real datasets, it can be seen that the proposed method produces the best balance between the computational complexity, accuracy, and generalization ability among these three methods. Thus, GMM-SCMS is more practical for accurate and efficient road centerline extraction from classified images.

5.5 Summary

In this chapter, an automatic method, so-called Gaussian-Mixture-Model-Based Subspace Constrained Mean Shift (GMM-SCMS), was developed for road centerline extraction from the classified image. The GMM-SCMS method can extract smooth road centerlines with high accuracy. Benefited from GMM, the proposed method can represent the classified images using a small number of points while retain the correct spatial topology, which in turn improves the computational efficiency. By using SCMS, the approximate points are iteratively adjusted to produce the accurate road centerline.

The proposed method was compared with cutting-edge SCMS and Feature-Points-Based SCMS (F-SCMS). The experimental results indicate that 1) GMM-SCMS is more efficient than SCMS, and 2) GMM-SCMS has a higher generalize ability than F-SCMS. In other words, GMM-SCMS can process the classified image with arbitrary shape. By fully consideration of efficiency, accuracy, and generalize ability, it can be concluded that GMM-SCMS can provide a practical solution for delineating road centerlines from classified images.

Chapter 6

Road Gap Connection

6.1 Introduction

Although tensor voting (TV) can partly be used to eliminate road discontinuities, it is challenging to process large road gaps, as suggested in Chapter 2. Therefore, this chapter focuses on the issue of large road gap elimination. It should be noted that the problem of large road gap elimination can be converted to a semi-automated road extraction issue, which is relying on road end points, also known as seed points. Although TV is an automatic process, the semi-automated method can guarantee correct road connection topology, which in turn improves road connection accuracy and robustness.

Based on the aforementioned analysis, this chapter proposes a semi-automatic method for road delineation from VHR satellite images. The proposed method is an extension of the geodesic method (Peyré et al., 2010) in which low spatial resolution images are processed. Although the previous integrated method shows certain advantages in processing low spatial resolution images, it is not suitable for VHR images. The proposed method solves this limitation.

6.2 The methodology

6.2.1 The geodesic method

In this chapter, the geodesic method (Peyré et al., 2010) is selected to link seed points and identify the central line for roadways. Let f denote the probability density estimation map, which is modeled as a 2D function $f : \Omega \to \mathbb{R}$. The image domain Ω is usually defined as $\Omega = [0,\ 1]^2$. In road linking, a road that connects two seed points $\boldsymbol{x}_s$ and $\boldsymbol{x}_e$ can be approximately defined as a smooth curve that has a constant grey value $c \in \mathbb{R}$. Based on this definition of the road model, a saliency map $W(\bullet)$ at the pixel $\boldsymbol{x}$ is then defined as:

$$W(\boldsymbol{x}) = |f(\boldsymbol{x}) - c| + \varepsilon \tag{6.1}$$

where ε (i.e., $\varepsilon = 0.01$) is a small value that prevents $W(\boldsymbol{x})$ from vanishing, and c is the grey value at the start point $\boldsymbol{x}_s$ (i.e., $c = f(\boldsymbol{x}_s)$). According to Equation 6.1, the area through which the road passes should have low values of $W(\boldsymbol{x})$.Based on the saliency map $W(\boldsymbol{x})$, the length of a smooth curve on the image $\gamma :\ [0,\ 1] \to \Omega$ is defined as a weighted length as follows:

$$L(\gamma) = \int_0^1 W(\gamma(t)) \|\gamma'(t)\| \, dt \tag{6.2}$$

where $\gamma'(t) \in \mathbb{R}^2$ is the derivative of γ. For linking road seed points, γ is generally constrained by

$$\gamma : [0, 1] \to \Omega \backslash \gamma(0) = \boldsymbol{x}_s \text{ and } \gamma(0) = \boldsymbol{x}_e \tag{6.3}$$

The minimal path γ^* is defined as

$$\gamma^* = \underset{\gamma \in (\boldsymbol{x}_s, \boldsymbol{x}_e)}{\arg\min} L(\gamma) \tag{6.4}$$

The geodesic path is computed on the grey image, which cannot be applied directly to the multispectral image. There are many algorithms to tackle this limitation, such as spectral angle (SA) and principal component analysis (PCA) transformation. For ease of computation in this study, multispectral satellite images are first summed and averaged to obtain a single band image. The experimental results show that this simple step is sufficient. The geodesic method is subsequently applied to the averaged image for extracting the minimal path between seed points. An example of the geodesic method is depicted in Figure 6.1. It can be seen that this method can correctly infer the spatial connection, and there is no need to guarantee the spatial connection topology. A well-recognized problem for the geodesic method is that of bias. Figure 6.1(b) illustrates such a problem, showing a minimal path that tends to follow the boundary of the road, but does not coincide with the true centerline. This limitation must be considered, and the following subsections describe the proposed method for dealing with such issues.

6.2.2 The road probability estimation

The initial road segments generated by the geodesic method are taken as training samples. Then we use the Mahalanobis distance (Hastie et al., 2009) to measure the probability of a pixel $\boldsymbol{x}$ being of the road class, as follows:

$$p_M(\boldsymbol{x}) = \sqrt{(I(\boldsymbol{x}) - S)^T S^{-1} (I(\boldsymbol{x}) - S)} \tag{6.5}$$

where μ and S denote denote the mean value and covariance matrix of training samples, respectively, and $I(\boldsymbol{x})$ is the spectral value at the pixel $(\boldsymbol{x})$. Using the Mahalanobis distance to compute the road probability is valid, as this measure is unit-less and efficient to compute. After the computation of the road probability map, a simple thresholding is applied to segment the images into two classes: road class and non-road class. The thresholding result is

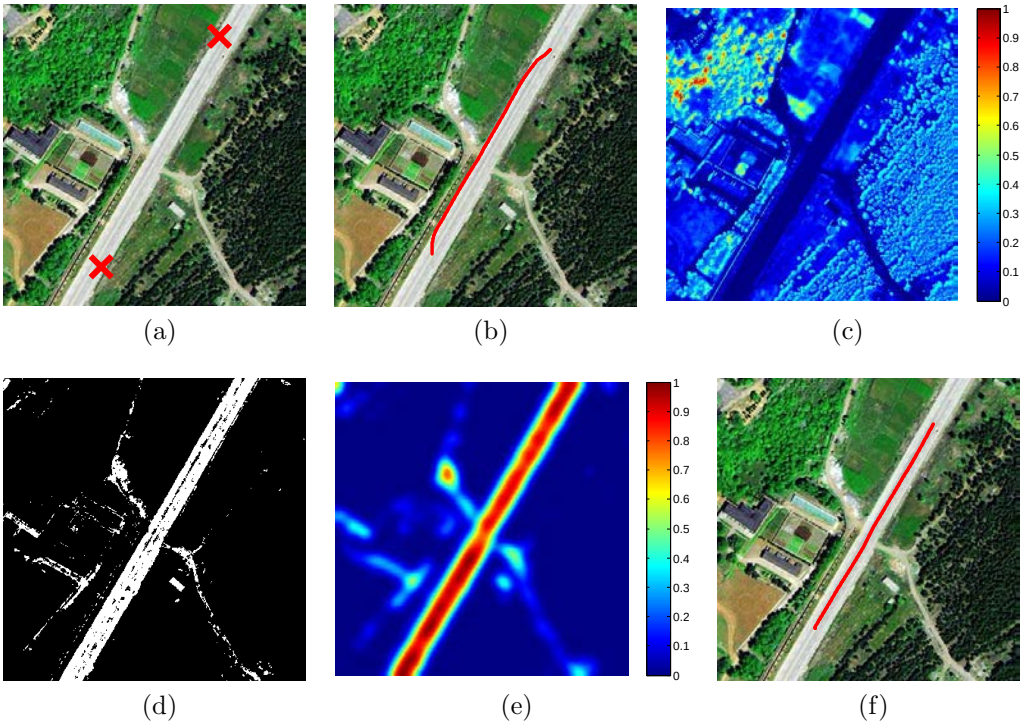


FIGURE 6.1: (a) Test image where seed points are shown in blue crosses. (b) Minimal path extracted by the geodesic method shown in red. (c) Mahalanobis distance map. (d) Thresholding result, in which 1 and 0 representroad and nonroad class, respectively. (e) Result of kernel density estimation. (f) Minimal path extracted by repeating the geodesic method on the KDE result.

defined as

$$\mathrm{CL}(\boldsymbol{x}) = \begin{cases} 1, & \text{if } p_M(\boldsymbol{x}) \leq \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{6.6}$$

where $\mathrm{CL}(\boldsymbol{x})$ is the class label of the pixel $\boldsymbol{x}$, the threshold value is automatically obtained by Otsu's method (Otsu, 1975), and 1 and 0 represent the road class and nonroad class, respectively. Figures 6.1(c) and 6.1(d) shows the Mahalanobis distance matrix and its corresponding thresholding result, respectively. Although the thresholding method misclassifies some road pixels, this error has little influence on the accuracy of the road centerline extraction as the connection of seed points in the proposed method relied on the geodesic method, which is robust to noise.

6.2.3 Kernel density estimation and mean shift

After the road probability estimation, kernel density estimation (KDE) technology (Ahamada and Flachaire, 2010) is introduced to assess the probability that any given pixel lies on the road centerline. Let $y_1, y_2, \ldots, y_n$ be a given set of d-dimensional random samples. The kernel density estimator is defined as

$$\hat{f}(y) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{y - y_i}{h}\right) \tag{6.7}$$

where n is the number of observations, h denotes the bandwidth parameter (which is determined by generalizing Scott's rule of thumb (Ahamada and Flachaire, 2010)), and K is the kernel function. We selected the Gaussian kernel as the kernel function, as follows:

$$K(y) = e^{-\frac{y^2}{2h^2}} \tag{6.8}$$

An example of a KDE result is shown in Figure 6.1(e). It can be seen that pixels on road centerlines have higher KDE values than pixels that are uncentered. It is worth pointing out that seed points selected by users may not precisely locate on road centerlines, and this factor affects the accuracy of the minimal path. To achieve robustness to such distortions, we rely on the mean shift method (Cheng, 1995) to obtain precise centered positions of seed points. Mean shift, a procedure to iteratively detect modes of the kernel function, has the following generic formula:

$$m(y) = \frac{\sum_{i=1}^{n} g\left(\frac{y-y_i}{h}\right) y_i}{\sum_{i=1}^{n} g\left(\frac{y-y_i}{h}\right)} - y \tag{6.9}$$

where $g(y) = -K'(y)$. After projecting road seed points onto the ridge through KDE, the geodesic method is repeated to trace road centerlines using the KDE map, with the result shown in Figure 6.1(f). It can be seen that the proposed method enables us to obtain the centered result.

As discussed in the preceding sections, the procedures of the proposed method are summarized in Algorithm 9.

Algorithm 9 Linking of seed points using the proposed method

Input
The original satellite image.
Seed points selected by users.
Output
The road network.

Manually label seed points on the input satellite image.
Extract the minimal path between seed points using the geodesic method.
Compute the road probability map using the Mahalanobis distance, in which training samples consist of pixels that locate on the minimal path produced by Step 2.
Segment the Mahalanobis distance map by applying a simple thresholding.
Apply kernel density estimation to compute the probability that a pixel locates on the centerline.
Project seed points onto the ridge of the KDE.
Based on the KDE map, repeat the geodesic method to extract the minimal path between seed points.

6.3 Experiment and results

6.3.1 Tests on sensitivity of seed positions

To demonstrate the performance of the proposed method, an aerial image with a spatial resolution of 0.3 m/pixel was tested, and the geodesic method was selected as the reference method. The results are shown in Figure 6.2. As can be observed, the proposed method could provide better results than the geodesic method. The experimental results show that the original geodesic method was sensitive to the positions of seed points, in that different positions for the seed points produced different road centerlines. Thus, the original geodesic method was less robust. The second column in Figure 6.2 also shows that road centerlines produced by the original geodesic path method were wrongly extracted. In particular, even though the seed points were located at centerlines, the original geodesic method still could not achieve correct results.

In contrast, the proposed method was insensitive to the positions of seed points, and could achieve reliable, accurate road centerlines. The proposed method is more accurate as it has two advantages. First, the mean shift can project initial seed points onto the ridge of the KDE. Second, the geodesic path using KDE traces the road centerline along the ridge line. Furthermore, a visual comparison of the results in Figure 6.2 evidently validates the advantage of the proposed method in road centerline extraction. The poor performance of the original geodesic approach demon-strates its inability to extract road centerlines from VHR satellite images. However, as suggested above, the proposed method solves the limitation of the geodesic method and provides a practical solution for the road centerline extraction.

6.3.2 Experiments

A QuickBird image downloaded from VPLab (2015) was used to evaluate the proposed method. The study area had a spatial size of pixels. The satellite image used in this study had a spatial resolution of 0.6 m/pixel. Figure 6.3(a) shows the study area in which the seed points selected by users are shown with blue crosses. The geodesic method was first applied to obtain the minimal path that linked seed points, with the results shown in yellow in Figure 6.3(b). Finally, the proposed method is depicted in red in Figure 6.3(b). A comparison of the two methods is presented in Figure 6.3(c). As can be seen, the result of the proposed method is more centered than that of the geodesic method. Hence, the proposed method more accurately extends the geodesic method to extract road centerlines from VHR images.

In the second experiment, an image with a spatial size of pixels, downloaded from (Laboratory, 2015), was used to test the proposed method's performance. Figure 6.4(a)-6.4(c) show the test image, the geodesic method result, and the proposed method result, respectively. It can be seen that both the geodesic method and the proposed method can achieve results with the desired 'U'

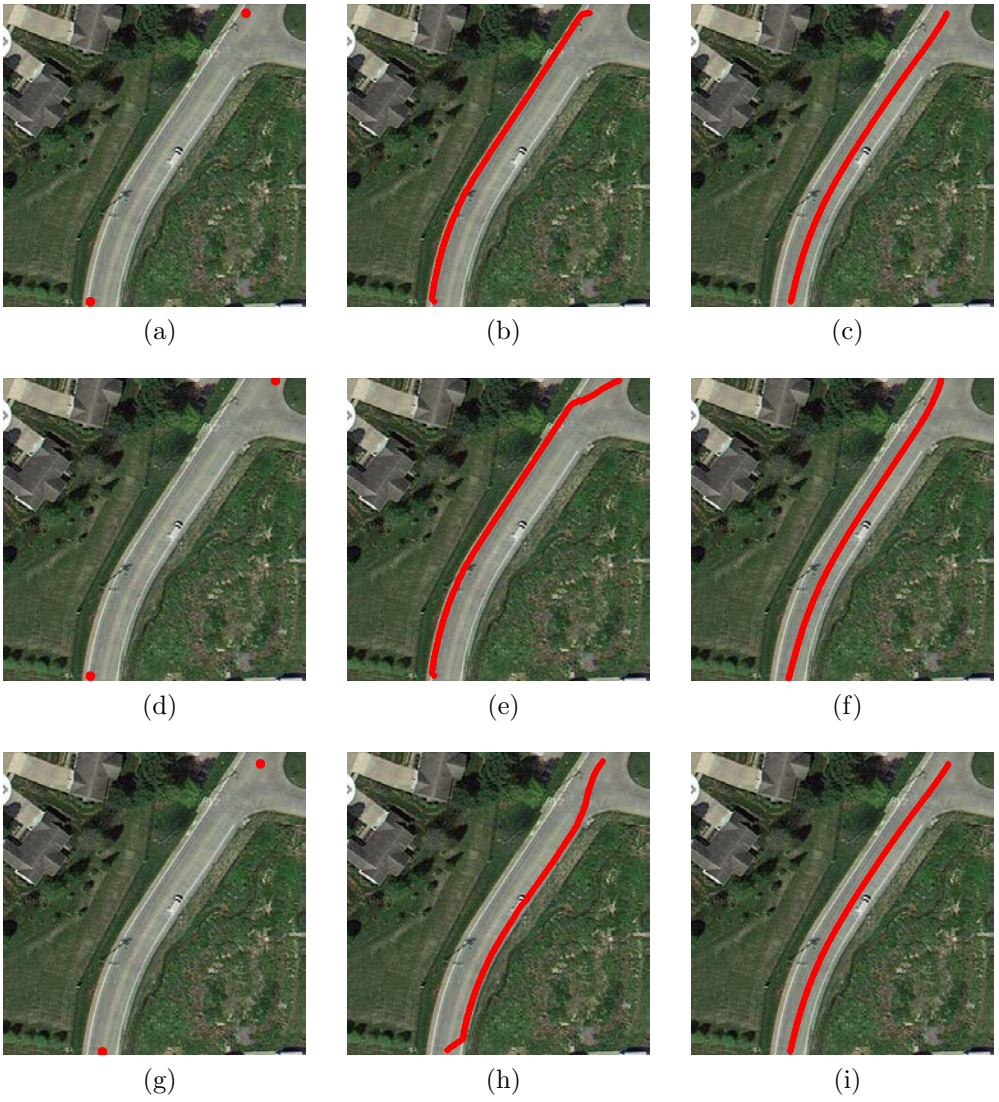


FIGURE 6.2: Comparison of road centerline extraction results. (a) The original images with different seed points, which are shown with red circles. (b) Results by the geodesic method. (c) Results by the proposed method. The centerline results are shown in red.

shape. However, the result of the proposed method is more centered than that of the geodesic method.

The road extraction result for the third test image was obtained using another QuickBird image downloaded from VPLab (2015). This test image, with a spatial size of pixels, is shown in Figure 6.5(a). It can be seen that two seed points are located in regions with different materials. In other words, the grey values of the two seed points are significantly different. The road connection results of the geodesic method and the proposed method are shown in Figure

(a) (b) (c)

FIGURE 6.3: Results on the first image. (a) Test image in which seed points which are shown in blue cross. (b) The geodesic method result which is shown in yellow. (c) The proposed method result which is shown in red. (d) Comparison results of two methods where the geodesic method and the proposed method are shown in yellow and red, respectively.

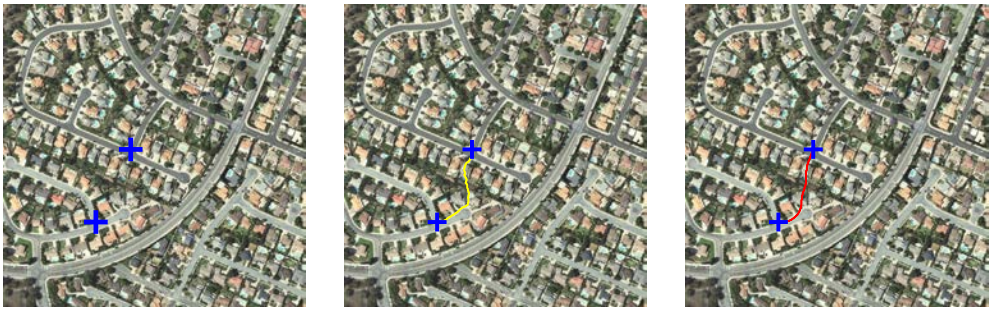
(a) (b) (c)

FIGURE 6.4: Results on the second image. (a) Test image in which seed points which are shown in blue cross. (b) The geodesic method result which is shown in yellow. (c) The proposed method result which is shown in red.

6.5(b) and Figure 6.5(c), respectively. These results show that both methods failed to extract roads with the desired topology. Therefore, the limitation of the proposed method is that it fails in a situation of material change. A possible solution to this problem would be to use the object-space method, based on intensity-change-invariant road shape features.

6.3.3 Comparison with the existing method

The proposed method was compared with a semi-automatic road centerline extraction method drawn from the existing literature. Figure 6.6 depicts comparison results of the proposed method and the method used by Hu et al.



(a) (b) (c)

FIGURE 6.5: Results on the third image. (a) Test image in which seed points which are shown in blue cross. (b) The geodesic method result which is shown in yellow. (c) The proposed method result which is shown in red.

(a) (b)

FIGURE 6.6: Comparison with the existing semi-automatic method. (a) The proposed method results which are shown in red. (b) The results produced by Hu et al. (2004) which are shown in red . Seed points are shown with blue crosses.

(2004). These two methods were selected because both of them rely on seed points selected by users, which guaranteed a fair comparison.

When quantifying the performance values, five accuracy measures (Hu et al., 2004, Wiedemann et al., 1998) were used to evaluate the semi-automatic centerline extraction methods. These measures were 1) number of seed points; 2) total road length; 3) Completeness; 4) Correctness; and 5) Quality. In this chapter, the true road centerline between two seed points was provided by the hand drawing method.

Table 6.1 shows the results of the two methods. It can be seen that the

proposed method and that proposed by Hu et al. achieved similar accuracy. However, the proposed method needed fewer seed points than the Hu et al. method. Hu's method needed more seed points because it requires that seed points be labeled where the road curvature changes. Compared to Hu's method, the proposed method needed fewer interactions with users. Therefore, the proposed method shows a stronger global inference for connection of seed points than Hu's method. The experimental results indicate a great potential of the proposed method for practical delineation of roads from VHR satellite images.

6.4 Summary

In this chapter, the large road gap connection problem is converted to a semi-automatic road extraction issue. To this end, a seed point based method for road centerline extraction from very high resolution satellite images has been proposed. More specifically, the proposed method was shown to incorporate the strengths of the geodesic method, kernel density estimation (KDE), and mean shift. In contrast to the geodesic method, a combination of the geodesic method and KDE provided a better solution for tracing an unbiased road centerline between seed points. The application of the mean shift method allowed the proposed method to obtain precise seed points from coarse points selected by users. These versatile functions, allowed us to extract robust road centerlines that were not subject to the positions of seed points that affect the geodesic method. One of the immediate practical applications of the proposed method is that the road discontinuities of other road extraction algorithms can be eliminated. In its present form, the proposed method is semi-automatic and needs seed points need to be manually selected. Future work will therefore focus on automatic selection of the road seed points, and on the use of graph theory for extracting a complete road network.

TABLE 6.1: Comparison of two semi-automatic road centerline extraction methods

	Experiment 1			Experiment 2	
	Proposed method	Hu's method		Proposed method	Hu's method
Seed point number	2	4	Seed point number	2	6
Total road length	520	487	Total road length	737	760
Completeness (%)	96.54	96.77	Completeness (%)	95.2	99.21
Correctness (%)	96.54	98.47	Correctness (%)	96.88	99.6
Quality (%)	93.31	95.34	Quality (%)	92.37	98.81

Chapter 7

Conclusion

7.1 Summary

In this thesis, road extraction from Very High Resolution (VHR) satellite images has been explored by utilizing a variety of road features (i.e. shape features and spectral information) and computer vision tools. To achieve this goal, the VHR images are firstly segmented/classified to extract road segments. The accurate road centerlines are subsequently extracted from these road segments, followed by a road gap connection step to eliminate road discontinuities. The contributions of this thesis can be concluded in three aspects.

7.1.1 Road segments extraction

The main contributions of this part are as follows:

1. Designing several measures to demonstrate road shape features.
2. Introducing advanced mathematical morphology, which is more suitable for roads than traditional mathematical morphology, to implement spectral-spatial classification to extract road segments from VHR satellite images.

3. Designing an information fusion scheme to integrate local homogeneous regions and classification results to produce a robust road extraction result.

4. Using road shape features to filter road extraction result to remove noise which in turn improves the extraction accuracy.

7.1.2 Accurate road centerline extraction from the classified images

The second part attempts to extract accurate road centerline from the classified road images. In this study, two methods are designed based on the geodesic method:

1. The first method integrates the tensor voting and the geodesic method. The combination of tensor voting and the geodesic method is able to extract the accurate road centerline that achieves the good balance between the accuracy and computational load.

2. The second method is using the k-Means and the geodesic method. This method has a loose spatial constraint than the first one and can guarantee a similar accuracy.

7.1.3 Road gap connection

The third study concerns road gap connection. In this thesis, two means are studied to achieve this goal, as listed in the follow:

1. The first method is involving the use of tensor voting. The advantage of tensor voting is that it is automatic and has a strong global connection inference. Therefore, there is no need to set connection hypothesis in advance. Another advantage of tensor voting is that it only has one parameter to be tuned, and hence it is a robust method.

2. The second means is using the geodesic method in a semi-automatic manner. The geodesic require the user to provide two end points to connect road gaps. Experimental results show that the geodesic method can generally produce correct results with the minimum interactions with users.

7.2 Future research

The main recommendations of this thesis are list as follows:

1. Any road shape features has its own limitation when it is used alone. To obtain a robust road shape feature measure, a more complicated and universe road shape feature should be designed.

2. Despite the good performance of using advanced mathematical morphology, such as path openings and closings and General Adaptive Neighborhood, there is still a large room to improve the spectral-spatial classification accuracy of road extraction. More road shape features should be explored and integrated to the existing methods.

3. Local methods should be considered to make the road centerline extraction in this thesis more general.

4. Tensor voting can be implemented with a small scale parameter in an iterative manner to solve the cross-talk phenomenon.

5. Tensor voting and the geodesic method can be integrated to eliminate road discontinuities in a more robust and fast manner.

Appendix A

Efficient GAN Dilation

Algorithm 10 Pseudo-code of the Proposed Efficient GAN Erosion

Input
The original grey image I_O
The homogeneity tolerance value m
Output
The GAN erosion result I_D

Initial I_D and set all the elements in I_D to be 0.
for all $i = 0$ to 255 **do**
    Set $T_i = i - m$.
    Set $T_s = i + m$.
    Extract the area A_i that satisfies $I_O \geq T_i$ and $I_O \leq T_s$.
    Extract the seed points set S_i that satisfies $I_O = i$.
    Run MorphologicalReconstruction(S_i, A_i, 8) to extract the GANs I_m.
    Run LabelConnectedComponents(I_{mj}, 8) to compute the connected components number n_r in I_m.
    Set $T_{s2} = T_s$.
    if $T_{s2} > 255$ **then**
        $T_{s2} = 255$
    end if
    for $j = T_{s2}$ to i **do**
        Let $R = I_O$ and set all the elements in R to be 0.
        Extract the seed points set S_{ij} that satisfies I_O==j and $I_m \neq 0$.
        Run MorphologicalReconstruction(S_{ij},I_m, 8) to extract the area I_{mj}.
        for all points x, y in I_{mj} **do**
            if $I_{mj}(x, y) \neq 0$ and $I_D(x, y) \leq j$ **then**
                $I_D(x, y) = j$
            end if
        end for

Algorithm 11 Pseudo-code of the Proposed Efficient GAN Erosion (continued)

Run LabelConnectedComponents(I_{mj},8) compute the connected components number n_{rj} in I_{mj}.
if $n_r - n_{rj} \neq 0$ **then**
$n_r \leftarrow n_r - n_{rj}$
else
break
end if
end for
end for

Appendix B

Efficient GAN Dilation

Algorithm 11 Pseudo-code of the Proposed Efficient GAN Dilation

Input
The original grey image I_O
The homogeneity tolerance value m
Output
The GAN erosion result I_D

Initial I_D and set all the elements in I_D to be 0.
for all $i = 0$ to 255 **do**
    Set $T_i = i - m$.
    Set $T_s = i + m$.
    Extract the area A_i that satisfies $I_O \geq T_i$ and $I_O \leq T_s$.
    Extract the seed points set S_i that satisfies $I_O = i$.
    Run MorphologicalReconstruction(S_i, A_i, 8) to extract the GANs I_m.
    Run LabelConnectedComponents(I_{mj}, 8) to compute the connected components number n_r in I_m.
    Set $T_{s2} = T_s$.
    if $T_{s2} > 255$ **then**
        $T_{s2} = 255$
    end if
    for $j = T_{s2}$ to i **do**
        Let $R = I_O$ and set all the elements in R to be 0.
        Extract the seed points set S_{ij} that satisfies I_O==j and $I_m \neq 0$.
        Run MorphologicalReconstruction(S_{ij},I_m, 8) to extract the area I_{mj}.
        for all points x, y in I_{mj} **do**
            if $I_{mj}(x, y) \neq 0$ and $I_D(x, y) \leq j$ **then**
                $I_D(x, y) = j$
            end if
        end for

Algorithm 12 Pseudo-code of the Proposed Efficient GAN Dilation (continued)

```
22:         Run LabelConnectedComponents(I_mj,8) compute the connected
    components number n_rj in I_mj.
23:         if n_r − n_rj ≠ 0 then
24:             n_r ← n_r − n_rj
25:         else
26:             break
27:         end if
28:     end for
29: end for
```

Bibliography

Ibrahim Ahamada and Emmanuel Flachaire. Non-parametric econometrics. *OUP Catalogue*, 2010.

Luc Anselin. Local indicators of spatial association-lisa. *Geographical Analysis*, 27(2):93–115, 1995.

Martin Baatz and Arno Schäpe. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII*, pages 12–23, 2000.

GH Ball and DJ Hall. Isodata, a novel method of data analysis and classification, technique report, 1965.

EP Baltsavias. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3): 129–151, 2004.

Ursula C Benz, Peter Hofmann, Gregor Willhauck, Iris Lingenfelder, and Markus Heynen. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3):239–258, 2004.

Munuele Bicego, Silvio Dalfini, Gianni Vernazza, and Mttorio Murino. Automatic road extraction from aerial images by probabilistic contour tracking. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 3, pages III–585. IEEE, 2003.

Thomas Blaschke. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16, 2010.

Thomas Blaschke and Josef Strobl. What's wrong with pixels? some recent developments interfacing remote sensing and gis. *GeoBIT/GIS*, 6(01):12–17, 2001.

Thomas Blaschke, Stefan Lang, and Geoffrey Hay. *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Springer Science & Business Media, 2008.

Christopher JC Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

Matthias Butenuth and Christian Heipke. Network snakes: graph-based object delineation with active contour models. *Machine Vision and Applications*, 23(1):91–109, 2012.

Chuqing Cao and Ying Sun. Automatic road centerline extraction from imagery using road gps data. *Remote Sensing*, 6(9):9014–9033, 2014.

Dengfeng Chai, Wolfgang Forstner, and Florent Lafarge. Recovering line-networks in images by junction-point processes. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1894–1901. IEEE, 2013.

Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

Yizong Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):790–799, 1995.

Casey Cleve, Maggi Kelly, Faith R Kearns, and Max Moritz. Classification of the wildland–urban interface: A comparison of pixel-and object-based classifications using high-resolution aerial photography. *Computers, Environment and Urban Systems*, 32(4):317–326, 2008.

Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

Aluir P Dal Poz, Rodrigo AB Gallis, João FC da Silva, and Erico FO Martins. Object-space road extraction in rural areas using stereoscopic aerial images. *IEEE Geoscience and Remote Sensing Letters*, 9(4):654–658, 2012.

Sukhendu Das, TT Mirnalinee, and Koshy Varghese. Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10):3906–3931, 2011.

Johan Debayle and Jean-Charles Pinoli. General adaptive neighborhood image processing. *Journal of Mathematical Imaging and Vision*, 25(2):245–266, 2006.

Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

Peter Doucette, Peggy Agouris, Anthony Stefanidis, and Mohamad Musavi. Self-organised clustering for road extraction in classified imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(5):347–358, 2001.

Manfred Ehlers, Monika Gaehler, and Ronald Janowsky. Automated techniques for environmental monitoring and change analyses for ultra high resolution remote sensing data. *Photogrammetric Engineering and Remote Sensing*, 72(7):835–844, 2006.

Mathieu Fauvel, Jón Atli Benediktsson, Jocelyn Chanussot, and Johannes R Sveinsson. Spectral and spatial classification of hyperspectral data using

svms and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3804–3814, 2008.

Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

Jan Flusser. Moment invariants in image analysis. In *Proceedings of World Academy of Science, Engineering and Technology*, volume 11, pages 196–201, 2006.

Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Interventation-MICCAI' 98*, pages 130–137. Springer, 1998.

Robert C Frohn, Kenneth M Hinkel, and Wendy R Eisner. Satellite remote sensing classification of thaw lakes and drained thaw lake basins on the north slope of alaska. *Remote Sensing of Environment*, 97(1):116–126, 2005.

Paolo Gamba, Fabio Dell'Acqua, and Gianni Lisini. Improving urban road extraction in high-resolution images exploiting directional filtering, perceptual grouping, and simple topological concepts. *IEEE Geoscience and Remote Sensing Letters*, 3(3):387–391, July 2006.

Rafael C Gonzalez, Richard Eugene Woods, and Steven L Eddins. *Digital Image Processing Using MATLAB*. Gatesmark, Knoxville, TN, USA, 2nd ed edition, 2009.

Anne Grote, Christian Heipke, and Franz Rottensteiner. Road network extraction in suburban areas. *The Photogrammetric Record*, 27(137):8–28, 2012.

Armin Gruen and Haihong Li. Semi-automatic linear feature extraction by dynamic programming and lsb-snakes. *Photogrammetric Engineering and Remote Sensing*, 63(8):985–994, 1997.

Youkyung Han, Hyejin Kim, Jaewan Choi, and Yongil Kim. A shape–size index extraction for classification of high resolution multispectral satellite images. *International Journal of Remote Sensing*, 33(6):1682–1700, 2012.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd ed edition, 2009.

Martin Herold, Joseph Scepan, André Müller, and Sylvia Günther. Object-oriented mapping and analysis of urban land use/cover using ikonos data. In *22nd Earsel Symposium Geoinformation for European-Wide Integration*, pages 4–6, 2002.

Jiuxiang Hu, Anshuman Razdan, John C Femiani, Ming Cui, and Peter Wonka. Road network extraction and intersection detection from aerial images by tracking road footprints. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):4144–4157, 2007.

Xiangyun Hu, Zuxun Zhang, and C Vincent Tao. A robust method for semi-automatic extraction of road centerlines using a piecewise parabolic model and least square template matching. *Photogrammetric Engineering & Remote Sensing*, 70(12):1393–1398, 2004.

Xin Huang and Liangpei Zhang. Road centreline extraction from high-resolution imagery based on multiscale structural features and support vector machines. *International Journal of Remote Sensing*, 30(8):1977–1987, 2009.

Xin Huang and Liangpei Zhang. An svm ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1):257–272, 2013.

Jungho Im, JR Jensen, and JA Tullis. Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, 29(2):399–423, 2008.

Ron Kimmel. Numerical geometry of images: Theory, algorithms, and applications. 2004.

Computer Vision Laboratory. Data. http://cvlab.epfl.ch/data, 2015. Accessed: 2015-09-17.

Andrea S Laliberte, Albert Rango, Kris M Havstad, Jack F Paris, Reldon F Beck, Rob McNeely, and Amalia L Gonzalez. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern new mexico. *Remote Sensing of Environment*, 93(1):198–210, 2004.

Shawn Lankton and Allen Tannenbaum. Localizing region-based active contours. *IEEE Transactions on Image Processing*, 17(11):2029–2039, 2008.

Shawn M Lankton. *Localized statistical models in computer vision*. PhD thesis, Georgia Institute of Technology, 2009.

Ivan Laptev, Helmut Mayer, Tony Lindeberg, Wolfgang Eckstein, Carsten Steger, and Albert Baumgartner. Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications*, 12(1):23–31, 2000.

Zhongbin Li, Wenzhong Shi, Qunming Wang, and Zelang Miao. Extracting man-made objects from high spatial resolution remote sensing images via fast level set evolutions. *IEEE Transactions on Geoscience and Remote Sensing*, 2015.

XG Lin, JX Zhang, ZJ 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. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34:309–312, 2008.

Xiangguo Lin, Jixian Zhang, Zhengjun Liu, Jing Shen, and Minyan Duan. Semi-automatic extraction of road networks by least squares interlaced template matching in urban areas. *International Journal of Remote Sensing*, 32(17):4943–4959, 2011.

Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.

Bibo Lu, Yongxia Ku, and Hui Wang. Automatic road extraction method based on level set and shape analysis. In *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, volume 3, pages 511–514. IEEE, 2009.

Ramesh Marikhu, Matthew N Dailey, Stanislav Makhanov, and Kiyoshi Honda. A family of quadratic snakes for road extraction. In *Computer Vision–ACCV 2007*, pages 85–94. Springer, 2007.

Helmut Mayer, Stefan Hinz, Uwe Bacher, and Emmanuel Baltsavias. A test of automatic road extraction approaches. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, 36(3): 209–214, 2006.

Gerard Medioni, Mi-Suen Lee, and Chi-Keung Tang. *A Computational Framework for Segmentation and Grouping*. Elsevier, New York, USA, 2000.

Juan B Mena. State of the art on automatic road extraction for gis update: a novel classification. *Pattern Recognition Letters*, 24(16):3037–3058, 2003.

Zelang Miao and Wenzhong Shi. Road centreline extraction from classified images by using the geodesic method. *Remote Sensing Letters*, 5(4):367–376, 2014.

Zelang Miao, Wenzhong Shi, Hua Zhang, and Xinxin Wang. Road centerline extraction from high-resolution imagery based on shape features and multivariate adaptive regression splines. *IEEE Geoscience and Remote Sensing Letters*, 10(3):583–587, 2013.

Zelang Miao, Bin Wang, Wenzhong Shi, and Hao Wu. A method for accurate road centerline extraction from a classified image. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(12): 4762–4771, 2014a.

Zelang Miao, Bin Wang, Wenzhong Shi, and Hua Zhang. A semi-automatic method for road centerline extraction from vhr images. *IEEE Geoscience and Remote Sensing Letters*, 11(11):1856–1860, 2014b.

Zelang Miao, Bin Wang, Shi Wenzhong, Hao Wu, and Yiliang Wan. Use of gmm and scms for accurate road centerline extraction from the classified image. *Journal of Sensors*, 2014c.

Zelang Miao, Wenzhong Shi, Paolo Gamba, and Zhongbin Li. An object-based method for road network extraction in vhr satellite images. 2015. doi: 10.1109/JSTARS.2015.2443552.

Volodymyr Mnih and Geoffrey E Hinton. Learning to detect roads in high-resolution aerial images. In *Computer Vision–ECCV 2010*, pages 210–223. Springer, 2010.

Volodymyr Mnih and Geoffrey E Hinton. Learning to label aerial images from noisy data. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 567–574, 2012.

Mehdi Mokhtarzade and MJ Valadan Zoej. Road detection from high-resolution satellite images using artificial neural networks. *International Journal of Applied Earth Observation and Geoinformation*, 9(1):32–40, 2007.

Vincent Morard, Etienne Decenciere, and Petr Dokládal. Efficient geodesic attribute thinnings based on the barycentric diameter. *Journal of Mathematical Imaging and Vision*, 46(1):128–142, 2013.

Philippos Mordohai and Gérard Medioni. Tensor voting: a perceptual organization approach to computer vision and machine learning. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1):1–136, 2006.

Sahar Movaghati, Alireza Moghaddamjoo, and Ahad Tavakoli. Road extraction from satellite images using particle filtering and extended kalman filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7):2807–2817, 2010.

Yoichi Nakaguro, Stanislav S Makhanov, and Matthew N Dailey. Numerical experiments with cooperating multiple quadratic snakes for road extraction. *International Journal of Geographical Information Science*, 25(5):765–783, 2011.

Matteo Negri, Paolo Gamba, Gianni Lisini, and Florence Tupin. Junction-aware extraction and regularization of urban road networks in high-resolution sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2962–2971, 2006.

Richard Nock and Frank Nielsen. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1452–1458, 2004.

Richard Nock and Frank Nielsen. On weighting clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1223–1235, 2006.

Ali Ozgun Ok, Caglar Senaras, and Baris Yuksel. Automated detection of arbitrarily shaped buildings in complex environments from monocular vhr optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1701–1717, 2013.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

Umut Ozertem and Deniz Erdogmus. Nonparametric snakes. *IEEE Transactions on Image Processing*, 16(9):2361–2368, 2007.

Umut Ozertem and Deniz Erdogmus. Principal curve time warping. *IEEE Transactions on Signal Processing*, 57(6):2041–2049, 2009.

Umut Ozertem and Deniz Erdogmus. Locally defined principal curves and surfaces. *The Journal of Machine Learning Research*, 12:1249–1286, 2011.

Ting Peng, Ian H Jermyn, Véronique Prinet, and Josiane Zerubia. Incorporating generic and specific prior knowledge in a multiscale phase field model for road extraction from vhr images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(2):139–146, 2008.

Martino Pesaresi and Jon Atli Benediktsson. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2):309–320, 2001.

Renaud Péteri, Julien Celle, and Thierry Ranchin. Detection and extraction of road networks from high resolution satellite images. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–301. IEEE, 2003.

Gabriel Peyré, Mickaël Péchaud, Renaud Keriven, and Laurent D Cohen. Geodesic methods in computer vision and graphics. *Foundations and Trends® in Computer Graphics and Vision*, 5(3–4):197–397, 2010.

J-C Pinoli and Johan Debayle. General adaptive neighborhood mathematical morphology. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 2249–2252. IEEE, 2009.

J-C Pinoli and Johan Debayle. Spatially and intensity adaptive morphology. *IEEE Journal of Selected Topics in Signal Processing*, 6(7):820–829, 2012.

Charalambos Poullis and Suya You. Delineation and geometric modeling of road networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (2):165–181, 2010.

William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

M Rajeswari, KS Gurumurthy, SN Omkar, J Senthilnath, and L Pratap Reddy. Automatic road extraction using high resolution satellite images based on level set and mean shift methods. In *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, volume 2, pages 424–428. IEEE, 2011.

Marie Rochery, J Zerubia, et al. New higher-order active contour energies for network extraction. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II–822. IEEE, 2005.

J Senthilnath, Shivesh Bajpai, SN Omkar, PG Diwakar, and V Mani. An approach to multi-temporal modis image analysis using image classification and segmentation. *Advances in Space Research*, 50(9):1274–1287, 2012.

Jean Serra. Introduction to mathematical morphology. *Computer vision, graphics, and image processing*, 35(3):283–305, 1986.

Yuanzheng Shao, Bingxuan Guo, Xiangyun Hu, and Liping Di. Application of a fast linear feature detector to road extraction from remotely sensed imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(3):626–631, 2011.

Wenzhong Shi and Changqing Zhu. The line segment match method for extracting road network from high-resolution satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 40(2):511–514, 2002.

Wenzhong Shi, Zelang Miao, and Johan Debayle. An integrated method for urban main-road centerline extraction from optical remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6):3359–3372, 2014a.

Wenzhong Shi, Zelang Miao, Qunming Wang, and Hua Zhang. Spectral–spatial classification and shape features for urban road centerline extraction. *IEEE Geoscience and Remote Sensing Letters*, 11(4):788–792, 2014b.

Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4):427–437, 2009.

Mingjun Song and Daniel Civco. Road extraction using svm and image segmentation. *Photogrammetric Engineering and Remote Sensing*, 70(12): 1365–1371, 2004.

Yuliya Tarabalka, Jón Atli Benediktsson, and Jocelyn Chanussot. Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (8):2973–2987, 2009.

Engin Türetken, Germán González, Christian Blum, and Pascal Fua. Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors. *Neuroinformatics*, 9(2-3):279–302, 2011.

Engin Türetken, Fethallah Benmansour, and Pascal Fua. Automated reconstruction of tree structures using path classifiers and mixed integer programming. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 566–573. IEEE, 2012.

Engin Turetken, Fethallah Benmansour, Bjoern Andres, Hanspeter Pfister, and Pascal Fua. Reconstructing loopy curvilinear structures using integer programming. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1822–1829. Ieee, 2013.

Silvia Valero, Jocelyn Chanussot, Jon Atli Benediktsson, Hugues Talbot, and Björn Waske. Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters*, 31(10):1120–1127, 2010.

VPLab. Downloads. http://www.cse.iitm.ac.in/~sdas/vplab/satellite.html, 2015. Accessed: 2015-05-15.

Jan Dirk Wegner, Javier Montoya-Zegarra, Kaspar Schindler, et al. A higher-order crf model for road network extraction. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1698–1705. IEEE, 2013.

Christian Wiedemann, Christian Heipke, Helmut Mayer, and Olivier Jamet. Empirical evaluation of automatically extracted road axes. *Empirical Evaluation Techniques in Computer Vision*, pages 172–187, 1998.

David H Wolpert and William G Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997.

World Bank and the Development Research Center of the State Council, P. R. China. *Urban China: Toward Efficient, Inclusive, and Sustainable Urbanization*. World Bank, Washington, DC, 2014. ISBN 0195209931, 9780195209938.

World Bank Group. *East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth*. World Bank, Washington, DC, 2015. ISBN 0195209931, 9780195209938.

Yuan Yuan, Yishan Luo, and Albert Chung. Ve-lli-vo: Vessel enhancement using local line integrals and variational optimization. *IEEE Transactions on Image Processing*, 20(7):1912–1924, 2011.

Nima Zarrinpanjeh, Farhad Samadzadegan, and Toni Schenk. A new ant based distributed framework for urban road map updating from high resolution satellite imagery. *Computers and Geosciences*, 54:337–350, 2013.

Qiaoping Zhang and Isabelle Couloigner. A wavelet approach to road extraction from high spatial resolution remotely-sensed imagery. *Geomatica*, 58 (1):33–39, 2004.

Qiaoping Zhang and Isabelle Couloigner. Accurate centerline detection and line width estimation of thick lines using the radon transform. *IEEE Transactions on Image Processing*, 16(2):310–316, 2007.

Jun Zhou, Walter F Bischof, and Terry Caelli. Robust and efficient road tracking in aerial images. In *Joint Workshop of ISPRS and the German Association for Pattern Recognition, Vienna, Austria, 29-30 August*, pages 35–40, 2005.