Semi-automatic Road Extraction from Very High Resolution Remote Sensing Imagery by RoadModeler

by

Yao Lu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Geography

Waterloo, Ontario, Canada, 2009

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Accurate and up-to-date road information is essential for both effective urban planning and disaster management. Today, very high resolution (VHR) imagery acquired by airborne and spaceborne imaging sensors is the primary source for the acquisition of spatial information of increasingly growing road networks. Given the increased availability of the aerial and satellite images, it is necessary to develop computer-aided techniques to improve the efficiency and reduce the cost of road extraction tasks. Therefore, automation of image-based road extraction is a very active research topic.

This thesis deals with the development and implementation aspects of a semi-automatic road extraction strategy, which includes two key approaches: multidirectional and single-direction road extraction. It requires a human operator to initialize a seed circle on a road and specify a extraction approach before the road is extracted by automatic algorithms using multiple vision cues. The multidirectional approach is used to detect roads with different materials, widths, intersection shapes, and degrees of noise, but sometimes it also interprets parking lots as road areas. Different from the multidirectional approach, the single-direction approach can detect roads with few mistakes, but each seed circle can only be used to detect one road. In accordance with this strategy, a RoadModeler prototype was developed. Both aerial and GeoEye-1 satellite images of seven different types of scenes with various road shapes in rural, downtown, and residential areas were used to evaluate the performance of the RoadModeler. The experimental results demonstrated that the RoadModeler is reliable and easy-to-use by a

non-expert operator. Therefore, the RoadModeler is much better than the object-oriented classification. Its average road completeness, correctness, and quality achieved 94%, 97%, and 94%, respectively. These results are higher than those of Hu et al. (2007), which are 91%, 90%, and 85%, respectively. The successful development of the RoadModeler suggests that the integration of multiple vision cues potentially offers a solution to simple and fast acquisition of road information. Recommendations are given for further research to be conducted to ensure that this progress goes beyond the prototype stage and towards everyday use.

.

Acknowledgements

First of all, I am extremely grateful to Professor Dr. Jonathan Li, as my supervisor, for his support, expertise and enthusiasm for this work. His efforts to ensure my financial support during my studies are also appreciated. Much of my research style and philosophy, as well as my background in both remote sensing and computer vision, can be traced directly to Dr. Jonathan Li. His constant stream of ideas, insights and suggestions has been invaluable throughout this study.

The financial supports of the University of Waterloo and the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery Grant awarded to Dr. Jonathan Li, for my M. Sc. study, are also greatly appreciated.

Special thanks also go to Professor Dr. Michael A. Chapman, as my thesis committee member, for his constructive problem solving discussions in the past two years. I would like to express my gratitude to the two thesis readers, Professor Dr. David Clausi and Professor Dr. Alexander Bernning for their critical comments and valuable suggestions on my thesis.

Many thanks go to the Remote Sensing and Geospatial Technology team: Yu Li, Yipeng Yuan, Yuanming Shu, Gary Goems, Hamad Yousif and Flora Guan for their collaboration and assistance. My research has profited from their teamwork. I would further like to thank all staff members in the Department of Geography and Environmental Management,

particularly, Ms. Lynn Finch and Ms. Diane Ridler who helped me greatly in various ways. In particular, I express my heartfelt thanks to Ms. Anne Grant at the Mapping, Analysis and Design (MAD) services of the Faculty of Environment for her technical support in terms of providing various datasets and maintaining the computing facilities in the Remote Sensing and Geospatial Technology Lab.

Finally, and most importantly, I would like to express my indebtedness to my parents and my oldest aunty. They have never failed to support me in all endeavours, and this thesis has been no exception. Without them, I would not have been able to succeed in this M. Sc. study.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

This chapter outlines the research problems and challenges in automatic road extraction from very high resolution (VHR) satellite and aerial images. Next the defined research objectives ensue. Lastly, the thesis structure is outlined.

1.1 Statement of Existing Problems

Accurate and up-to-date geospatial information about road networks is of great importance for effective urban and transportation planning, land development, and urban disaster management (Mena, 2003). Road information is normally integrated into a Geographical Information System (GIS) database for effective management. Due to rapid urban development, the GIS database needs to be updated with timely and accurate road network information.

Using traditional ground survey techniques to collect road data is out of date as it is labour-intensive and time-consuming, particularly in mapping large urban areas. However, remote sensing has proven to be a powerful technology for spatial data collection and change detection. Along with the development of innovative sensors and platforms, road network spatial information can be acquired from aerial and satellite imagery, which includes optical imagery, Synthetic Aperture Radar (SAR) imagery, Light Detection and Ranging (LiDAR)

range data, and image sequences taken from land-based mobile mapping systems with different spatial and spectral resolutions (Quackenbush, 2004).

The preferred technology for road extraction is VHR satellite imagery, as it covers larger areas than aerial imagery and provides updated information on a regular basis. Moreover, it is more economic than either terrestrial or aerial imaging technologies. Furthermore, unlike SAR and LIDAR data collection, it allows for the extraction of more detailed and accurate road information.

Satellite and aerial imagery is raw data, and rarely contains information in an explicit form. It is impossible to use the imagery directly without information extraction. Accurate and efficient information extraction is crucial in spatial data applications. Manually extracting roads from satellite imagery, although easy along simple stretches, is impractical and time consuming when the scenes are very complex. However, not only are such complex maps required for large geographic areas, but frequent updating is needed (Gibdaughter, 2003). Therefore, the automatic extraction of satellite imagery is necessary to obtain accurate and current maps from complex satellite images.

Research on extracting roads from aerial and satellite images can be traced back to the pioneers Bajcsy and Tavaloki (1976). Even though research in this field began over three decades ago, advances in image processing and computer vision remain inadequate. Many automatic and semi-automatic algorithms have been attempted, but flawed, and never

became operational. Further research on an automatic or semi-automatic road extraction method is needed to reach the operational stage.

Clearly there is still no single automatic method that is universally applicable to all types of roads from satellite imagery (Quackenbush, 2004). Furthermore, in order to recognize a road automatically, automatic methods must check every pixel in the entire image, which is time consuming, especially for large volumes of VHR imagery. As people are able to quickly, accurately, and effortlessly identify road areas under a variety of conditions, it is practical to have a human operator, rather than a computer, provide information about road conditions, such as starting points and directions. Starting points are used as seed points and starting directions assist road detection (Vosselman and de Knecht, 1995). An algorithm is then used to predict the route in incremental steps until it reaches a stopping criterion. Thus, the running time will depend on the area of the road not the area of the whole image. In addition, with the assistance of an operator, most roads in various environmental conditions could be detected by computer. Moreover, results from semi-automatic methods are more accurate than results from fully automatic methods (Gruen and Li, 1997). In summary, semi-automatic methods for road extraction are more practical than automatic methods.

1.2 High Resolution Imagery for Road Extraction

Nowadays, optical imagery, especially VHR satellite imagery, has received considerable attention because it provides accurate, spatial information. Table 1.1 shows the main

parameters of currently operational and future optical VHR satellite systems. All of the VHR satellites mentioned in Table 1.1 simultaneously collect panchromatic (Pan) and multispectral (MS) images at higher and lower spatial resolution, respectively. Most MS images are taken within the visible and near-infrared wavelength (VNIR) and are recorded into 3 or 4 multispectral bands (Zlatanova and Li, 2008).

Table 1.1 List of currently operational and future optical VHR satellite systems

Optical satellite	Spatial resolution (m) and (#bands)		Swath (km)	Repeat cycle (days)	Year launch
	PAN	VNIR			
IKONOS	1	4(4)	11	3	1999
QuickBird 2	0.6	2.5(4)	16	3	2001
Orb View-3	1	4(4)	8	3	2003
KOMPSAT-2	1	4(4)	15	28	2004
Resurs DK-1	1	2-3(3)	4.7-28.3	6	2006
WorldView-1	0.55	1 (4)	17.6	1. 7-5.9	2007
GeoEye-1	0.41	1.65(4)	15.2	1-3	2008
WorldView-2	0.25	1 (4)	16.4	1-4	2009
Pleiades-1 and 2	0.7	2.8(4)	20	1-2	2009-2010

Source: (Zlatanova and Li, 2008)

The image-fusion techniques have been developed to combine Pan and MS images to generate a high-resolution, pan-sharpened-colour image. To date, the PCI Pansharp module produces the best Pan-MS fusion results among all commercially available software tools (PCI, 2004; Gorin, 2005). By image pan-sharpening, the highest-resolution coloured satellite

imagery available is GeoEye-1 imagery, which reaches 0.5m spatial resolution. Currently, DigitalGlobe is developing a VRH imaging satellite named WorldView with a higher spatial resolution of 0.25m (Pan), and 1m (MS), respectively. Furthermore, there will be 8 MS bands in total within the VNIR range. This occurrence proves that spatial resolution may be further improved; eventually matching the resolution of aerial imagery (around 0.15m). It is vital to understand aerial orthoimage in the event that spatial resolution of satellite imagery reaches the level of aerial photographs. Thus, the images used in this thesis will include pan-sharpened colour GeoEye-1 images and aerial orthoimages.

1.3 Challenges in Automatic Road Extraction

Road characteristics in general can be classified as radiometrical, geometrical, topological, and contextual characteristics (Vosselman and Knecht, 1995). Together, they correspond with the challenges listed below:

(1) Radiometrical characteristics and challenges: The road surface is made from different materials (e.g., asphalt and cement). On the satellite and aerial imagery, the road surface displays different greyscales representing the construction phase of the road. Additionally, adjacent regions are often different. Sometimes, even building roofs and parking lots appear to similar spectral information to roads. When multiplied and inconsistent intensity of roads arise and spectral similarity of neighbourhood areas increase, creating radiometrical models for roads will become more difficult.

(2) Geometric characteristics and challenges: Since their lengths are far larger than their widths, roads appear as elongated regions in VHR imagery. Roads vary in width and curvatures. In each road, the width and direction change smoothly. Various road shapes may increase the difficulty of building geometric models.

(3) Topological characteristics and challenges: Roads form networks when they link certain places together. In some areas, roads and surrounding buildings are connected through driveways and alleys. This may cause confusion between them on the VHR imagery since they both carry similar physical characteristics.

(4) Context characteristics and challenges: The appearances of vehicles, pedestrian lines, overpasses and shadows cast by trees and buildings can produce a negative influence on homogeneous intensity and connectivity of roads. In some cases, shadows may cover certain sections of a road. On the other hand, pedestrian crossings and trees located on roadsides may help to imply the presence of roads on a VHR image.

To build models of roads in a large area is difficult as proven by the characteristics and challenges listed above. Methods that had already been developed for medium to coarse imagery may not apply to VRH imagery. This is because there is more noise (e.g. vehicles, traffic lines, shadows, etc.) to lead to blockage problem in VRH imagery (Niu, 2006). Furthermore, there is another problem. It is difficult to separate the certain roads from their surroundings with similar spectral information, leading to leakage problem (Niu, 2006). Thus, a new methodology for road extraction in VHR imagery is required.

1.4 Research Objectives

The overall purpose of this study is to develop a semi-automatic method towards a prototype software tool called RoadModeler, which can be used to extract diverse types of roads from VHR imagery. Specifically, the objectives of the study are as follows:

(1) To analyze the spatial complexity and heterogeneity of roads in VHR imagery;

(2) To design and implement a semi-automatic road strategy through a combination of digital image analysis techniques in order to develop an operational software tool called RoadModeler. Such a tool should be accurate, reliable and easy to use.

(3) To assess the performance of the RoadModeler through a qualitative and quantitative evaluation method on a set of VHR images of various types of roads.

1.5 Thesis Organization

The rest of the thesis consists of the following five chapters:

Chapter 2 provides a review of previous studies on road extraction from VHR imagery.

Chapter 3 presents a theoretical background, framework and explanation of the proposed methodology, followed by a detailed description of the developed software tool named RoadModeler.

Chapter 4 illustrates the results of the proposed RoadModeler with different test images. A comparison and discussion between the developed RoadModeler and the object-oriented classification methods are included in the experimental results.

Chapter 5 presents a qualitative and quantitative evaluation method. It also focuses on an evaluation of the developed RoadModeler's performance.

Chapter 6 provides the conclusions based on the findings of this study and recommendations for future research.

Chapter 2
Related Work on Automatic Road Extraction

This chapter reviews the principle of semi-automatic method for road extraction from VHR imagery. Previous research regarding methods and techniques are reviewed.

2.1 Automatic Road Extraction Method: An Introduction

Research on extracting roads from aerial and satellite images can be traced back to the work of Bajcsy and Tavaloki in 1976. Many methods have been proposed during the past three decades. Conventionally, three steps are required for road extraction from imagery (Trinder and Wang, 1998): road finding, road tracking, and road linking. Generally, when human interaction is involved in the first step (road finding), the algorithm is referred to as semi-automatic. If no human interaction is required, the algorithm is considered as a fully automatic approach (Niu, 2006). Many techniques rely on preprocessing to enhance edges or lines, or segment the imagery into homogeneous regions (Guindon, 1998).

Roads, in medium and coarse resolution imagery (e.g., 10~30m), appear as curvilinear structures, while in higher resolution imagery they appear as homogenous regions that satisfy certain shape or size constraints (Hinz et al., 2001). Methods such as morphological operators, ratio edge detector, edge detector, Hough transform have been used to extract curvilinear structures from medium and coarse resolution imagery is not appropriate for road extraction

in VHR imagery (Hu et al., 2007; Quackenbush, 2004). Therefore, the emphasis in this chapter is placed on the study of road extraction from VHR imagery, and the relevant problems are discussed.

In the last three decades, there have been many studies on automatic or semi-automatic road extraction from aerial and satellite imagery. As a result, many strategies, methodologies and algorithms for road extraction were presented which had reached various degree of success (Mena, 2004). All these studies firstly described the characteristics of the road reflected from the specific remotely sensed data. After that, a feature model was built based on the characteristic information. The more information explored, the better results can be obtained. According to the applied extraction techniques, the existing road extraction methods using VHR remotely sensed images can be classified into: multi-resolution techniques, segmentation and classification methods, artificial intelligence approaches, snakes, and road tracking methods (Mena, 2003).

2.2 Multi-Resolution Techniques

As the objects are represented with more details in VHR imagery, the task of road finding requires complicated methods for the abstraction of these details. In contrast, roads in medium and coarse resolution imagery can be extracted by using simple road models. Line extraction at low resolution can guide detailed image analysis at high resolution and restrict it

to image parts with high probability of containing a road. Many multi-resolution approaches first generate lower resolution imagery by degrading a high resolution image.

Shneier (1982) successively applied a 2*2 filter over the image to replace the four-pixel neighbourhood with the median value, creating a pyramid of images with progressively lower resolution from which lines was extracted by a line detector for identifying roads in the higher resolution imagery. Each line in the lower resolution image was assumed to correspond to an elongated region in the original image, and could be utilized to identify the position and the extent of these regions. As an alternative, Couloigner and Ranchin (2000) used a wavelet transform to generate pyramid layers.

Instead of degradating high resolution imagery for multi-resolution analysis, some researchers integrate different types of satellite imagery for extraction. Bonnefon *et al*. (2002) used SPOT imagery to approximately identify linear features which are used to identify roads in IKONOS imagery.

Trinder and Wang (1998) found distinct advantages by combining the abstraction of the coarser scale with the detailed information found at the finer scale. Pairs of edges in high resolution imagery were first identified and then combined with lines in lower resolution (re-sampled) imagery in order to fully extract the road network.

Baumgartner et al. (1999) combined roadsides extracted from the original high resolution image (0.2–0.5m) and lines extracted from an image of reduced resolution to build the hierarchical road structure. The road area can be identified if a pair of edges are approximately parallel, have an approximately homogeneous region between them, and have a corresponding line in the reduced resolution image.

Road detection in low resolution is the key step for guiding road extraction in high resolution. Multi-resolution road extraction is appropriate for rural area where the majority of roads are easily extracted from low resolution image but not for high density downtown areas where extra edges which represent roads are extracted in low resolution imagery.

2.3 Segmentation and Classification Methods

Segmentation is the automated process of partitioning an image into several clusters that are homogeneous with respect to some characteristics such as colour, texture, reflection signal, or context, etc. (Chen et al., 2002). Classification is the process of assigning segmented individual pixels or homogeneous clusters to specific and more meaningful information classes (Asano et al., 1996). These two techniques are used for road extraction in order to obtain a binary image where the road network is depicted. Therefore, many researchers have chosen to include these techniques in their road extraction methods (Mena, 2003). Baraldi and Parmiggiani (1994) followed an Iterative Self-Organizing Data Analysis Technique (ISODATA) classification with defining a model to extract linear features by using geometric

information such as thin, elongated regions. The ISODATA algorithm is unsupervised classification. It can refine the clusters automatically by merging similar clusters which have a few pixels or whose centers are very close and by splitting clusters with large standard deviations (Jensen, 2005). Doucette, et al. (1999) found that GIS data in a neural network could be used to extract linear features after performing a maximum likelihood classification on HYDICE imagery.

Gardner, et al. (2001) used Multiple Endmember Spectral Mixture Analysis (MESMA, Roberts et al., 1998) to classify the urban areas to extract road networks. However, the roofs are misclassified as roads because of the similarities of the spectral information of the roads and roofs. Then, they used Q-tree filter to improve the result.

Agouris, et al. (2001) extracted road from multispectral imagery by using a novel technique of spatiospectral cluster analysis, in which unsupervised classification, a K-Medians algorithm, was first applied to segment the image, and a Three-stage Fuzzy Inference System determined which cluster was belong to road class by checking whether a cluster demonstrates sufficient characteristics of elongated regions.

Chen et al. (2002a) introduced inductive learning methods to derive rules at three levels: parameter learning, algorithm learning, and cluster learning from training data. Then the rules were combined to be applied on the new image to extract roads. In their method, parameter learning is to select appropriate parameters for K-Means and K-Nearest-Neighbour

algorithms; algorithm learning is to choose more suitable algorithm for a given image; and cluster learning is to separate the road clusters from non-road clusters.

Mena and Malpica (2002) represented a new supervised classification mainly based on the texture analysis named Texture Progressive Analysis on three RGB bands of real image. Here, the Dempster - Shafer theory of evidence was applied in order to fuse the information from texture analysis to achieve a binary segmentation. The road could be extracted after the selecting of road training data by a operator.

Amini et al. (2002) also used a segmentation called split and merge algorithm (SMA). They firstly generated a simpler image by Grey scale Morphological Algorithms (GMA), and then SMA was applied on the simplified image, which would be converted to a binary image. Next, the binary image map objects were labelled using the connected component analysis (CCA) algorithm. After that, the straight line segments and roadsides were extracted by applying chain coding and perceptual grouping. Finally, the skeleton of roads could be extracted from in the binary reduced resolution image by using SMA and Binary Morphological Algorithms (BMA).

Haverkamp (2002) extracted road networks from road objects and intersection objects based on size, eccentricity, length of the objects and spatial relationships between neighbouring intersection objects. After the a vegetation mask derived from multispectral IKONOS imagery, these objects were generated by grouping pixels with similar road directional

information based on texture analysis in panchromatic IKONOS imagery. This method requires extreme predetermination that the straight roads have a specific level of contrast and a low along-road variance.

Daughterg and Civco (2004) employed the support vector machine to classify IKONOS imagery into a road group and a non-road group. Then, a region growing technique was applied on the road group image to generate geometrically homogeneous objects. A simple thresholding on the shape measurements derived from these objects was performed to extract road features. Finally, a thinning process and a vectorization procedure were carried out to obtain road centerlines. However, road gaps caused by shadow or obscuring land features still exist in the results.

Long and Zhao (2005) utilized the mean shift algorithm to clean and segment a red band of IKONOS image to extract roads in urban area. A classification step was based on the analysis of a grey histogram of the segmented image and a grey threshold was used to extract road areas. After that, morphological opening and closing were applied to remove small objects and connect useful objects. Finally, a convex hull for each street-block region was extracted as road-side edges. However, the successful extraction strictly assumed that the greylevel values in the road do not change dramatically and distribute in a narrow value range.

Zhang et al. (2006) proposed a novel road identification approach on VHR multi-spectral imagery. It integrated a traditional k-means clustering for segmentation, a fuzzy logic

classifier for road area identification and the angular texture signature for separating the roads from the parking lots that had been misclassified as roads. This approach is only tested successfully on the scenes where there is no vehicle scattered in the parking lots, because of the almost same width and length. However, when there are a lot of vehicles in it, the detected area is not the entire parking lots but part of it which shows a shape similar to road. In this case, parking lots are still misclassified as road.

Most traditional classification algorithms, such as ISODATA, perform classification using spectral analysis based on single pixels. The feature extraction techniques utilized by the eCognition® software consider image objects rather than single pixels when performing classifications (Definiens, 2008). Figure 2.1 shows the framework of the object-oriented classification approach. A multi-resolution segmentation is first used to produce image objects, so as to calculate the features which could be used to classify these objects to a specific class. Another segmentation and classification will be operated until the best result could be obtained. In these processes, the objects produced by segmentation will be propitious to subsequent classification, and the classified objects will also advantage the subsequent segmentation (Definiens, 2008).

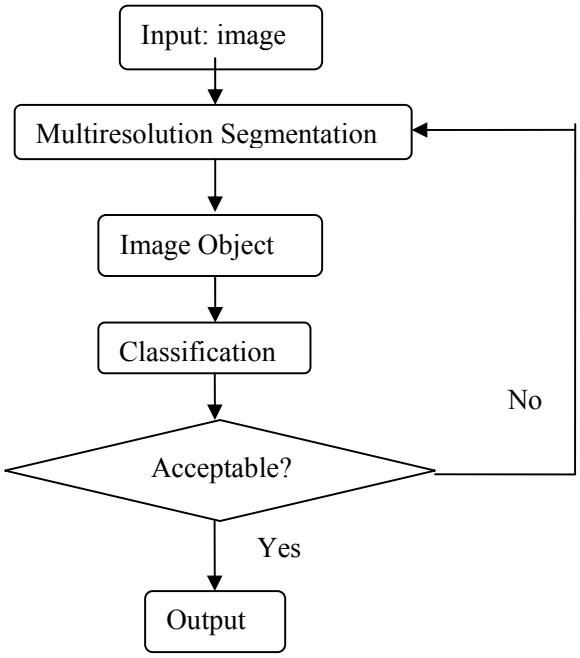


Figure 2.1 Flowchart of object-oriented classification

Accordingly, this method is appropriate for the extraction of roads from high resolution imagery. All the studies mentioned above have the same steps as common grounds, which are to segment the image, to build a model to define a requirement based on the road characteristics, and to classify the road class by clusters which satisfy the specific requirement. However, the segmentation algorithms need to scan the whole image, and every pixel needs to be proceeded at least once, costing much more time. Additionally, only when the road areas have a homogeneous colour, the segmentation can generate satisfied segments for subsequent classification. If the road does not have obvious different colour from its neighbour area, the satisfied segments cannot be obtained, which makes the subsequent classification step difficult to generate an appropriate criterion to classify.

2.4 Artificial Intelligence Approaches

Artificial intelligence approaches convert human knowledge into an exploitable form, aiming to carry out reasoning in the same way as a human being (Pigeon, *et al*., 1999a). In this way, combining rules from a variety of sources to build a road model allows the computer to provide correct, flexible and effective results. Fuzzy logic, neural networks and genetic algorithms are the most popular mathematical tools used in artificial intelligence systems (Mohammadzadeh et al., 2006).

Mohammadzadeh et al. (2006) proposed a fuzzy logic approach to detect main road centerline from pan-sharpened IKONOS images. Depending on the complexity of the scenes, one or up to three pixels might be manually selected from the road surface and the RGB grey levels of these pixels were considered as initial values. The best mean values in each band from fizzification processing were used in defuzzification to generate a segmented image. After that, morphological algorithm was applied to the segmented image for small paths removal, holes filing, and centerline extraction. However, road covered by large shadows cannot be extracted in this method. Moreover, this fuzzy model is mainly based on spectral information, which may be ascribed to failure for complicated scenes.

Mokhtarzade and Zoej (2007) used artificial neural networks for road detection from high-resolution satellite image. The discrimination ability of the network is highly affected by the choosing of input parameters. After testing multi-spectral IKONOS and Quick-Bird images,

the network can be empowered in road detection by inputting neighbour pixels and in background detection by inputting the distance of each pixel to the road mean vector. The successful detection should base on the assumption that roads are homogeneous area. These two methods make good use of three bands information to extract road networks. The more characteristics of road are modeled, the more accurate results will be obtained. However, only spectral and few shape information is used in these two models, therefore, they are only appropriate for simple scenes where roads are salient.

2.5 Snakes

The concept “snake”, also called “active contour model” was first introduced by (Kass et al 1987), which has been used to seek any shape in the image that was smooth and forms a closed contour. Since the boundaries of road network were of diverse shapes including various degrees of curvature, snakes were well suited for this task. The extraction process was started by initializing a curve, called snake, close to the object boundary by the operator. Then, the curve was associated with an objective function which combined internal smoothness constraints such as bending of a curve with image forces like the gradient. By optimizing the objective function iteratively, the curve started deforming and moved towards the desired object boundary. In the end, it completely “shrink-wraped” around the object (Kass et al 1987).

However, traditional snakes are extremely sensitive to parameters and its convergence is

dependent on initial position. Moreover, it has a small capture range since no external force acts on those points far away from the boundary. Furthermore, it fails to detect concave boundaries because external force cannot pull control points into boundary concavity. Since road networks have various degrees of curvature a close initialization often cannot be provided. As a result, traditional snakes can easily get stuck in an undesirable local minimum (Niu, 2006).

Neuenschwander et al. (1997) proposed the ziplock snake model allowing a user to only specify the distant end points of the curve, simplifying far less initialization effort. The image information around the end points was used to provide boundary conditions, Moreover, when a snake was near its extremities, the image information would be taken into account at first. Accordingly, this modified snake model yields excellent convergence properties for the snakes even if the initialization is far away from the solution. They also extended the snake-based approach to ribbon snakes to extract roads in rural area. The extraction process could be achieved by optimizing the position and the width of the ribbon. However, it will be stopped in the presence of disturbances.

Gruen and Li (1997) used either a dynamic programming approach or LSB-Snakes (Least Squares B-spline Snakes) to extract linear feature after providing a few seed points. Dynamic programming was used to build the cost function among the seeds and solve optimization problems to extracts roads with the parameter model. LSB-Snakes which combine least-squares template matching (Gruen, 1985) and B-spline Snakes (Trinder and Li, 1995)

improved the performance of active contour models and controlled blunders such as occlusions very well.

Mayer et al. (1998) and Laptev et al. (2000) used ribbon snakes to extracted salient roads based on the detected lines at a coarse scale and the variation of the road width at a fine scale. Non-salient roads were extracted by connecting two adjacent ends of salient roads with a road hypothesis, which was then verified based on homogeneity and the constancy of width. Finally, a closed snake was initialized inside the central area of the junction and expanded until delineating the junction borders. Nevertheless, cars, traffic islands and road markings in urban area can block the snake's movement; therefore, this approach is more intended for rural areas.

Amo et al. (2006) used the region competition algorithm to extract roads from aerial images. The region competition was a mixed approach which combined region growing techniques with active contour model. Region growing made the first step faster and region competition delivers more accurate results. However, this method is only appropriate for handling roads in agricultural fields, where roads are quite homogeneous and their homogeneity is sufficiently different from that of their surroundings.

Niu (2006) presented a semi-automatic framework for highway extraction based on a geometric deformable model which referred to the minimization of an objective function that connects the optimization problem with the propagation of regular curves. After the seed

points were placed at the end of the highway segments, the framework would incorporate the shape information of a highway segment into the seed point propagation scheme, thus it successfully prevented the leakage problem and blockage problem. However, it is based on the assumption that the highway is a continuous ribbon with no sudden ends or sharp turns, therefore, this method is not appropriate for roads in residential areas.

Most snake methods are applied on rural area, where roads' colour is homogeneous inside and different from neighbour areas. In urban areas, various types of features which often present inside the road areas in VHR imagery, such as cars, traffic islands, road markings, shadows cast by trees or buildings, which can block the snake's movement. Furthermore, expanding snakes can pass over weak junction borders, leading to some leakage problems.

2.6 Road Tracking Methods

Semi-automatic road tracking by template matching methods seems to be more useful in operational applications due to the participation of human operators in tracing (Baumgartner et al., 2002; Zhou et al., 2006). This kind of method is an iterative road segment growing process starting from a set of seed points automatically or manually selected. The templates can be categorized into two classes: profile (1-dimension cross-section of the image intensities taken orthogonal to the direction of a road segment) and rectangular template (2-dimension cross-section of the image intensities taken orthogonal to the direction of a road

segment) based on dimension of road template. Template matching compares a reference template with the template at a pixel to predict the road direction.

McKeown and Denlinger (1988) introduced a road tracking method based on multiple cooperative methods which included a surface correlation tracker and a road edge tracker. When one method failed at one point, the alternative tracking method could be used to continue the tracking process. However, these two trackers have a lot of limitations, because they can work only when the road has constant width, gradually or suddenly changing intensity profile, slowly changing direction and so on.

Vosselman and Knecht (1995) extracted roads based on least squares matching of grey value profiles; Kalman filter was also used to continue the process when the profile matching failed. The two cooperative can trace roads with intersections, flyovers, and vehicle. This method can settle complex situations such as occlusion by trees, shadows which are left to the operator.

Baumgartner et al. (2002) built a graphical user interface for a profile matching method in a style of method mentioned in Vosselman and Knecht (1995). Therefore, the operator can monitor the tracking processing, and receive report which describe the problems occurred during the tracking through this user interface. This tracking tool is fast, but only appropriate for simple scene such as rural areas.

Shukla et al. (2002) applied a path following method to extract road form high-resolution satellite image by initializing two points to indicate the road direction. Scale space, Edge-detection techniques were used as pre-processing for segmentation and estimation of road width. The cost minimization technique was used to determine the road direction and generate next seeds. This method is better than the work of Park and Kim (2001) because it can generate seeds in different directions at intersection. The limitations are that the algorithm may not work on the road cast by shadows.

Zhao et al. (2002) imposed a semi-automatic method by matching a rectangular road template with both road mask and road seeds to extract roads from IKONOS imagery. Road mask is the road pixels generated from maximum likelihood classification, and the road seeds can be generated by tracing the long edge of the road mask. The problem is that neither all of extracted road mask are road area nor all of the extracted long edges are road edge, which would result in misclassification.

Hu et al. (2004) presented a semi-automatic road extraction method based on a piecewise parabolic model with 0-order continuity, which was constructed by seed points placed by a human operator. Road extraction became the problem of estimating the unknown parameters for each piece of parabola, which could be solved by least square template matching based on the deformable template and the constraint of the geometric model. In densely populated areas, where roads have sharp turns and orthogonal intersections, a plenty of seed points are needed to be located, resulting in a degrading the efficiency.

Kim et al. (2004) used least squares correlation matching to extract road centerlines from IKONOS images based on the orientation of the initial seed calculated through Burns line extraction algorithm and a road template built around the seed. The limitations of this algorithm are that it cannot work on the road with shadows which may terminate the tracking process, the initial seed must be selected on road central lines by the operator, and one seed can only extract road with one direction, leading to too many seeds when the scene is large and complex.

Zhou et al. (2006) proposed a framework which consists of the user, a human–computer interface, computer vision algorithms, knowledge transfer schemes and an evaluation criterion for semi-automatic road extraction. Extended Kalman filter and particle filter were applied to solve profile matching issues for road tracking to enhance the robustness of the tracker. Two profiles were required, one perpendicular to the road direction and the other one parallel to the road direction.

Another tracking method is based on texture signature. Hu et al. (2007) extracted road networks from aerial images by tracking road footprints obtained by a spoke wheel operator based on texture information. The first road footprint was generated from a road seed initialized by an operator or automatically generated based on rectangular approximations. Then, a toe-finding algorithm was used to classify footprints for growing a road tree. Finally, a Bayes decision model based on the area-to-perimeter ratio of the footprint was applied to prune the paths that leaked into the surroundings. This method can successfully extract

various shapes of roads and intersections, while the footprints often fail to be generated because of the marking lines or shadows from trees or buildings.

The road tracking methods are mostly semiautomatic at which the running time is proportional to the area of the road rather than the area of the whole image, and as such they are all efficient. However, most of the road trackers mentioned above would fail when they encountered blockage problems caused by radiometric changes due to shadows, vehicle congestions, pavement changes, lane markings, overpasses, etc. (Niu, 2006).

2.7 Summary of Existing Methods

In conclusion, most methods can extract salient roads in simple scenes such as rural areas. While for some complicated scenes, various problems may occur in different kinds of degree, reducing the correctness or completeness of the road extraction. In multi-resolution, results from the medium and coarse resolution images bring more effectiveness for the road finding processing in VHR imagery, while results from VHR imagery provide more detailed information. Nevertheless, this method still cannot detect roads covered by big shadows, because no clue can be obtained from different resolution images used for detecting that blockage area. Segmentation and classification methods usually separate the images into several homogeneous segments, and propose a rule to detect road area, while the noisy road can affect the generated segments. Consequently, no satisfied rule can be made to extract all road areas. Moreover, the running time of segmentation is usually very long, because every

pixel in the whole image needs to be calculated to get a homogeneous segment. Artificial intelligence approaches usually builds a model, which includes rules with different weights, the area which satisfies all the rules to some degree can be considered road area, and nevertheless, rules for very complex scenes are difficult to be exploited into some exploitable form that can be performed in computers. Snake methods reduce the road extraction process as the optimization of an objective function and it can detect the roads with various shape but not the ones with too much noise. Road tracking method tracks the road by comparing it with the template to determine the road direction, thus only pixels on the road area and around it are calculated, however, blockage on the roads may lead the stop of tracking, and not all detected areas which matches with the template are guaranteed as the road area. Based on the discussions above, road tracking method demonstrates its superiority over others in practical use for VHR imagery, and the human intervention can be imported to resolve its problems in this study.

Chapter 3
Semiautomatic Road Extraction by RoadModeler

This chapter describes the research approach applied to develop a semi-automatic road extraction strategy and a prototype tool, the RoadModeler. Firstly, Section 3.1 entails RoadModerler strategy. Section 3.2 describes grey-like colour in the Red Green Blue (RGB) colour space. From Sections 3.3 to 3.8, the proposed strategy is further explained in detail.

3.1 RoadModeler Strategy

In VHR aerial and satellite images, roads are continuous and elongated homogeneous regions with nearly constant width. Most tracking methods, especially work of Hu et al. (2007), can obtain satisfied results in the simple scene based on radiometrical, geometric, topological characteristics of roads. However, roads and parking lot surfaces are made of the same construction materials and thus have similar spectral information. Moreover, they are connected with each other. Therefore, the traditional tracker will extract the parking lots mistakenly. The leakage from roads to the surroundings because of spectral similarity and connection is called leakage problem (Hu et al., 2007). Besides of this situation, the appearance of shadows, vehicle congestions, road material changes, overpasses, etc. stop road-tracking process, which is called blockage problem (Niu, 2006). To resolve these two problems, a RoadModeler consisting of a multidirectional approach and a single-direction approach is proposed. The multidirectional approach which can detect a road network is

mainly used to resolve most of the blockage problems mentioned above. However, its powerful tracking ability can result in leakage problem. Therefore, the single-direction approach which can detect one road according to one optimal direction is needed to avoid the leakage problem.

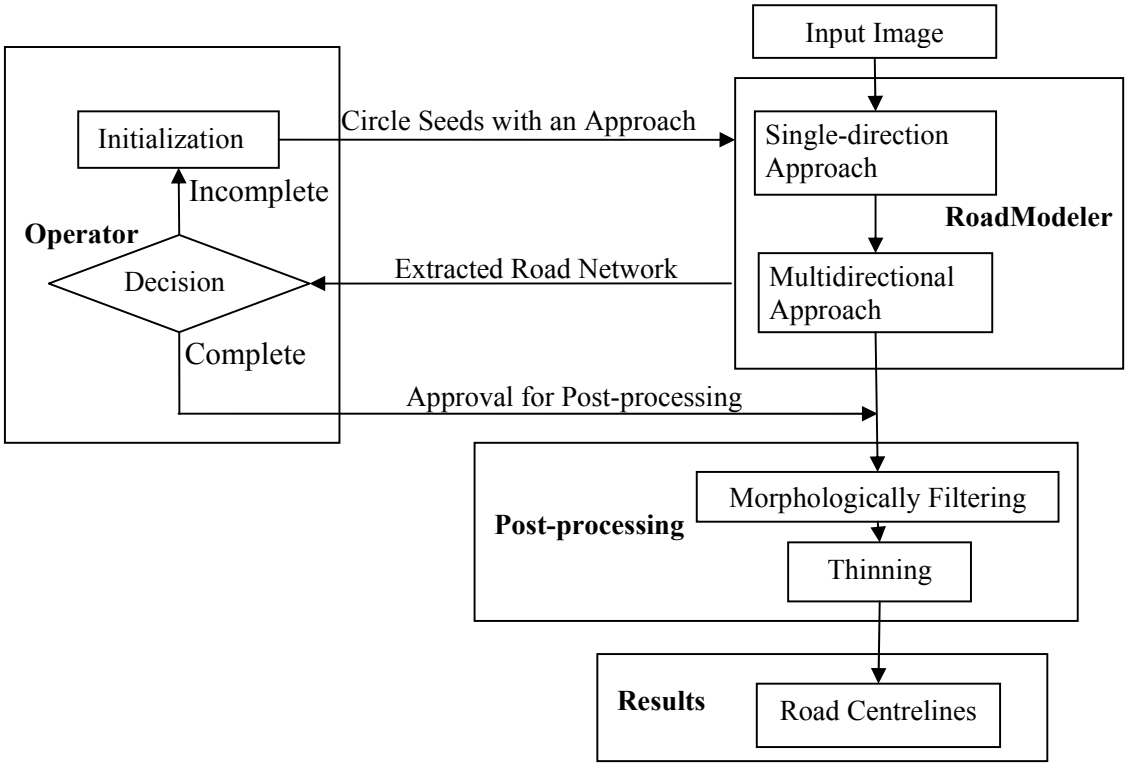


Figure 3.1 Flowchart of proposed semi-automatic road extraction system

Figure 3.2 illustrates the flowchart of the proposed semi-automatic road extraction system. Compared with the traditional one, the one difference is that the initialized seed by a human operator is not a point but a circle, so called initial circle seed, which means that the operator need to provide the centre position and its radius. The other difference is that the applied

approach, which could be the single-direction approach or the multidirectional approach, are also required to be initialized for the circle seed so that it can track roads through the corresponding approach.

For a scene without potential leakage problem, the multidirectional approach is enough to extract all the road networks. The initial circle seed tracks the road based on the selected approach and continue to track a road until no more roads can be detected. At this point, the human operator will check whether all the roads in the image have been detected. If not, the operator needs to place a new circle seed together with the same approach. This process continues until all the roads are extracted.

On the other hand when dealing with a scene with potential leakage problem, the single-direction approach is first applied to extract the roads which link with the area with similar spectral information, such as parking lots. After that, the undetected roads without potential leakage problem can be extracted by using multidirectional approach. Because based on one circle seed, the multidirectional approach which can track a road network is much more efficient than single-direction approach which can only track one road.

When roads are detected, morphological filtering can be applied to remove noises (e.g., vehicles on the road, shadows cast by trees and buildings along the sides of the road). Finally a thinning algorithm is applied to delineate the centreline of the road.

3.2 Grey-Like Colour in RGB Colour Space

Colour images have a great advantage over pan-chromatic or other grey-level imagery as it enhances the capability to discriminate road surface material from most of the other types of landscape materials (Niu, 2006). It is very significant to understand the road's spectral information represented in the RGB colour space.

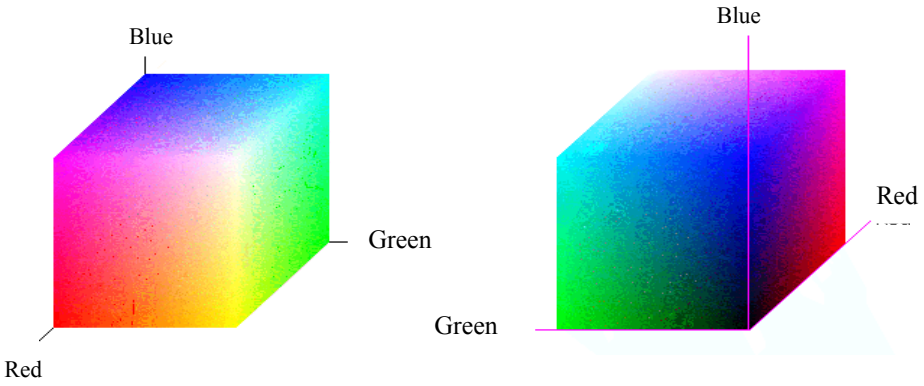


Figure 3.2 RGB colour space.

In the RGB colour space, a colour is represented by the combination of its three primaries (Red, Green, and Blue). Figure 3.1 illustrates the RGB colour model, which uses a rectangular coordinate system with three coordinate axes assigned to each of three primaries to indicate their intensities which start at the origin and increase along each axis (Wyszecki and Stiles, 1967).

Each colour in the RGB colour space can only have a discrete value between zero and the maximum intensity (e.g., 255 for 8-bit length). Thus, the structure of this colour space would

result in a cube. Any colour can be defined by giving its red, green, and blue values or coordinates within the colour cube. A colour image V with n pixels defined in the RGB colour space can be viewed as a vector set,

$$\boldsymbol{V} = \{\boldsymbol{v}_1, \ldots \boldsymbol{v}_j, \ldots, \boldsymbol{v}_n\} \tag{3.1}$$

where $\boldsymbol{v}_j = (v_{jr}, v_{jg}, v_{jb})$ is the colour vector of the jth pixel, and v_{rj}, v_{gj}, v_{bj} are its three colour components, specifically the red, green, and blue, respectively.

The greyscale pixels have the same value in three colour components in the RGB colour space. However, not all the road surface displays grey scales. There are usually some differences between three colour components. Therefore, a grey-like colour which has similar colour components to greyscale can be used to represent the spectral information of road. The grey-like pixels set GP meets the two requirements: (1) Red, green, and blue values of the pixel are close to each other:

$$| \boldsymbol{v}_{jr} - \boldsymbol{v}_{jb} | \leq \boldsymbol{GD}, | \boldsymbol{v}_{jr} - \boldsymbol{v}_{jg} | \leq \boldsymbol{GD}, | \boldsymbol{v}_{jb} - \boldsymbol{v}_{jg} | \leq \boldsymbol{GD}, \text{ for } j = 1, 2, \ldots, n \tag{3.2}$$

(2) All the three values cannot be close to either its minimum (0) or maximum (255) values. This way, black or white can be avoided.

$$\boldsymbol{GMin} \leq \boldsymbol{v}_{jr}, \boldsymbol{v}_{jg}, \boldsymbol{v}_{jb} \leq \boldsymbol{GMax}, \text{ for j} = 1, 2, \ldots, n \tag{3.3}$$

where *GD* (acceptable grey difference), *GMin* (grey minimum), and *GMax* (grey maximum) are selected as 40, 20, 200 respectively based on experiments. The lower *GD* is, the larger proportion the greyscales have in grey-like colour. Presenting grey-like colour can be used as a radiometrical requirement for a road pixel. When the road displays very dark, or the road is covered by shadows, *GMin* can be changed to 0.

3.3 Initialization of Circle Seeds

A circle seed can be initialized by two given points; one is used to indicate the centre point whereas the other is used to indicate the nearest boundary of a road. The radius *R* of the circle seed can be calculated by:

$$\boldsymbol{R} = \sqrt{(\boldsymbol{X_{centre}} - \boldsymbol{X_{boundary}})^2 + (\boldsymbol{Y_{centre}} - \boldsymbol{Y_{boundary}})^2} \tag{3.4}$$

where X_{centre} and Y_{centre} are coordinate of the centre point, while $X_{boundary}$ and $Y_{boundary}$ are coordinate of the boundary point. The initial (or original) circle seed pixels $OSP = \{osp_1, osp_2, ..., osp_l\}$ are the set of the pixels covered by the initial circle seed, l is the number of the pixels. $bsp_i = \{osp_{ir}, osp_{ig}, osp_{ib}, osp_{ix}, osp_{iy}\}$ is the colour and position vector of the ith pixel, and $osp_{ir}, osp_{ig}, osp_{ib}$ are its three colour components, namely the red, the green, and the blue, respectively. osp_{ix} and osp_{iy} are the coordinate of the ith pixel and satisfy:

$$\sqrt{(\boldsymbol{osp_{ix}} - \boldsymbol{X_{centre}})^2 + (\boldsymbol{osp_{iy}} - \boldsymbol{Y_{centre}})^2} <= \boldsymbol{R} \quad (3.5)$$

The Grey initial (or original) circle seed pixels $GOSP = \{gosp_1, gosp_2, \ldots, gosp_m\}$ are the set of the grey-like pixels covered by the initial circle seed, m is the number of the pixels. $gosp_i = \{ gosp_{ir}, gosp_{ig}, gosp_{ib}, gosp_{ix}, gosp_{iy}\}$ is the colour and position vector of the ith pixel, and $gosp_{ir}$, $gosp_{ig}$, $gosp_{ib}$ are its three colour components, namely the red, the green, and the blue, respectively. *GOSP* satisfies:

$$\boldsymbol{GOSP} = \boldsymbol{GP} \cap \boldsymbol{OSP} \quad (3.6)$$

The reference colour ($R_{reference}$, $G_{reference}$, $B_{reference}$) can be calculated by

$$\boldsymbol{R_{reference}} = \frac{\sum_{i=1}^{m} gosp_{ri}}{m} \quad (3.7)$$

$$\boldsymbol{G_{reference}} = \frac{\sum_{i=1}^{m} gosp_{gi}}{m} \quad (3.8)$$

$$\boldsymbol{B_{reference}} = \frac{\sum_{i=1}^{m} gosp_{bi}}{m} \quad (3.9)$$

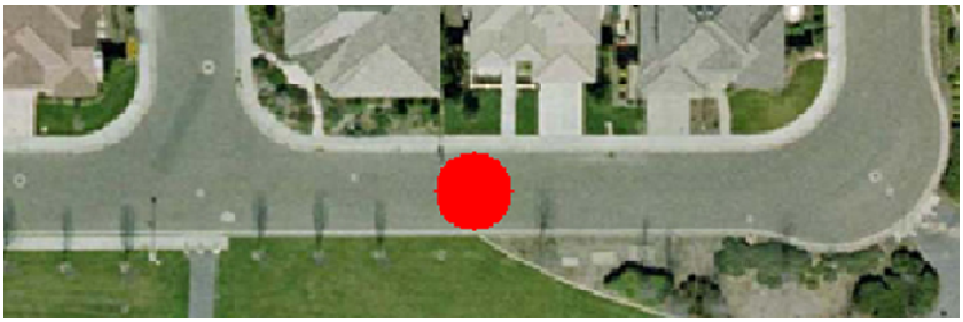


Figure 3.3 A good initial circle seed

The position and the size of the circle seed may significantly influence the further road-tracking process. To explain, a good circle seed would have the circle's centre point positioned in the middle of the road, with its diameter equally matched to the measure of the road's width. Moreover, the area covered by the circle seed should appear as a homogeneous colour. On the contrary, if the circle seed is too big or positioned out of the road's boundary, the covered area would not appear as a homogeneous colour. This becomes a problem because the reference colour generated from that area would not be effective for further road detection processing. The red circle shown in Figure 3.3 is an example of a good circle seed which meets the requirements mentioned above.

3.4 Road Tracking

After the circle seed is placed, the algorithm in of the selected approach will generate adjacent circle seeds. The generated circle seeds can generate other circle seeds. In this thesis, only initial circle seeds are generated by a human operator, while other circle seeds are

automatically generated by the RoadModeler. Therefore, roads can be tracked in the manner of being covered by circle seeds.

Each generated circle seed's information (e.g. coordinate, reference colour) always guide the generation of the sequential circle seeds, thus, it is significant to store every generated circle seed's information. As illustrated in Figure 3.4, the initial circle seed is first added into a circle seeds array (defined in Section 3.4.1) which is used to store the information of all the generated circle seeds, and it would appear later in the processed image. The circle seeds array is enlarged when new circle seeds are generated. Every circle seed in the array will be selected and processed once, and newly generated seeds will be added in this array.

When a circle seed is selected, its peripheral condition (defined in Section 3.4.3), which is the spectral information of the selected circle seed's neighbour area, is calculated. Based on the analysis of this condition, new circle seeds may be generated. The information of the new circle seeds is then added in the existing circle seeds array and the circle seeds are drawn on the processed image. The area covered by newly generated circle seeds is considered as the road area. When all the circle seeds in the circle seeds array have been processed once, and no more circle seeds can be generated to enlarge the array, the road extraction based on this initial circle seed is finished.

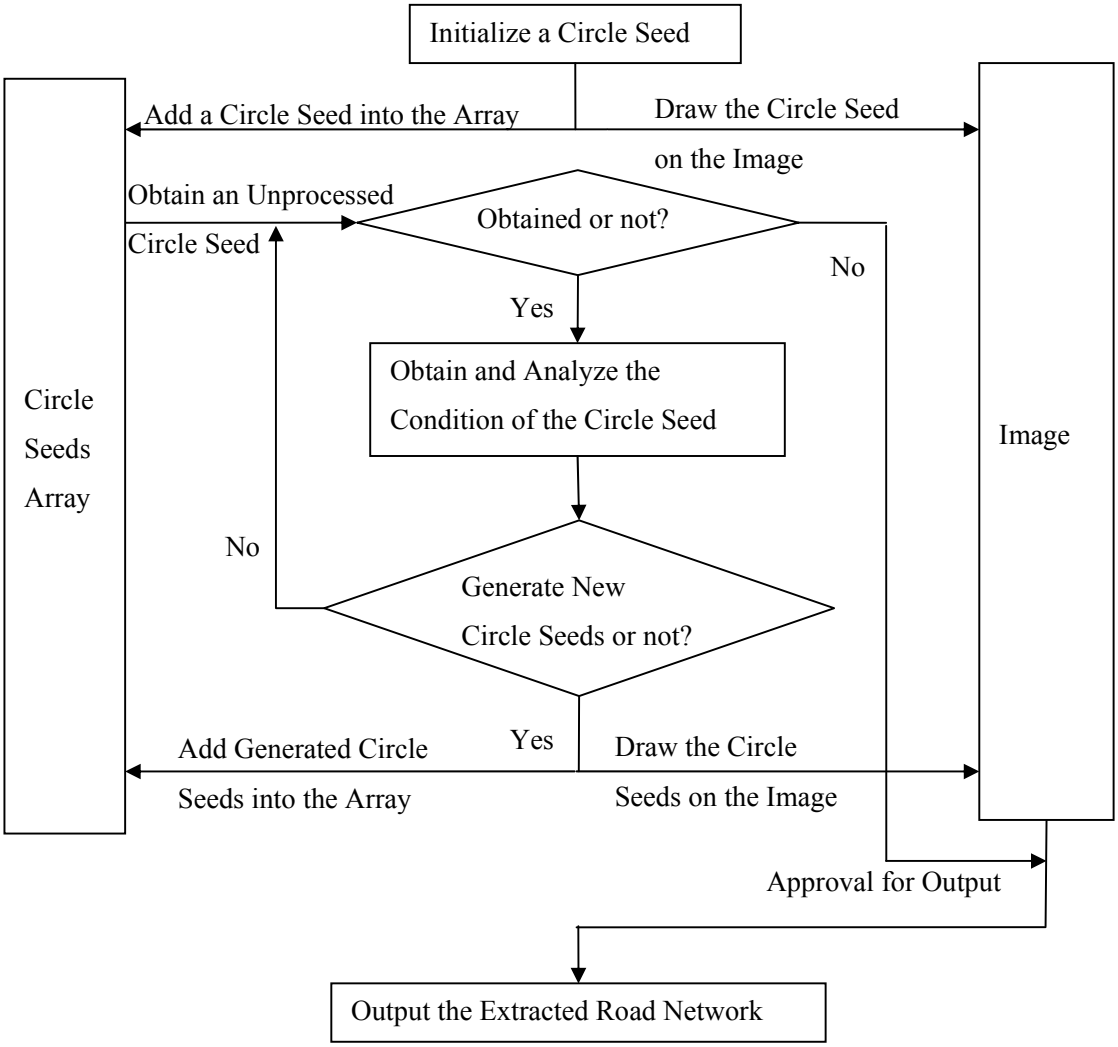


Figure 3.4 Flowchart of road tracking

3.4.1 Initialization of Circle Seeds Array

As mentioned previously, the circle seeds array stores the information of all the circle seeds. This information includes: (1) the identification number used to recognize the circle seed , (2) The X and Y coordinate of the centre point of circle seed used to indicate its position on the

image, (3) the radius of the circle seed used to determine its size, (4) the direction of the circle seed used to determine its origin direction, (5) the detection strategy used for generating the circle seed , (6) the parent circle seed (7) the daughter circle seed generated by this circle seed (8) the reference colour of the circle seed, which is used to determine whether or not the pixels covered by the circle seed belong to road area. The information of each circle seed is significant in the generating and testing of new circle seeds.

Firstly, the circle seeds array will be set by an initial circle seed generated by a human operator. The information of the initial circle seed can be set as the following: The identification number is set to 1; the x and y coordinate and the radius are set according to the original circle seed; The direction is set to -1, which implies the seed is the initial circle seed The reference colour is an ordered triplet (R, G, B) which can be calculated by Equation 3.7-3.9; The detection strategy is set to 0 which indicates that it is being generated by an operator; Parent seed is set to -1, which means this seed does not have a parent seed; The daughter seeds must wait to be set by the newly generated circle seeds.

3.4.2 Two Approaches in RoadModeler

For large complex scenes, using only one type of detection algorithm would not be enough to detect all the roads. For instance, in some areas, there is an inconsistency in colour which leads to confusion with neighbouring areas. Also, some roads are surrounded by trees or

vehicles. There are five detection algorithms composed of two approaches used to generate a circle seed. They are in detail as follows:

(1) General similarity detection algorithm: This is used to detect ribbon road with typical intersections such as cross road, T-junction and Y-junction which have clear boundaries and less noises.

(2) Narrow similarity detection algorithm: This is used to detect roads which are partly covered by vehicles or shadows cast by buildings or trees.

(3) Grey similarity detection algorithm: This is used to detect roads with sudden radiometrical changes due to construction material and hidden by the trees or covered by shadows.

(4) Jump similarity detection algorithm: This is used to detect roads completely covered by vehicles or shadows cast by buildings or trees.

(5) Single-direction detection algorithm: This is used to detect roads in one direction.

The two approaches are as follows:

(1) Single-direction approach: This approach is integrated by operating single-direction detection, narrow similarity detection, grey similarity detection, and jump similarity detection.

(2) Multidirectional approach: This is formed by replacing single-direction detection by general similarity detection.

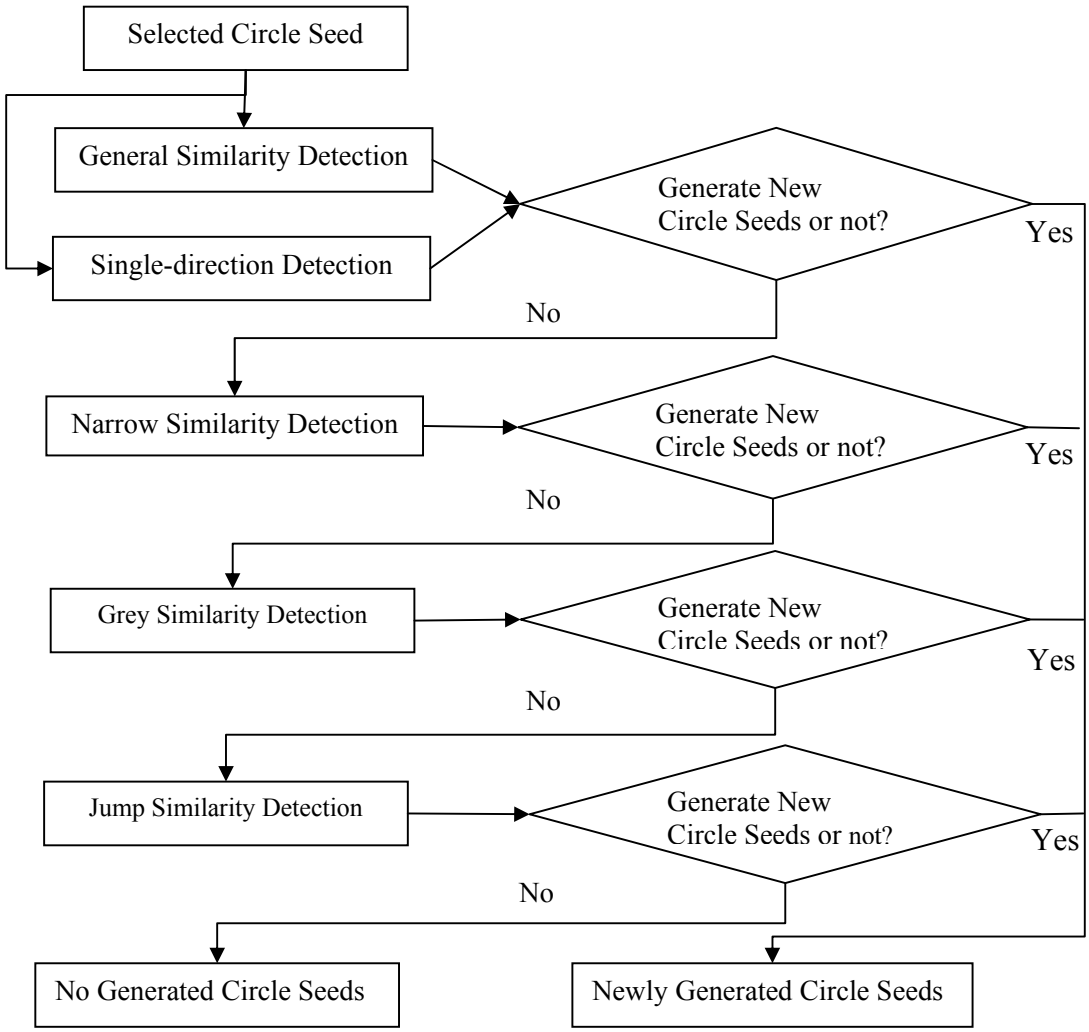


Figure 3.5 Flowchart of different detections

General similarity detection and single-direction detection have the highest priority and are operated first in each approach since the requirements for a circle seed in those detections are strict. On the contrary, the other three detections have relatively low priorities because of the tolerant requirements for the circle seed. As shown in Figure 3.5. When the detection with higher priority is unable to generate new circle seeds, detections with lower priorities will

continue to do its job until all detections in one approach are used. When a new circle seed is generated by the detection with higher priority, the sequent detections will not be processed.

From the detection algorithms which multi-directional approach consists of, this approach can extract various roads with different materials, different widths, different shapes of intersection, and different degrees of noise. However, it usually results in leakage problem. To solve this problem, the single-direction approach could be used in detecting roads according one direction because of the single-direction detection. Thus, when a circle seed passes by a parking lot, it would only extend in its original direction without extracting parking lots. The limitation of the single-direction model is that a circle seed can only move down a road in one direction.

In addition, detections with relatively low priority such as narrow similarity detection algorithm, grey similarity detection algorithm, and jump similarity detection algorithm enables the two approaches to detect more road areas than that detected only by general similarity detection and single-direction detection algorithms. However, this detection algorithm is also more prone to mistakes since non road areas are frequently detected. Thus, the validation should be applied on the circle seed generated by the detections with low priority. If no circle seed can be generated within the neighbour area, only jump similarity detection can detect the area far from the selected seed to obtain more information for analyzing. For single-direction detection, only one circle seed needs to be generated based on the road direction and spectral information, and circle seed validation is not necessary.

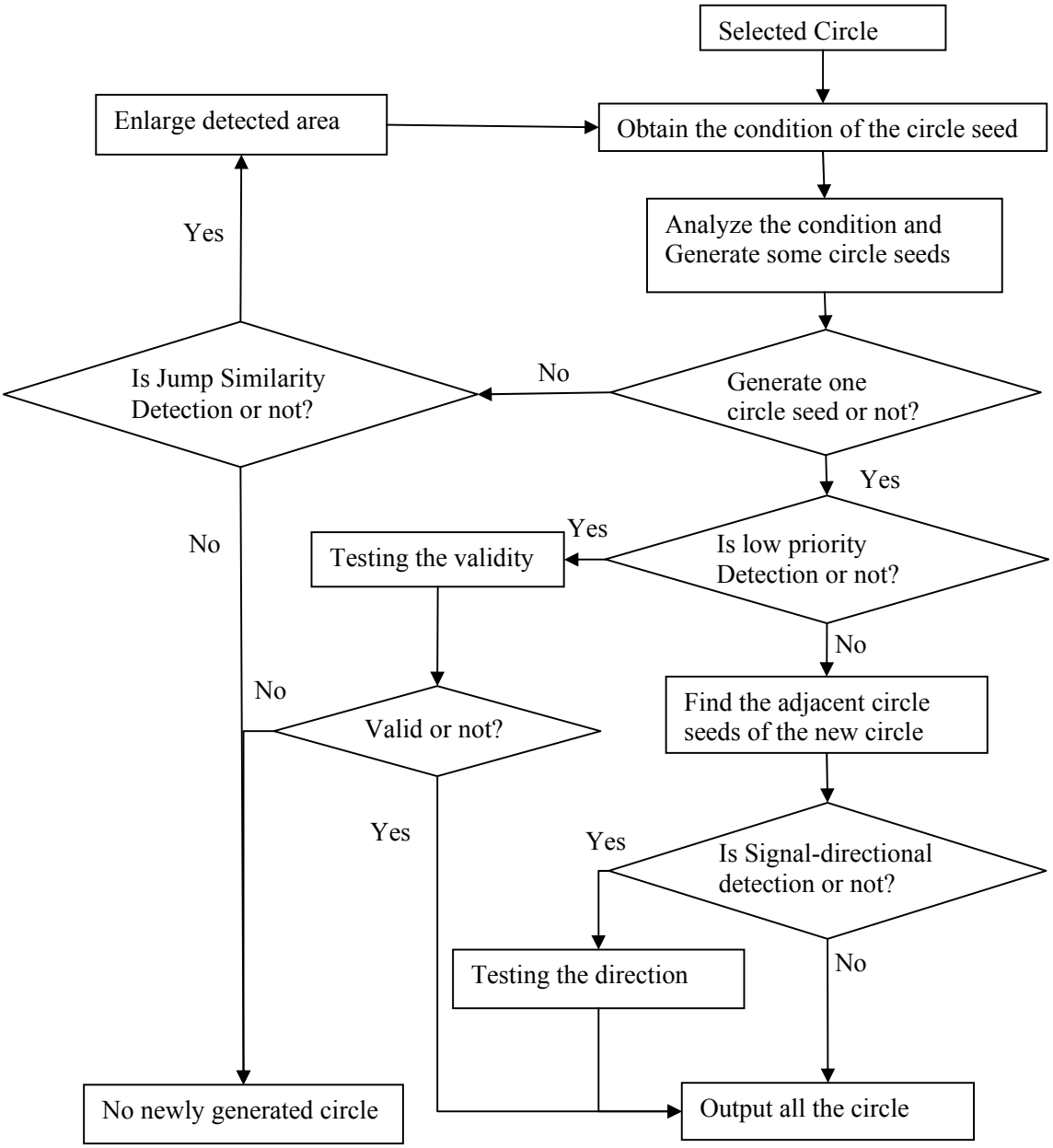


Figure 3.6 Flowchart of the road detection strategy

3.4.3 Peripheral Condition of the Circle Seed

Candidate circle seeds

As mentioned before, each circle seed has its own information in the circle seeds array. When a circle seed is selected as a parent circle seed (also called selected circle seed or processed circle seed) to generate new daughter circle seeds, the parent circle seed's information which consists of x and y coordination: X_{parent} and Y_{parent}, radius R, direction D_{parent}, and reference colour (RC_{parent_r}, RC_{parent_g}, RC_{parent_b}). Some of them are firstly used to generate the candidate circle seeds which are the peripheral circle seeds with the same radius surrounding the parent circle seed. The centre of a candidate circle seed can be calculated by:

$$\boldsymbol{X_{candidate\_k}} = \boldsymbol{X_{parent}} + \boldsymbol{MD} * \cos(\boldsymbol{\theta_k}) \tag{3.10}$$

$$\boldsymbol{Y_{candidate\_k}} = \boldsymbol{Y_{parent}} + \boldsymbol{MD} * \sin(\boldsymbol{\theta_k}) \tag{3.11}$$

where MD is the moving distance which is between the centres of the parent circle seed and a candidate circle seed. θ_k is the direction represented as a horizontal angle measured anticlockwise from a south base line, which cannot exceed the highest number of units in a circle (a 360° circle). So the direction θ_k satisfies:

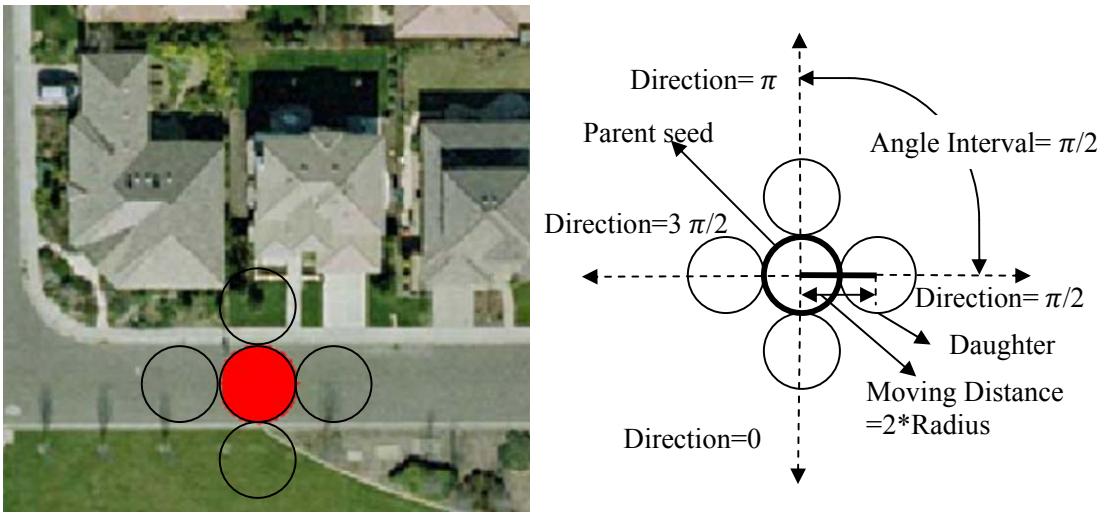
$$\boldsymbol{\theta_k} = \boldsymbol{k} * \boldsymbol{AI}, \text{ for } \boldsymbol{k} = 1, 2, \ldots, \boldsymbol{N_{direction}} \tag{3.12}$$

$$N_{direction} = \frac{360^{\circ}}{AI} \tag{3.13}$$

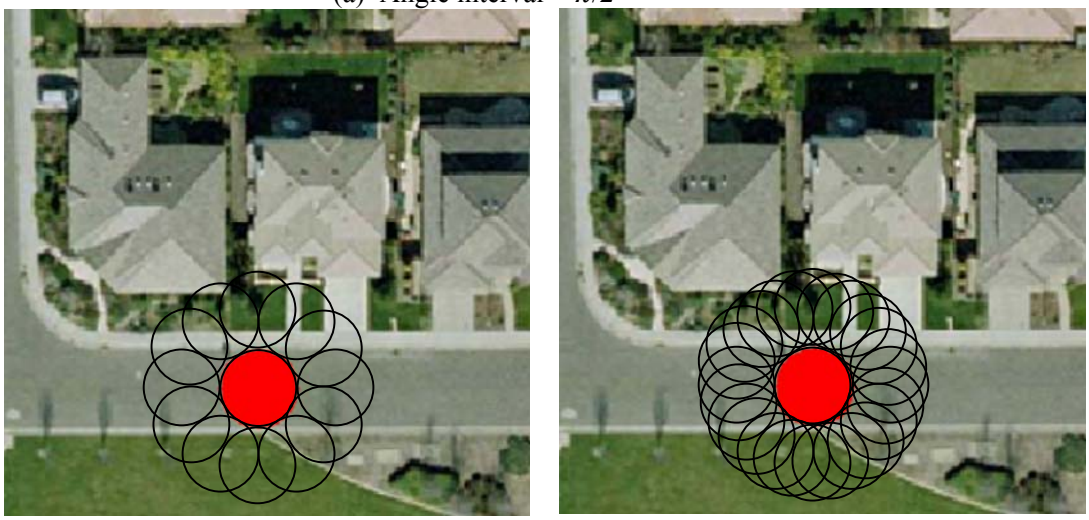
where *AI* is the angle between two adjacent directions, known as the angle interval. The lower the angle interval is the more accurate the direction of a daughter seed. However, more calculating time is also spent. Based on experiments, the angle interval is selected as $\pi/12$, and the moving distance is set as the diameter of the circle (see Figure 3.7 c). For the jump similarity detection, this distance can be enlarged to detect more areas around a selected circle seed.

Candidate circle seed pixels $BSP = \{bsp_1, bsp_2, ..., bsp_m\}$ are the set of the pixels covered by the circle seed, m is the number of the pixels. $bsp_i = \{bsp_{ir}, bsp_{ig}, bsp_{ib}, bsp_{ix}, bsp_{iy}\}$ is the colour and position vector of the *i*th pixel, and bsp_{ir}, bsp_{ig}, bsp_{ib} are its three colour components, namely the red, the green, and the blue, respectively. bsp_{ix} and bsp_{iy} are the coordinate of the pixel and satisfy:

$$\sqrt{(\boldsymbol{bsp_{ix}} - \boldsymbol{X_{candidate\_k}})^2 + \left(\boldsymbol{bsp_{iy}} - \boldsymbol{Y_{candidate\_k}}\right)^2} <= R \tag{3.14}$$



(a) Angle interval = $\pi/2$



(b) Angle interval = $\pi/6$ (c) Angle interval = $\pi/12$

Figure 3.7 Position of the candidate seed circles with different angle intervals

Information of Candidate circle seeds

After the candidate circle seeds have been determined their condition information can be calculated. This includes the direction of the peripheral candidate circle seed, the proportion

of the similar grey pixels, the proportion of the grey pixels and the proportion of the classified road pixels. The proportion of the similar grey pixels (PSGP) is defined as the ratio between the number of similar grey pixels (SGPB) whose colour is similar to the selected circle seed's reference colour and the total number of circle seed pixels (BSP) that is covered by a candidate circle seed:

$$\boldsymbol{PSGP = \frac{N(SGPB)}{N(BSP)} = \frac{l}{m}} \tag{3.15}$$

where $SGPB = \{sgpb_1, sgpb_2, \ldots, sgpb_l\}$ is intersection of *BSP* and *GP*. l is the number of the pixels. $sgpb_i = \{ sgpb_{ir}, sgpb_{ig}, sgpb_{ib}, sgpb_{ix}, sgpb_{iy} \}$ is the colour and position vector of the ith pixel, and $sgpb_{ir}$, $sgpb_{ig}$, $sgpb_{ib}$ which are its three colour components, namely the red, the green, and the blue respectively. Besides the requirements mentioned in *BSP* and *GP*, those three colour components satisfy:

$$\boldsymbol{| sgpb_{ir} - RC_{parent\_r}| \leq CDiff} \tag{3.16}$$

$$\boldsymbol{| sgpb_{ig} - RC_{parent\_g}| \leq CDiff} \tag{3.17}$$

$$\boldsymbol{| sgpb_{ib} - RC_{parent\_b}| \leq CDiff} \tag{3.18}$$

where CDiff (Colour Difference) is represented as the acceptable difference between the reference colour and the colour of one specific pixel. The similarity of two circle seeds is

evaluated by a threshold which can be set as different values: MaxCDiff, MedCDiff, and MinCDiff which is used in general similarity detection, narrow similarity detection, and jump similarity detection, respectively.

The proportion of the grey pixels (PGP) is defined as the ratio between the number of grey pixels in the candidate circle seed and the total number of pixels covered by the candidate circle seed.

$$\boldsymbol{PGP} = \frac{\boldsymbol{N(GPB)}}{\boldsymbol{N(BS)}} \tag{3.19}$$

where GPB are the pixels whose position vector satisfies the requirements of BS and whereas whose colour vector satisfies the requirements of GP.

The proportion of the classified road pixels (PRP) is defined as the ratio between the number of detected road pixels in the candidate circle seed and the total number of pixels covered by the candidate circle seed.

$$\boldsymbol{PRP} = \frac{\boldsymbol{N(RPB)}}{\boldsymbol{N(BS)}} \tag{3.20}$$

where *RPB* are the pixels whose position vector satisfies the requirements of *BS* and who has detected as road pixels

3.4.4 Analysis of the Condition of Candidate Circle Seeds

From the information provided by peripheral candidate circles, a new circle seed could be the candidate circle seeds which have the largest grey proportion value in the continuous peripheral candidate circle seeds. The continuous peripheral candidate circle seeds should satisfy four requirements:

Firstly, PSGP of the candidate circle seeds should exceed the minimum value which is referred as Stand Similar Grey Proportion which is composed of three values: Maximum Stand Similar Grey Proportion (MaxSSGP), Median Stand Similar Grey Proportion (MedSSGP), and Minimum Stand Similar Grey Proportion (MinSSGP) which can be used in general similarity detection algorithm (or single-direction detection algorithm), narrow similarity detection algorithm, and jump similarity detection algorithm respectively. For the grey similarity detection algorithm, PGP of the candidate circle seeds should exceed a minimum value which is referred as Stand Grey Proportion (SGP). All the four thresholds are selected based on experiments. These four thresholds are used to validate a circle seed's spectral information.

Secondly, the difference between PSGP and PRP of all the candidate circle seeds is greater than a minimum value which is referred as Additional Grey Proportion (AGP), which is used in general similarity detection algorithm or single-direction detection algorithm. For other three detection algorithms, PRP of them is under a maximum value which is referred as

Stand Maximum Road Proportion (SMRP). These two parameters is used to increase the effectiveness of each circle seed which means it can reduce superposition of two circle seeds.

Thirdly, the candidate circle seed's direction $\boldsymbol{\theta_k}$ satisfies:

$$\boldsymbol{Min}\ (\ |\ \boldsymbol{\theta_k} - \boldsymbol{D_{parent}}\ |,\ 2\pi - |\ \boldsymbol{\theta_k} - \boldsymbol{D_{parent}}\ |\ )\ \leq \frac{\boldsymbol{ERA}}{\mathbf{2}} \tag{3.21}$$

where D_{parent} is the parent circle seed's direction which is obtained from the circle seeds array, *ERA* is effective range angle and is used to limit the direction of the candidate circle seed. ERA is set as two thresholds which is referred as wide effective range angle (WERA) and narrow effective range angle (NERA). As illustrated in Figure 3.8, the parent circle seed A's direction is $\pi/2$ and ERA is $4\pi/3$, so the directions of the candidate circle seeds B, C, D, E, and F in shadow area are effective directions.

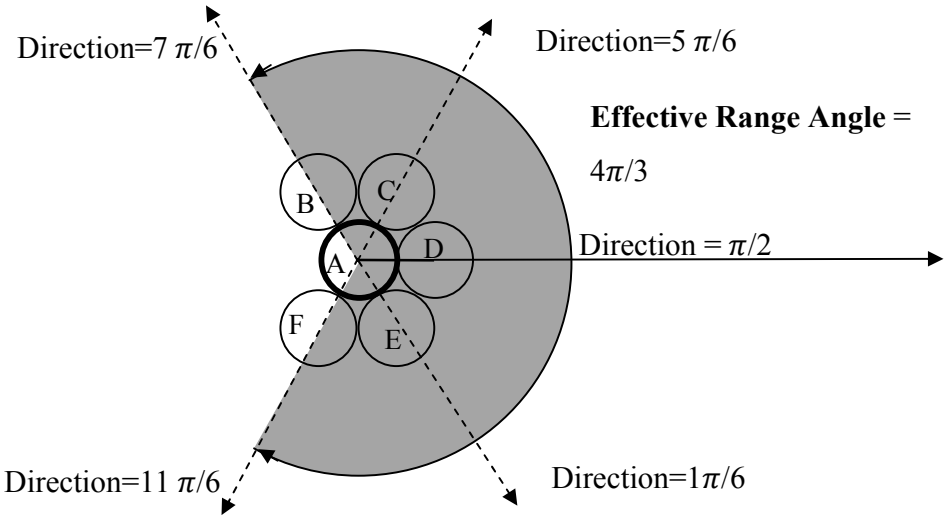


Figure 3.8 Effective range angle

Finally, the peripheral candidate circle seeds are in a circle. This means the first direction $\pi/12$ and the last direction $2*\pi$ can also be considered as continuous circle seeds. The continuous circle seed's directions θ_m and θ_n would satisfy:

$$\boldsymbol{Min}\ (\ |\ \boldsymbol{\theta_m} - \boldsymbol{\theta_n}\ |,\ \ 2\pi - |\ \boldsymbol{\theta_m} - \boldsymbol{\theta_n}\ |\ )\ \leq\ \boldsymbol{AI} \tag{3.22}$$

where *AI* is the angle interval. After the new circle seed have been detected, its coordinate, radius, direction, its parent identification number, and the used detection strategy will be recorded for updating of the circle seeds array.

3.4.5 Validation of Adjacent Circle Seeds

After the new daughter circle seeds in various directions are generated, considering that the road may become wider, the adjacent area of the daughter seed will be checked to see whether it satisfies the same requirements of circle seed generated by the general similarity detection algorithm. The neighbour circle seeds are also the candidate circle seeds whose direction is 60º departures from the direction of the newly generated circle seed. The adjacent circle seed will be used to extend to more neighbouring circle seeds until no more circle seed can satisfy the requirements.

As shown in Figure 3.9, circle seed B is generated from circle seed A, candidate circle seeds C and D are the adjacent circle seeds of B, the interval angle between the circle seed and its

neighbour circle seeds is 60º. If C or D satisfies the requirements, it will be considered as the new daughter circle seed whose adjacent circle seed (E or F) also need to be checked to generate more new daughter circle seeds. This step is used in general similarity detection algorithm to enlarge the width of the road and in single-direction detection algorithm to find the best direction.

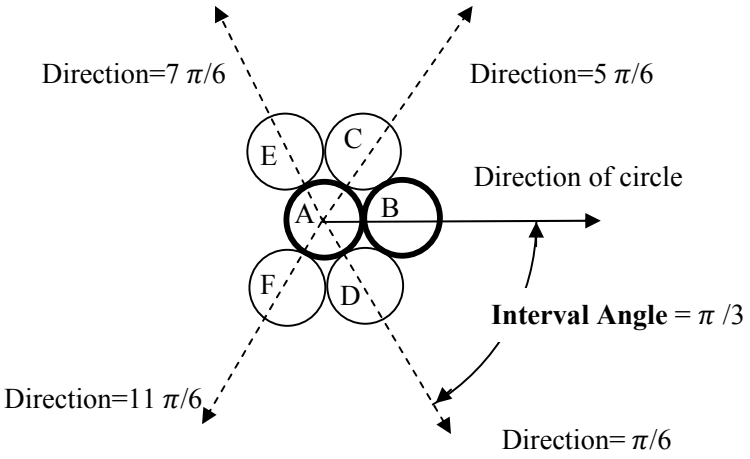


Figure 3.9 Neighbour circle seeds

3.4.6 Requirements for Different Detection Algorithms

As shown in Table 3.1, the standard thresholds of the different detection algorithms in the multidirectional approach are based on experiments. Grey Maximum, Grey Minimum, Grey Difference are the same. Other thresholds will be explained in detail in the following part.

Table 3.1 The Thresholds of different detection algorithms in Multidirectional Approach

	General Similarity Detection algorithm	Narrow Similarity Detection algorithm	Grey Similarity Detection algorithm	Jump Similarity Detection algorithm
Stand Similar Grey Proportion	MaxSSGP =70%	MedSSGP =30%	N/A	MinSSGP =25%
Additional Grey Proportion	AGP=70%	N/A	N/A	N/A
Stand Grey Proportion	N/A	N/A	SGP =90%	N/A
Stand Road Proportion	N/A	SMRP =1%	SMRP =1%	SMRP =1%
Colour Difference Criteria	MaxCDiff =25	MedCDiff =15	N/A	MinCDiff =15
Effective Range Angle	WERA=3 π/4	NERA= π/6	NERA= π/6	NERA= π/6
Grey Maximum	200	200	200	200
Grey Minimum	20	20	20	20
Grey Difference	40	40	40	40
Neighbour Detection or not	Yes	No	No	No
Number of generated circle seeds	No limit	1	1	1
Moving Distance	2*radius	2*radius	2*radius	2*radius
Max Moving Distance	N/A	N/A	N/A	6*radius
Distance Increment	N/A	N/A	N/A	2

General Similarity Detection algorithm

The general similarity detection algorithm is used frequently. It can be used to detect visible roads with different degrees of width and different types of intersections. This can be done as

long as spectral values along a road do not change greatly within a short distance, and few vehicles or shadows are evident in covering the road.

As shown in Table 3.1, the standard similar grey proportion and additional grey proportion are usually set as 70% (MaxSSGP and AGP), and the colour difference criteria is set as 25 (MaxCDiff), which is larger than other detection algorithms. Thus, the detected circle seeds would have large proportion of similar colour pixels to his parent circle seed. The effective angle, which is set as 1.5π (WERA), is the largest one in all the detection algorithms to detect all the possible directions so that different kinds of intersections can be detected. After the circle seeds which have the largest proportion of grey similar pixels and have different directions have been detected, the adjacent areas of them also need to be checked so that road whose width is wider than the diameter of the circle seed could be detected.

Narrow Similarity Detection algorithm

When general similar detection algorithm fails to generate new circle seeds, the narrow similar detection algorithm can be used to detect a circle seed, specifically for roads partly covered by vehicles or shadows cast by trees and buildings.

As shown in Table 3.1, the stand similar grey proportion is usually set as 30% (MedSSGP), which is lower than general similar detection algorithm. Therefore, the unclear road with low proportion of grey pixels can also be detected. Stand road proportion is set as 1% (SMRP**)** to

ensure the effectiveness of the newly generated circle seed. In order to reduce mistaken detection, the colour difference criteria is set as 15 (MedCDiff), which is much lower than general similar detection algorithm. Therefore, the circle seed with higher similar colour to its parent circle seed's reference colour can be generate through narrow similar detection algorithm than general similarity detection algorithm. Additionally, the effective range angle is set to $\pi/6$ (NERA), which is much lower than general similar detection algorithm, so that the direction of the possible newly generated circle seed keep consistent with its parent circle seed's direction. Therefore, the circle seed which has low proportion of pixels with higher spectral similarity and have a consistent direction could be detected. The neighbouring circle seeds do not need to be checked in this detection algorithm.

Grey Similarity Detection algorithm

Due to different materials or constructing time, the spectral information changes greatly along the road. No circle seeds can be generated through general similarity detection algorithm and narrow similarity detection algorithm because of the few similar pixels between the candidate circle seed and the selected circle seed. In this case, the grey similarity detection algorithm should be applied.

As shown in Table 3.1, the stand grey proportion is normally set to 90% (SGP). The range of the definition of the grey-like pixel in this thesis is quite large; even some non-grey pixel is included in this range. In order to reduce mistaken detection, the effective range angle and

stand road proportion is set to the same value as narrow similarity detection algorithm's to guarantee the generated circle seed's consistency in direction and effectiveness. The circle seed which has a large proportion of grey-like pixels and have consistent direction with its parent's can be detected. The checking of neighbouring circle seeds is not required in this detection algorithm.

Jump Similarity Detection algorithm

The appearances of vehicles, pedestrian lines, overpasses and shadows cast by trees and buildings can produce a negative influence on homogeneous intensity and connectivity of roads. In some cases, shadows may cover certain sections of a road. No circle seed can be generated through the detections mentioned before. This is because the near peripheral candidate circle seeds cannot satisfy the requirements of those three detections. However, there may be a circle seed satisfying the requirements of the jump similar detection algorithm in the area far from the selected circle seed.

As shown in Table 3.1, the parameters in narrow similar detection and jump similar detection are almost the same. The difference is that the stand similar grey proportion is usually set to 25% (MinSSGP), which is lower than that of the narrow similarity detection. The moving distance is dynamically changed to detect farther areas.

The moving distance is increased gradually by a specific value called Distance Increment. If the circle seed cannot be generated within the area of the increased moving distance, the moving distance will keep increasing until a threshold value called Max Moving Distance is obtained.

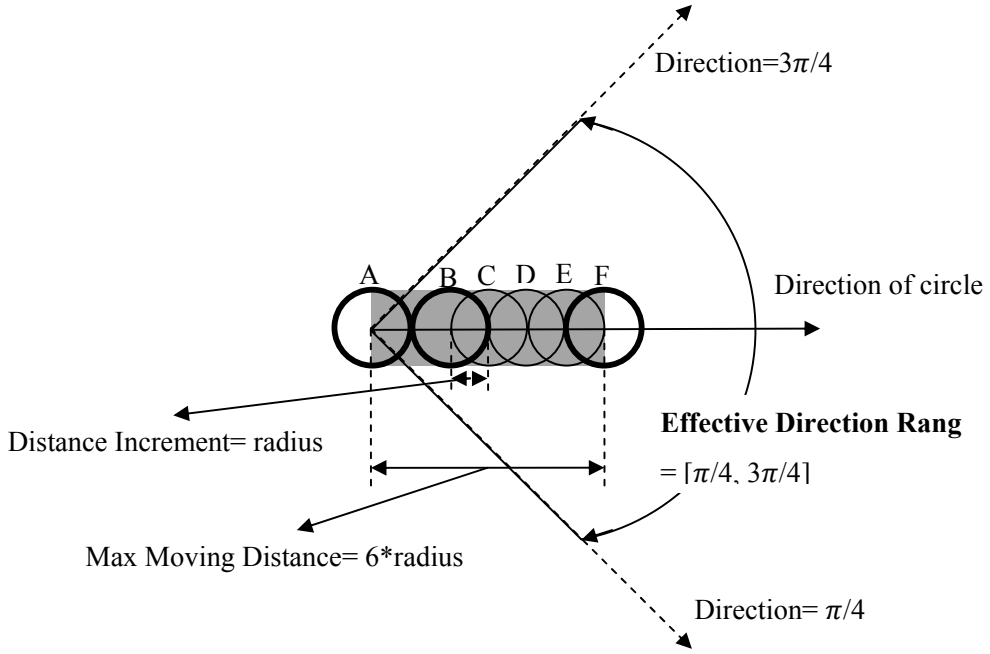


Figure 3.10 Jump similar detection

In Figure 3.10, assuming the distance increment and the Max Moving Distance are set as radius and 6*radius respectively. The circle seed A cannot generate any circle seeds by using the moving distance which is the diameter of the circle seed. There are still four more trials which can be made to obtain a circle seed. Since the distance increment is set to the length of 2 pixels, the circle seed can be detected as near as possible before the moving distance reaches its threshold value. After the circle seed which has low proportion of pixels with higher spectral similarity and have consistent direction has been detected, the area between

the daughter circle seed and the parent circle seed is considered as road area as well. As illustrated in Figure 3.10, if the candidate circle seed F satisfies all the requirements, the area between circle seed A and circle seed F which is the grey area will be detected as road area. The adjacent circle seeds do not need to be checked for detecting wider roads.

Single-direction Detection algorithm

Single-direction detection algorithm can replace the general similarity detection algorithm to generate a new approach called single-direction approach. When a parking lot is located beside a road, the general similar detection will consider the parking lot as a wider road or intersection mistakenly. Single-direction detection can resolve this problem effectively.

As shown in Table 3.2, the stand similar grey proportion and colour difference criteria are normally set to 60% and 35% respectively, which is more tolerant than general similar detections. This way, the detected circle seed would have large proportion of similar colour pixels to his parent circle seed. The requirement for additional grey proportion is not required, because the road tracking does not need to be blocked by the detected road. Additionally, the effective range angle is set to $\pi/6$, which is much lower than general similar detection algorithm. Therefore, the effective direction range is narrow, and the direction of the possible newly generated circle seed is consistent with its parent circle seed's direction. After the circle seed which has a high proportion of pixel with high spectral similarity and has consistent direction has been detected, its neighbouring circle seeds needs to be checked so

that more circle seeds can be generated to provide more directional choices, Only the circle seed with the most consistent direction will be selected and shown on the map . One parent circle seed can only generate one daughter circle seed through this type of detection.

Table 3.2 The Thresholds of different detection algorithms in Single-direction Approach

	Single-Directional Detection algorithm	Narrow Similarity Detection algorithm	Grey Similarity Detection algorithm	Jump Similarity Detection algorithm
Stand Similar Grey Proportion	MaxSSGP =60%	MedSSGP =21%	N/A	MinSSGP =15%
Additional Grey Proportion	N/A	N/A	N/A	N/A
Stand Grey Proportion	N/A	N/A	SGP =60%	N/A
Stand Road Proportion	N/A	SMRP =1%	SMRP =1%	SMRP =1%
Colour Difference Criteria	MaxCDiff =35	MedCDiff =15	N/A	MinCDiff =15
Effective Range Angle	WERA= $\pi/6$	NERA= $\pi/6$	NERA= $\pi/6$	NERA= $\pi/6$
Grey Maximum	200	200	200	200
Grey Minimum	20	20	20	20
Grey Difference	40	40	40	40
Neighbour Detection or not	No	No	No	No
Number of generated circle seeds	1	1	1	1
Moving Distance	2*radius	2*radius	2*radius	2*radius
Max Moving Distance	N/A	N/A	N/A	6*radius
Distance Increment	N/A	N/A	N/A	2

In single-direction approach, when no more circle seeds can be generated, narrow similar detection algorithm, grey similarity detection algorithm, and jump similarity detection algorithm will be used to detect circle seeds in order, only one circle seed with the most consistent direction will be shown on the image and be used to generate the next circle seed. The thresholds for these three detection algorithms in a single-direction approach are more tolerant than that in a multidirectional approach because the direction consistency can guarantee the correct road extraction.

3.4.7 Updating the Circle Seeds Array

After the newly circle seeds have been generated, the information of it will be stored in the circle seed array. The information of the new circle seed can be set as the following : The identification number is set by adding the previous circle seed's identification number by 1; the x and y coordinate and the radius are set according to the new circle seed; The direction is set to corresponding direction from its parent circle seed; The reference colour can be calculated by equation (3.7)-(3.9); The detection strategy is set to corresponding approach and detection algorithm; Parent seed is set to the identification number of its parent circle seed; The daughter seeds must wait to be set by the newly generated circle seeds. This circle seed will record its identification number to the daughter seed field of its parent circle seed.

3.4.8 Drawing of a Circle Seed

After the newly circle seeds have been determined, the grey pixels covered by them will be changed to red colour and shown in the image (Figure 3.11a). To connect the roads together, the area between the parent seed and the daughter seed need to be covered by a circle named Connected Circle (Figure 3.11b). The grey-like pixels covered by the connected circle can be considered as road area.

(a) Drawing of one generated seed

(b) Making two seeds more continual

(c) Drawing of another generated seed

(d) Making two seeds more continual

Figure 3.11 Drawing a seed circle

The connected circle's coordinate ($\boldsymbol{X_{connected}}$, $\boldsymbol{Y_{connected}}$) can be calculated by:

$$\boldsymbol{X_{connected}} = \frac{\boldsymbol{X_{father}} + \boldsymbol{X_{son}}}{\mathbf{2}} \tag{3.23}$$

$$\boldsymbol{Y_{connected}} = \frac{\boldsymbol{Y_{father}} + \boldsymbol{Y_{son}}}{\mathbf{2}} \tag{3.24}$$

where (X_{parent}, Y_{parent}) and ($X_{daughter}$, $Y_{daughter}$) are a parent circle seed and its daughter circle seed' coordinates respectively.

3.4.9 Validation

Not all the road areas detected by the detection algorithms mentioned above are correct. Different detections have different probabilities of detecting road areas correctly. From the descriptions of the detections, general similar detection and single-direction detection have higher probability of detecting road areas correctly, and the other three detections have relatively lower probability. Testing the validity of the circle seeds generated by those three detections helps to improve the effectiveness of the detection algorithm so that the number of the mistakenly generated circle seeds can be avoided.

The continuous usage of low priority detection algorithms to generate circle seeds has very high probability of detecting non-road area mistakenly. When both a circle seed's detection strategy and its father circle seed's detection strategy are one of those three detections, the

direction of the two circle seeds will be tested to find out whether or not they are consistent. The degree of the consistence of directions can be evaluated by the angle difference between the two directions. When the angle difference is less than the Max Angle Difference, two circles' directions are viewed consistent, and the two circle seeds are taken as valid circle seeds. If they are not, the two circle seeds will be considered as useless circle seeds and are removed from the images. Moreover, their information in circle seeds array are removed as well. Max Angle Difference is selected as $\pi/12$ based on experiments. The information of the valid circle seeds will be appended into in the circle seeds array, waited to be processed.

3.5 Morphological Filtering

Matheron (1975) and Serra (1982) developed Mathematical morphology as a geometry-based technique for image processing and analysis. Binary morphology uses the set operators such as union, intersection, complementation, and translation, while a binary image is represented as a set.

Let E be the discrete integer space Z^m, a binary image can be represented as a subset of E or a mapping f: E $\rightarrow$ {0, 1}. The translation operator, T_a, which translates the set of object vectors, $\boldsymbol{OV} \subset \boldsymbol{E}$, by a constant vector, $\boldsymbol{a} \in \boldsymbol{E}$, is defined as

$$T_a(\boldsymbol{OV}) = \{\boldsymbol{ov} + \boldsymbol{a} : \boldsymbol{ov} \in \boldsymbol{OV}\} \tag{3.25}$$

The reflection or symmetry operator, –, of OV is defined as

$$-\boldsymbol{OV} = \{-\boldsymbol{ov} : \boldsymbol{ov} \in \boldsymbol{OV}\} \tag{3.26}$$

Let $\boldsymbol{BI} \subset \boldsymbol{E}$ be a binary image and $\boldsymbol{SE} \subset \boldsymbol{E}$ be a binary structuring element. The binary dilation δ_{SE}^{T}, erosion ε_{SE}^{T}, closing χ_{SE}^{T}, and opening o_{SE}^{T} of BI by SE are binary images given by:

$$\delta_{SE}^{T}(\boldsymbol{BI}) = \{\boldsymbol{a} \in \boldsymbol{E} : T_a(-\boldsymbol{SE}) \cap \boldsymbol{BI} \neq \Phi\} \tag{3.27}$$

$$\varepsilon_{SE}^{T}(\boldsymbol{BI}) = \{\boldsymbol{a} \in \boldsymbol{E} : T_a(\boldsymbol{SE}) \subseteq \boldsymbol{BI}\} \tag{3.28}$$

$$\chi_{SE}^{T}(\boldsymbol{BI}) = \varepsilon_{SE}^{T}(\delta_{SE}^{T}(\boldsymbol{BI})) \tag{3.29}$$

$$o_{SE}^{T}(\boldsymbol{BI}) = \delta_{SE}^{T}(\varepsilon_{SE}^{T}(\boldsymbol{BI})) \tag{3.30}$$

Generally, the binary opening suppresses small peaks and eliminates other small details, while the binary closing fills up narrow channels and small gaps. It is notable that the size of larger structures or objects isn't really affected by binary closing and binary opening. Given the fact that noise (objects which does not belong to road) exists in the extracted images. Depending on the shape of noise objects, the appropriate combinations of binary closing or

opening operators can be used to remove the noise (e.g., small area out of the road area) or fill the holes within the road area.

3.6 Road Centreline Delineation

Zhang and Suen (1984) presented a thinning algorithm to extract a skeleton of an object from a binary image. This algorithm can also be used to extract road centerline from the morphological result, where road pixels have a value of 1 (white), and those background (non-road) pixels have a value of 0 (black). This method consists of two iterative steps, which are used to delete boundary pixels. The first step deletes south-east boundary pixels p, if its 8-neighbour pixels shown in Figure 3.12 are satisfied:

$$2 \leq N(p) \leq 6 \tag{3.31}$$

$$S(p) = 1 \tag{3.32}$$

$$p_0 \cdot p_2 \cdot p_4 = 0 \tag{3.33}$$

$$p_2 \cdot p_4 \cdot p_6 = 0 \tag{3.34}$$

where $N(p)$ is the number of neighbour road pixels:

$$N(p) = \sum_{i=0}^{7} p_i \tag{3.35}$$

and $S(p)$ is the number of 01 patterns in the ordered sequence of $p_0, p_1, \ldots, p_6, p_7$.

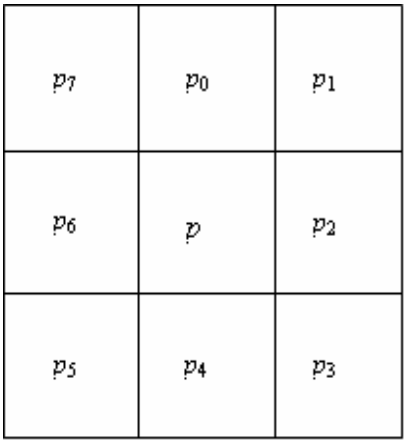


Figure 3.12 Neighbourhood arrangement

In the second step which is used to delete north-west boundary pixels, only Equations 3.33 and 3.34 are changed to

$$p_0 \cdot p_2 \cdot p_6 = 0 \tag{3.36}$$

$$p_0 \cdot p_4 \cdot p_6 = 0 \tag{3.37}$$

3.7 Degrading and Upgrading of Image Resolution

The width of a road in the aerial image with 0.15m spatial resolution is more than 50 pixels. Thus, this high resolution is not necessary for the road extraction. Moreover, the running time is usually proportional with the image size. It is meaningful to degrade the VHR imagery. Once the central line is extracted, this central line can be upgraded to overlay on the original image.

In order to degrade a high resolution image to a low resolution image, the value of each pixel in low resolution can be obtained by calculating the mean value of the reference pixels in a specific window from higher resolution image. The position of the centre reference pixel can be calculated by

$$X_{Creference} = X_{low} * (L_{large} / L_{small}) \quad (3.38)$$

$$Y_{Creference} = Y_{low} * (W_{large} / W_{small}) \quad (3.39)$$

where ($X_{low,}$ Y_{low}) is the coordinate of the pixel in low resolution image, and ($X_{Creference,}$ $Y_{Creference}$) is its centre reference pixel's coordinate in high resolution image, L_{small} and W_{small} are the length and width of the low resolution image. L_{large} and W_{large} are the length and width of the high resolution image. The centre reference pixel is in the centre of the window. The

number of reference pixels depends on the size of the window. The length and the width of the widow can be calculated by

$$L_{window} = L_{large} / L_{small} \tag{3.40}$$

$$W_{window} = W_{large} / W_{small} \tag{3.41}$$

After the centreline had been extracted, the pixels of that line are cast on the original image. The new position can be calculated by equations (3.38) and (3.39). The new enlarged image is a black and white image; the white point is cast from low resolution image. The closing binary morphology is used to connect those points into lines. Finally, the line is overlaid on the original high resolution image.

Chapter 4
Implementation

This chapter illustrates the implementation of the RoadModeler. The comparison between the proposed semi-automatic road extraction strategy and the object-oriented classification method is carried out according to the experimental results, the human operator' experience, and the degree of human intervention.

4.1 Datasets

A set of test images selected from the pansharpened GeoEye-1 colour images with 50 cm resolution, colour aerial orthoimages with 10 cm and 15 cm resolution, and black-and-white aerial orthoimages (see Table 4.1). The pan-sharpened GeoEye-1 colour satellite images and the colour aerial orthoimages covering the City of Davis, California, USA were downloaded from the two public websites (http://www.geoeye.com/CorpSite/) and (http://cityofdavis.org/gis/orthoimagerys/), respectively. The colour orthoimages covering the City of Waterloo, Ontario, Canada were collected from the Map Library at the University of Waterloo. The black-and-white aerial images used in Hu et al. (2007) were provided by the research group at Arizona State University in USA.

Table 4.1 List of the selected test images.

Image ID	Image type, spatial resolution, image size	Location
Figure 4.1 (a)	Colour GeoEye-1 0.5m 1000×1000 pixels	Giza, Egypt
Figure 4.1 (b)		Khalifa Sports City, Doha, Qatar
Figure 4.1 (c)	Colour aerial Orthoimage 0.1m 0.1m7433 × 7433 pixels	Waterloo, Ontario, Canada
Figure 4.1 (d)		Waterloo, Ontario, Canada
Figure 4.1 (e)	Colour aerial orthoimage 0.15m 3200×3200 pixels	City of Davis, California, USA
Figure 4.1 (f)		City of Davis, California, USA
Figure 4.2	Black-and-white aerial orthorimage N/A 823×744 pixels	N/A

In this study, MATLAB was used as a numerical computing environment and programming language. One of its important platforms is for image processing and analysis. It has strong capability for complicated mathematical calculation, friendly programming interface and convenient accesses to tabular data via interactive query.

Nevertheless, in order to make objective comparison with the proposed method, a commercial software package, Definiens Developer Trial, is also used in road network extraction. A 1.79 GHz Processor, 1.87GB of RAM DELL desktop computer was the major hardware for the program operation.

Figure 4.1 Selected test images

Figure 4.2 Black-and-white Aerial image

4.2 Extraction of Roads in Rural Areas

Roads in rural areas usually have few high buildings and trees along the road sides. Thus, there is normally no shadow blocking the roads in the image. However, the width of the rural road usually only accommodates two vehicles side by side in opposite directions, therefore, a big van and its shadow perpendicular to the road direction may block the road in some special cases. Moreover, the quality of the construction of rural roads is usually very crude. As a result, the boundary of the road is very weak, which will make the RoadModeler detect the neighbour areas mistakenly. Figure 4.1(a) shows typical roads in a rural area.

In the object-oriented classification, the multi-resolution segmentation is first operated. Three parameters including scale, the ratio between shape and colour, and the ratio between compactness and smoothness need to be set. For this scene, these parameters are set as 500, 0.1, and 0.5, respectively. The result of segmentation is shown in Figure 4.3(a).

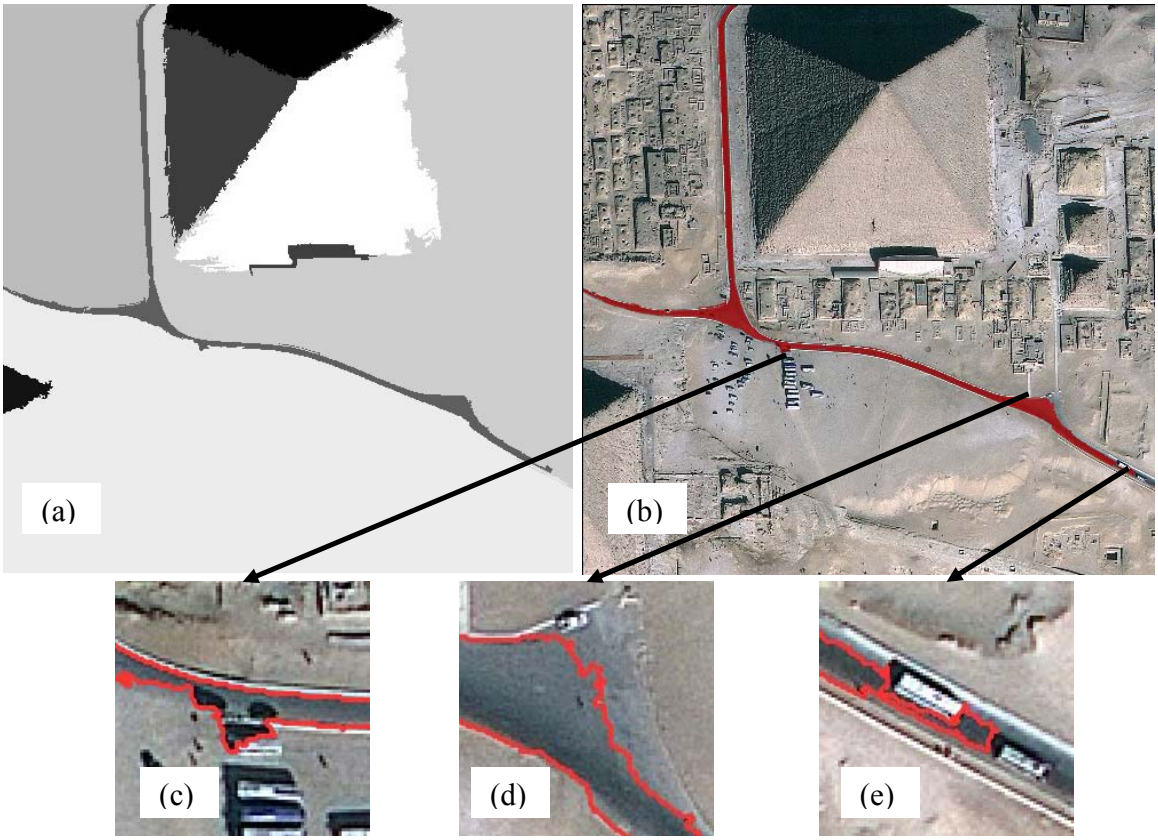


Figure 4.3 Roads in a rural area detected by the object-oriented classification

After segmentation, the image is classified into 9 objects. Since the road is composed of one object, the classification step becomes easier. Here, the classification rule is that objects whose mean values in the red layer are between 78 and 79 are considered as road areas. As

shown in Figures 4.3(b), (c) and (d) present the road segments leaking to the neighbour areas and (e) shows the road segment blocked by two vehicles.

Figure 4.4 demonstrates the results obtained by the RoadModeler by using standard thresholds shown in Table 3.1 and Table 3.2. To avoid the road leaking problem, the single-direction approach is applied first, and the result is shown in (a). After the white circle seed shown in (e) is initialized, the road is covered by a series of circles. From the original image (f), it is easy to see the obvious spectral change indicated by red arrow, so the grey similarity detection algorithm is automatically applied here to generate a blue circle seed shown in (e), and only the grey pixels are changed to blue. The yellow area in (e) is not a complete circle, since some area is covered by vehicles, the narrow similarity detection algorithm is automatically applied to generate a yellow circle seed, and only the pixels whose colour is similar to the reference colour are changed to yellow. (g) and (i) show that the proposed method does not leak to the adjacent area as object-oriented classification does, while the original image of them is shown in (h) and (j). When the circle seed meets with a block such as a vehicle shown in (l), jump similarity detection algorithm is automatically applied to generate a red circle. In this detection algorithm, the pixels with similar colour to the reference colour, together with the pixels in the area between the generated circle and its parent circle seed, are altered to red. Therefore, the vehicle can be part of the red area in (k).

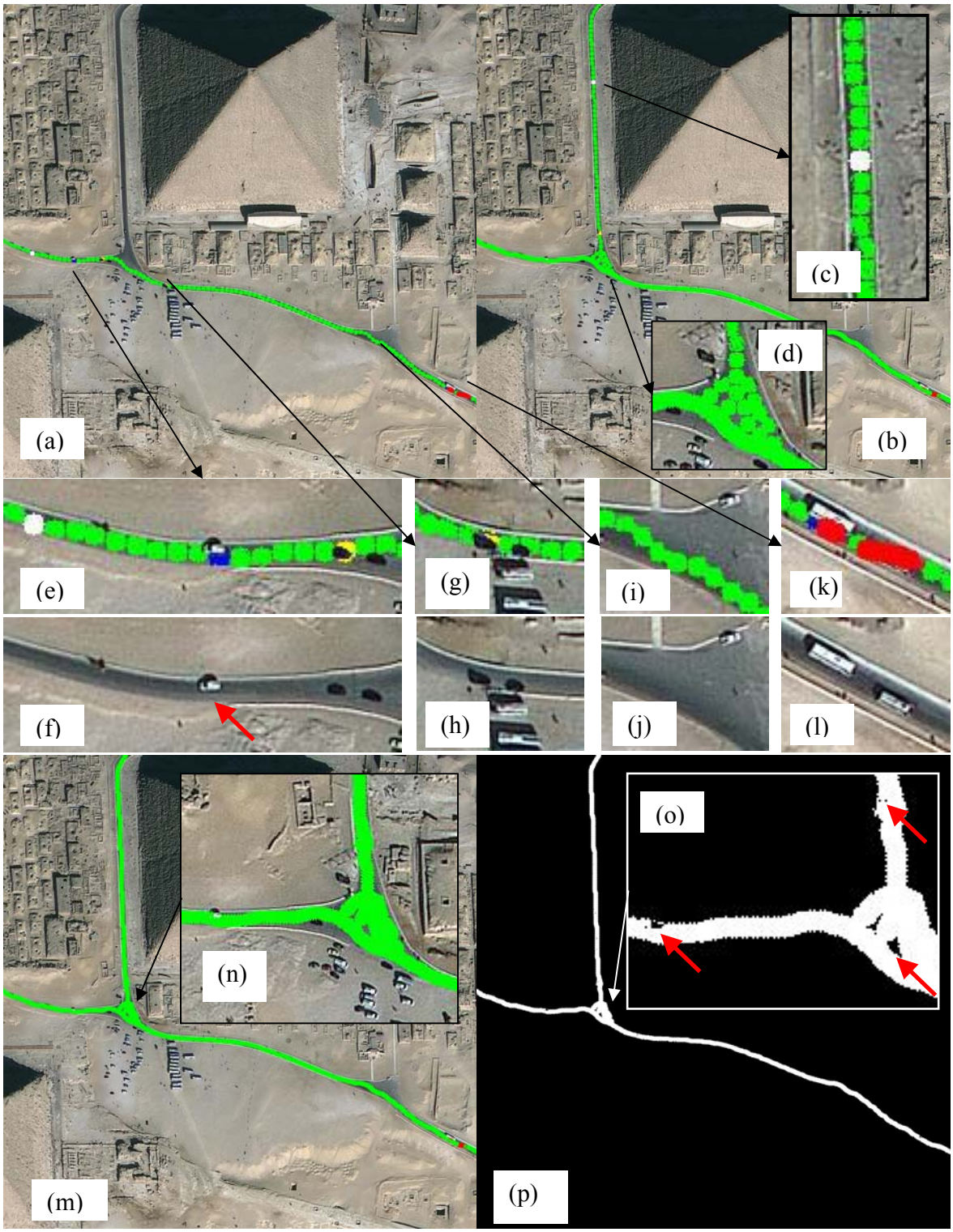


Figure 4.4 Roads in a rural area detected by the RoadModeler

Multidirectional approach follows to extract the undetected roads with width change, then, another circle seed shown in (c) is initialized and detects the remaining road shown in (b). As shown in (d), when a circle seed comes across a wider road area, more circle seeds are generated to cover it.

In order to make the detected road more continuous, the area between the parent circle seed and its daughter circle seeds is also considered as road area and changed to green. (m) shows the continuous road, more specifically, the result in (n) is much better than the result in (d). The continuous result is changed to black and white image for mathematical morphology, since there are some holes and cracks indicated by red arrow shown in (o) in the road (white area).

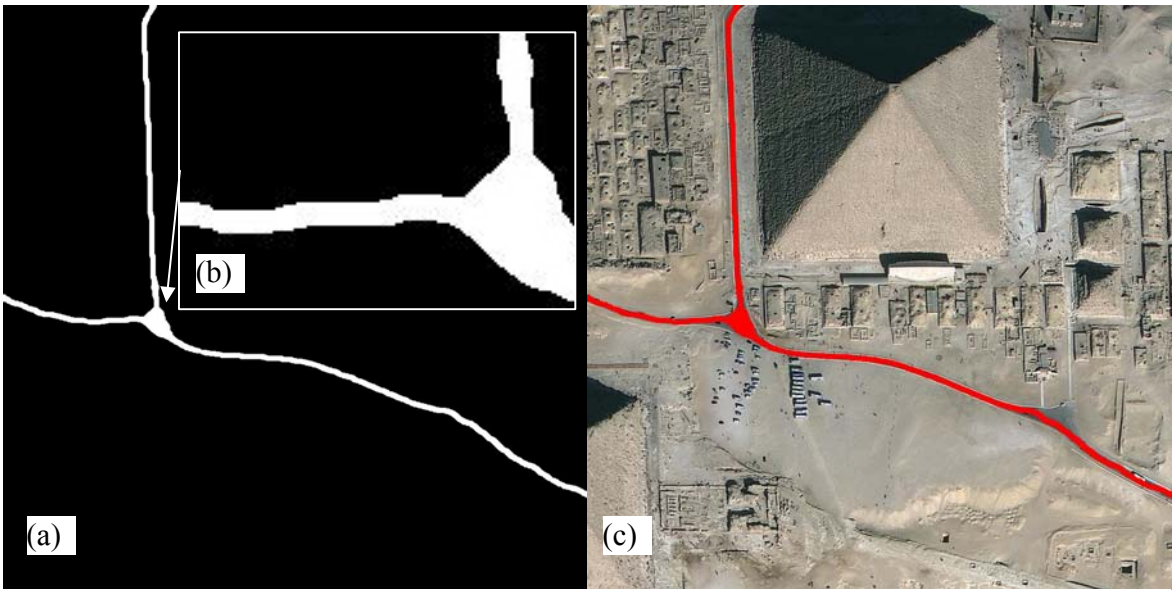


Figure 4.5 Morphologically filtered roads in a rural area and overlaid on the image

Figure 4.5 (a) shows the homogenous road regions (white) after using the binary morphological opening and closing operation that uses the square structuring element with size of 3 × 3 pixels and 11 × 11 pixels respectively. The result in (b) clearly shows that the holes and cracks have been filled. The final result is overlaid on the original image shown in (c). By visual evaluation and analysis, the result obtained by proposed approach, which does not have leakage problem and blockage problem, is better than that obtained by object-oriented classification in which those problems exist.

4.3 Extraction of Roads in Downtown Areas

Roads in downtown areas usually have some white or yellow lane markings and can accommodated several vehicles in parallel. There are also some high buildings casting large shadows on the road. Roundabout or cloverleaf junction makes roads show various shape and curvatures. Parking lots beside the road are always extracted mistakenly, because they have similar spectral information and link with the road.

4.3.1 Extraction of Roads with Roundabouts

Figure 4.1(b) shows roads with roundabouts. This scene embodies two difficulties. One is the road represents different spectral information and various curvatures. The other one is from the parking lot with similar spectral information beside the road. The scale is set as 200; the

other thresholds remain the same as the previous one. The result of segmentation is shown in Figure 4.6(a).

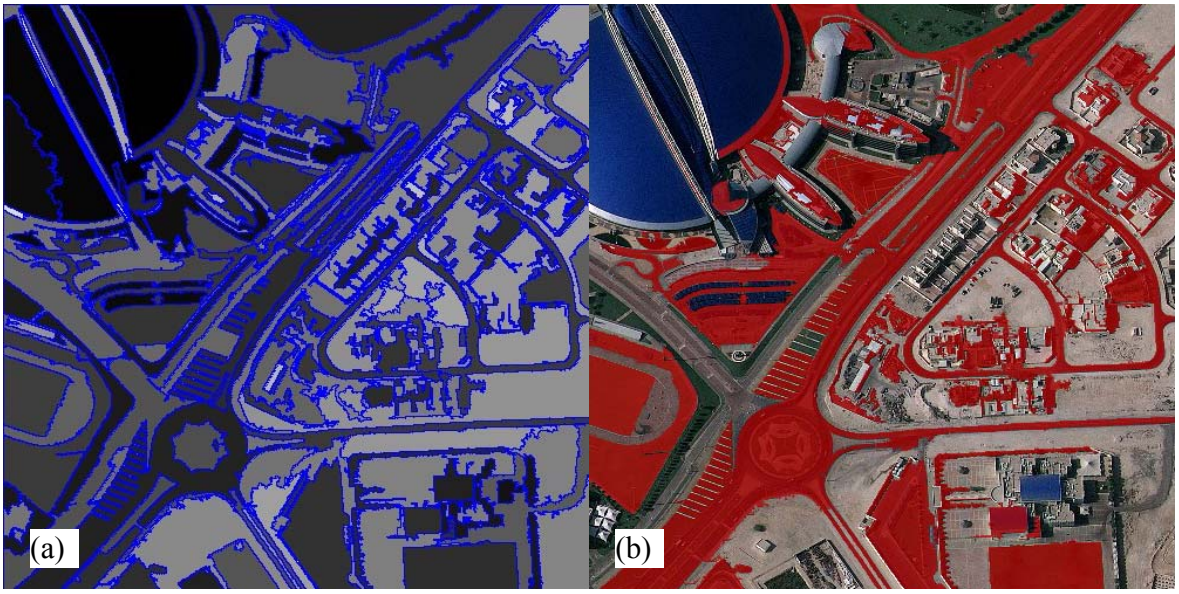


Figure 4.6 Roads with roundabouts detected by the object-oriented classification

After segmentation, the image is classified into 123 objects. In Figure 4.6 (a), the different degrees of grey show the mean values in red band in each segment. The lighter the segment is, the larger the mean value is. It easily can be seen that the segment of the non-road area has similar grey to the one of the road area. Accordingly, the classification rule is that objects with mean value between 35 and 78 in red layer are considered as road areas. From the result shown in Figure 4.6 (b), many non-road areas are extracted mistakenly.

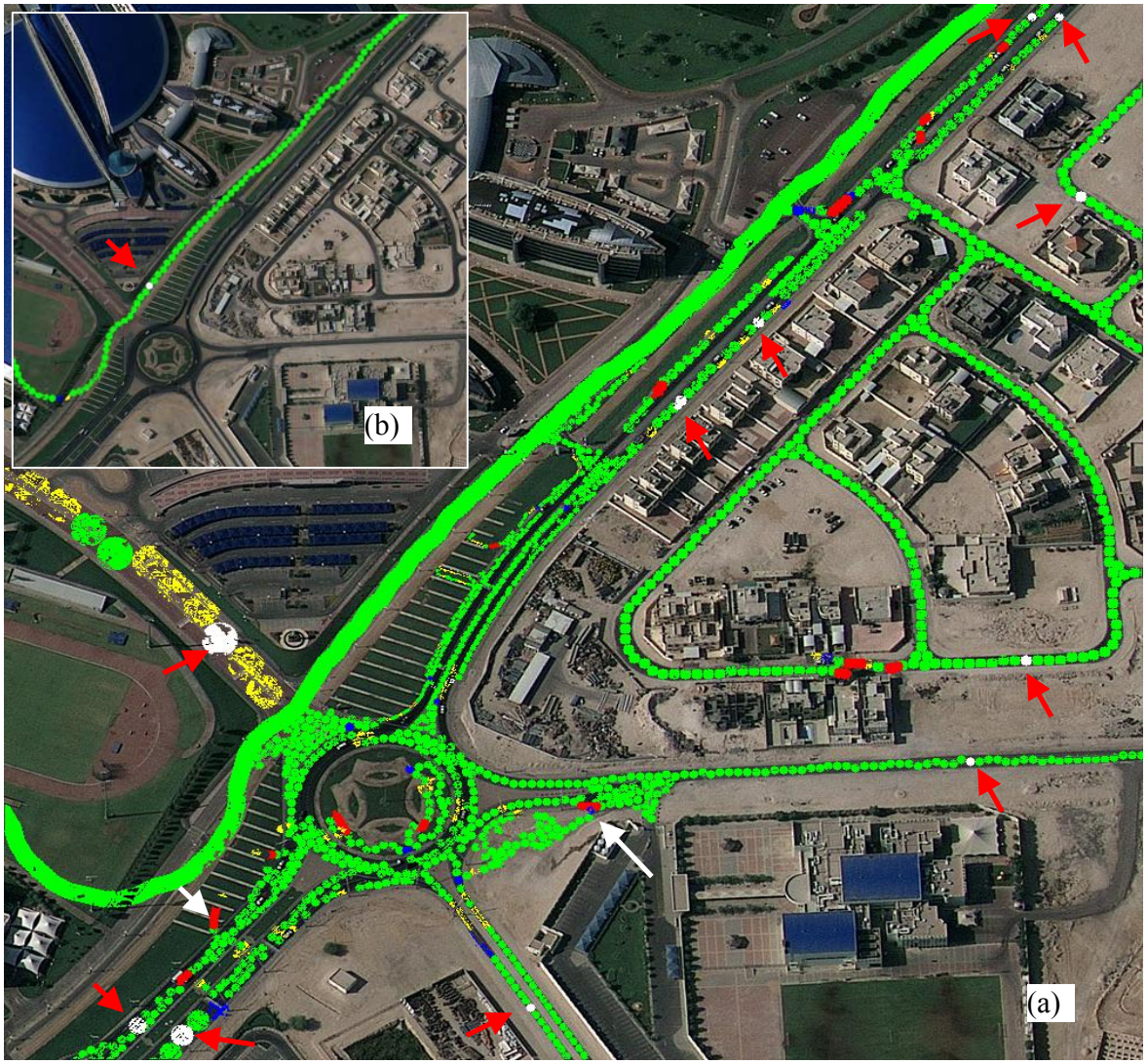


Figure 4.7 Roads with roundabouts detected by the RoadModeler

Figure 4.8 Morphologically filtered roads with roundabouts and overlaid on the image

For this scene, Figures 4.7 and 4.8 show the results of the proposed approach with the same standard thresholds and the morphology and overlaid results. The white circle shown in Figure 4.7 (b) is used in single-direction approach, and some parts of the directed road are

wrong because of the similar spectral information with the neighbour area. After the other white circles shown in (a) are applied by multidirectional approach, it can be seen that it is the jump similarity detection algorithm who leads to the wrong detection, as shown by the white arrow.

4.3.2 Extraction of Roads with Shadows

Figure 4.1(c) shows roads with some shadows cast by buildings. In this scene, not only does the shadow of the building cover the whole width of the road, but also the parking lot with similar spectral information is beside the road. In object-oriented classification, scale, the ratio between shape and colour, and the ratio between compactness and smoothness are set as 1000, 0.1, and 0.5, respectively. The result of the first segmentation is shown in the Figure 4.9(a). The classification rule is that objects whose mean values in red layer are between 131 and 180 are considered as road areas. From the classification result shown in Figure 4.9 (b), most of the roads can be extracted. However, road area covered by the shadow indicated by white arrow is not extracted, while some non-road areas such as parking lots are extracted mistakenly as well. The second segmentation is applied on red area shown in Figure 4.9(c), and the three parameters is set as 500, 0.3, and 0.5, and the classification rule is that the objects in red area whose densities are more than 1.2 are non-road areas. After the second classification, some parking lots and buildings are not considered as road areas in new results (see Figure 4.9d).

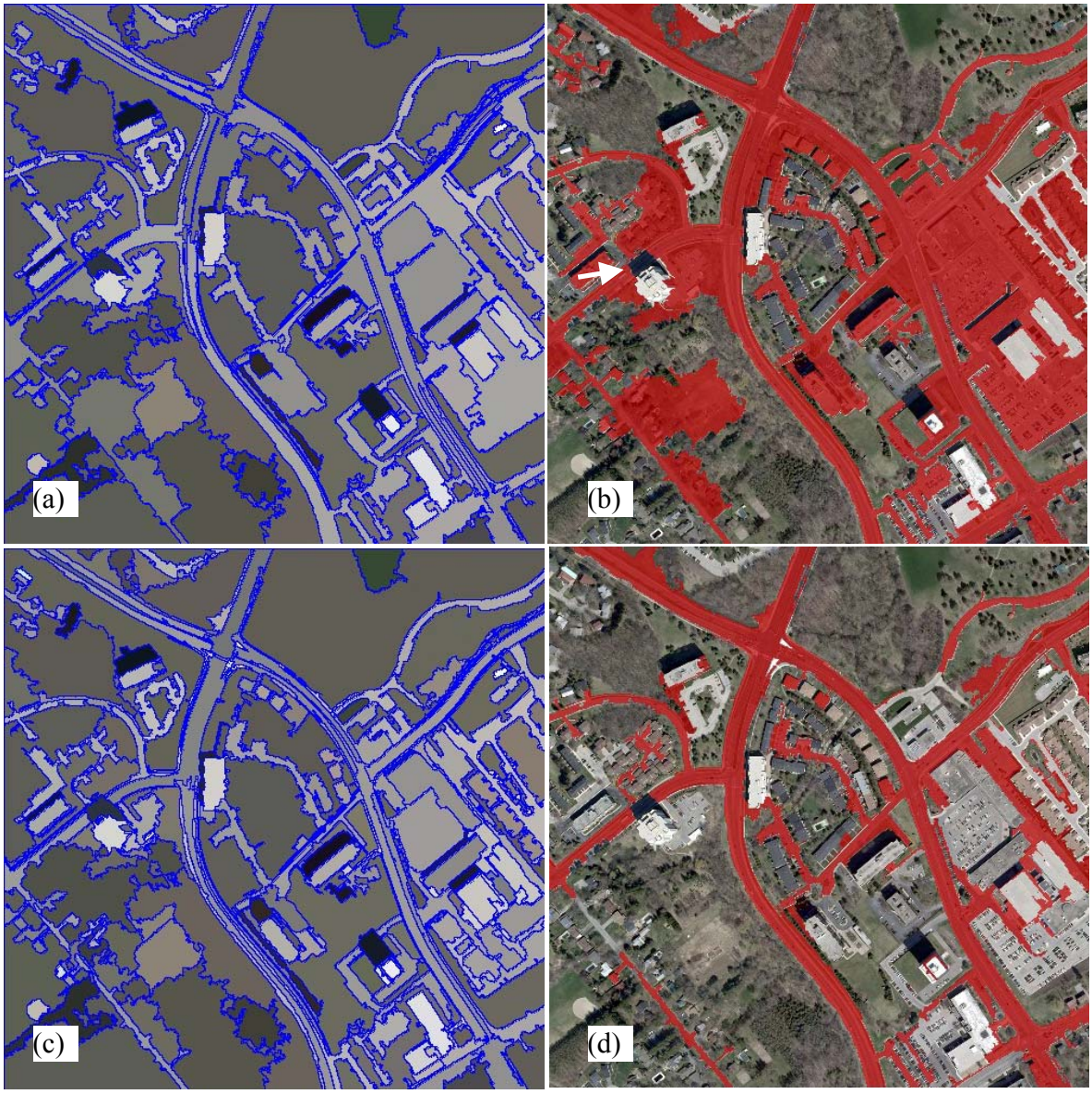


Figure 4.9 Roads with shadows detected by the object-oriented classification

Figure 4.10 Roads with shadows detected by the RoadModeler

For this scene, the RoadModeler with the same standard thresholds listed in Table 3-2 has also been implemented to generate results shown in Figure 4.10. Only single-direction approach is used in this scene to avoid extracting the parking lot beside the road mistakenly. This approach will not be obstructed by the detected road, and it can pass through the extracted road area and keeps its direction to continue detecting. Seven initialized circle seeds (white circle) are required for extracting the main road. The road area covered by shadow is also automatically extracted by jump similarity detection algorithm and covered by red circle seeds shown in Figure 4.10(a). The more continuous result and the black and white result are shown in (b) and (c), respectively. The result of mathematical morphology with the same thresholds and the final overlaid result are shown in Figures 4.11(a) and (b), respectively. The final result is much better than that of the object-oriented classification.

Figure 4.11 Morphologically filtered roads with shadows and overlaid on the image

4.3.3 Extraction of Roads with Parking Lots

Figure 4.1(d) shows roads with parking lots. For the multi-resolution segmentation, scale, the ratio between shape and colour, and the ratio between compactness and smoothness are set as 1000, 0.1, and 0.5, respectively. The result of segmentation is shown in Figure 4.12(a). The classification rule is that objects whose mean values in red layer are between 149 and 200 are considered as road areas. From the result shown in Figure 4.12 (b), most of the road areas can be extracted, while some non-road areas such as parking lots are extracted mistakenly as well.

Figure 4.12 Roads with parking lots detected by the object-oriented classification

Figure 4.13 Roads with parking lots detected by the RoadModeler

In Figure 4.13, the results of RoadModeler with standard thresholds are shown. Three white circle seeds shown in (a) are initialized for single-direction approach. Another six circle seeds indicated by red arrow shown in (b) are initialized for the application of

multidirectional approach. The more continuous result and the black and white result are shown in (c) and (d), respectively. The mathematical morphology results and the final overlaid result are shown in Figures 4.14(a) and (b), respectively. The final result is much better than that of object-oriented classification, as all the roads have been extracted and few non-road areas are extracted as road areas mistakenly.

Figure 4.14 Morphologically filtered roads with parking lots and overlaid on the image

4.4 Extraction of Roads in Residential Areas

Roads in the residential area are usually built by the same material with homogeneous colour, and have salient curbs which yield a good boundary of road. However, some tree shadows

may cause the blockage problem, and the roofs of the houses with similar colour to the road area may cause the leakage problem.

4.4.1 Extraction of Roads with Few Trees

Figure 4.1(e) shows roads with few trees. In this scene, the roads have homogeneous spectral information. In object-oriented classification, the ratio between shape and colour is set as 0.1 in order to generate segment mainly based on spectral information. The other two thresholds remain the same as the previous scene. The result of segmentation is shown in Figure 4.15(a). After segmentation, the image is classified into 47 objects. The classification rule is that objects whose mean values in red layer are between 127 and 137 are considered as road areas. From the result shown in Figure 4.15 (b), some of the house roofs are extracted as road areas mistakenly.

Figure 4.15 Roads with few trees detected by the object-oriented classification

Figure 4.16 Roads with few trees detected by the RoadModeler

Figure 4.17 Morphologically filtered roads with few trees and overlaid on the image

Figure 4.16 shows the results of the RoadModeler with the same standard thresholds and Figure 4.17 shows the morphological filtered and overlaid results. Four white circle seeds are initialized for multidirectional approach, and only an extremely thin path is extracted mistakenly by narrow similarity detection algorithm, which can almost be removed by opening morphology operation.

4.4.2 Extraction of Roads with Numerous Trees

Figure 4.1(f) shows the roads with numbers of trees. In this scene, the tree shadows or their canopy covers some parts of the roads. In object-oriented classification, three parameters: scale, the ratio between shape and colour, the ratio between compactness and smoothness are

set as 500, 0.3, and 0.5, respectively. The result of segmentation is shown in the Figure 4.18(a). After the segmentation, the image is classified into 47 objects. The classification rule is that objects whose mean values in red layer are between 111 and 154 are considered as road areas. From the result shown in Figure 4.18(b), road areas hidden from trees' canopy cannot be extracted. In addition, non-road areas are extracted as roads mistakenly.

Figure 4.18 Roads with numerous trees detected by the object-oriented classification

Figure 4.19 Roads with numerous trees detected by the RoadModeler

Because some of the roads are hidden, the GMin is changed from 20 to 0. For this scene, Figure 4.19 shows the results of the RoadModeler with the decreased GMin and Figure 4.20 shows the morphology and overlaid results. Four white circle seeds are initialized for multidirectional approach, and only a few non-road areas are extracted by jump similarity detection algorithm mistakenly. Most areas are detected by general similarity detection algorithm. Narrow similar detection algorithm is used when half of the road is covered by shadow of trees, and grey similarity detection algorithm is used when most of the road is covered by shadow of trees.

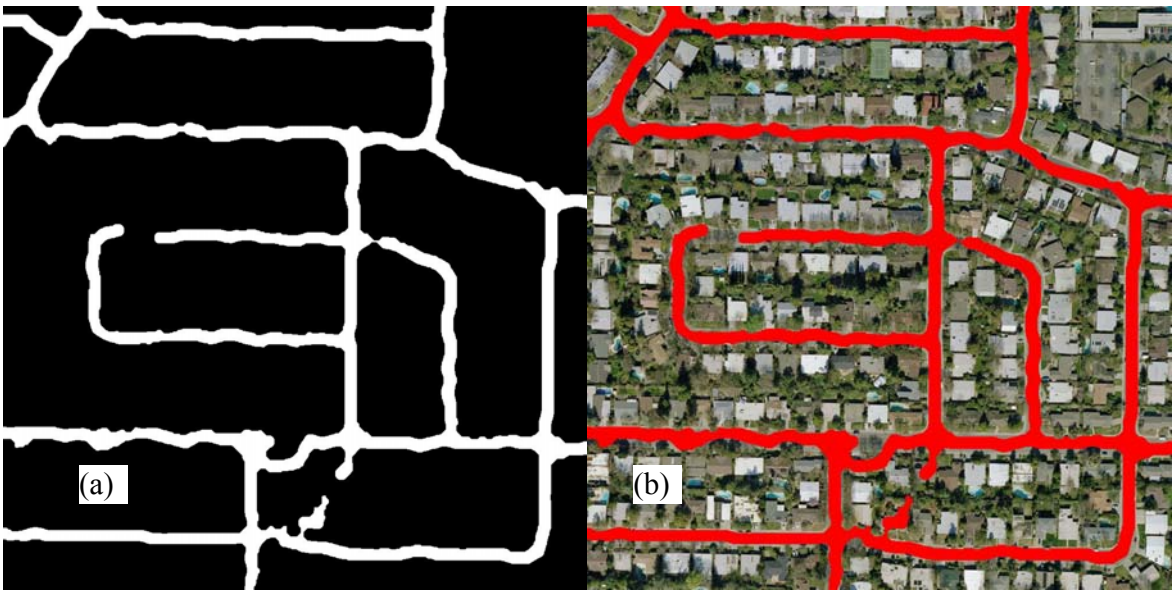


Figure 4.20 Morphologically filtered roads with numerous trees and overlaid on the image

4.5 Extraction of Roads in Black and White Aerial Image

In order to compare the RoadModeler with the method presented by Hu et al, (2007), the same image shown in Figure 4.2 will be used to extract roads by the RoadModeler. For this scene, Figure 4.21 shows the results of the RoadModeler and Figure 4.22 shows the morphological filtered and overlaid results. Six white circle seeds are initialized for multidirectional approach, and only a few non-road areas are extracted by jump similarity detection algorithm mistakenly. Most areas are detected by general similarity detection algorithm. Narrow similar detection algorithm is used when half of the road is covered by shadow of trees, and jump similarity detection algorithm is used when the material of the road changed suddenly. The centerline will be extracted and compared with the results of Hu et al, (2007) in the next Chapter.

Figure 4.21 Roads in black and white aerial image detected by the RoadModeler

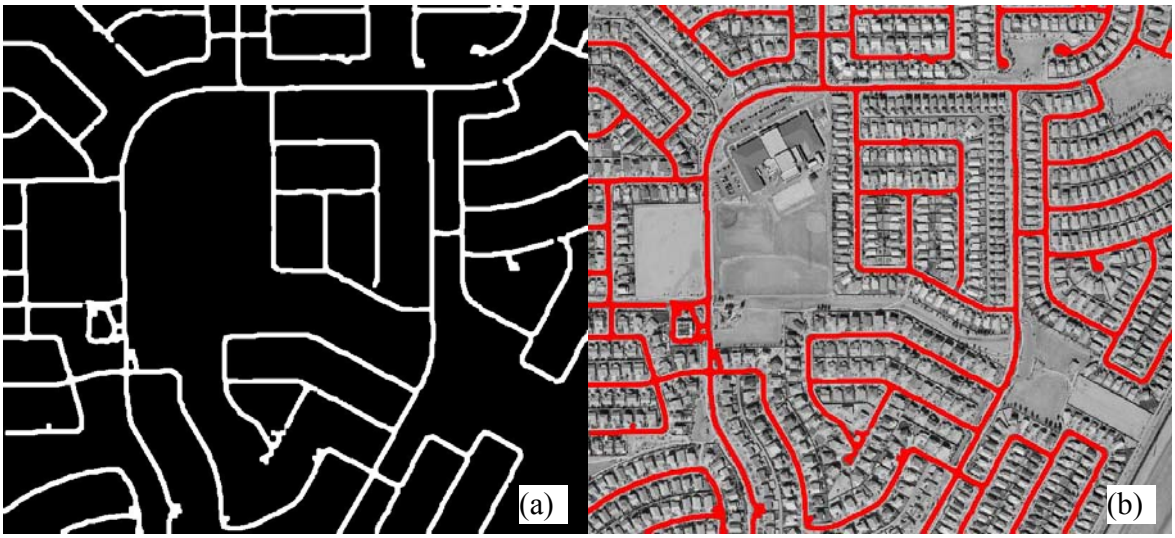


Figure 4.22 Morphologically filtered roads and overlaid on the black and white aerial image

4.6 Chapter Summary

From the comparison between object-oriented classification and the RoadModeler, although the object-oriented classification has only three parameters which control the size, the homogeneous degree, the compactness degree, and smoothness degree of the generated objects, it is difficult to generate desired objects with the uniform feature classified by the a single rule when the scene is complicated.

By visual evaluation and analysis, the RoadModeler can obtain much better results than the object-oriented classification approach. In addition, the RoadModerle is easier to be used than object-oriented classification. The RoadModeler with the same thresholds is used to extract roads. What the human operator is required to do is to choose the positions and the

radii of some circle seeds and the corresponding approaches. For object-oriented classification, different parameters are used in different scenes for the segmentation. Different classification rules also need to be made by a human operator. Although the parameters and the classification rules used for different scenes are not the best, because the objective of a tool is to reduce the labour of the human operator and save the time, it is meaningless to spend much more time in finding the best parameters and classification rules for different scenes when the post-extraction of human is become much easier.

After testing eight different types of scenes, the strategy for two approaches is that when the parking lot or any other area with the similar spectral information is located beside the road, one-directional approach should firstly be adapted to avoid the leakage problem. Besides the scenes discussed above, RoadModeler can also be used to extract bridges, because the material of a bridge is the same and appears as homogeneous spectral information on the remote sensing image. When a road extends from paved area to unpaved area, if the road keeps its direction without turning, grey similarity detection can be applied to detect roads when spectral information of the roads change suddenly. When grey similarity detection does not work because of the road curvature, a human operator can initial another circle seed to track the unpaved road. RoadModeler cannot extract long hidden roads such as channels, which require an operator to manually extract based on human knowledge.

The objective of RoadModeler is to facilitate a human operator, so it can provide the most convenient assist to handle different road situations. In most cases, RoadModeler cannot

extract all the roads, and there are usually some ruptures in the road network. Therefore, before the GIS database updating, the manually post-extraction by human operator is required to fill the rupture, to extract the undetected road, and to remove the mistakenly detected road.

Chapter 5
Performance Assessment

This chapter introduces a qualitative and quantitative evaluation method for the accuracy assessment of RoadModeler. The effectiveness of the proposed RoadModeler is further investigated by comparison with the results mentioned in Hu et al. (2007).

5.1 Evaluation Method

From an algorithmic point of view, the extraction accuracy typically defines the success of a feature extraction method (Agouris et al., 2004). Accuracy is commonly measured by comparing the algorithm output against the reference data either from the GIS database or through manually digitizing. According to Wiedemann et al. (1998) and Hu et al. (2007), three indexes to evaluate the quality of road extraction are as follow:

$$\textit{Correctness} = \frac{L_m}{L_r} \tag{5.1}$$

$$\textit{Completeness} = \frac{L_m}{L_e} \tag{5.2}$$

$$\textit{Quality} = \frac{L_m}{L_{ur}+L_e} \tag{5.3}$$

where L_r is the number of pixels composing of the reference road R, L_e is the number of the pixels composing of the extracted road V, L_{me} shown in Figure 5.1(a) is the number of the pixels composing the extracted road that is in the reference buffer (seethe shaded area shown in Figure 5.1a), which is the area in which the distance between pixels and the reference road is less than a given tolerance T. L_{mr} shown in Figure 5.1(b) is the number of the pixels composing the reference road that is in the extraction buffer (the shaded area shown in Figure 5.1b), which is the area in which the distance between pixels and extracted road is less than a given tolerance T. $L_m = \min (L_{me}, L_{mr})$, and L_{ur} is the number of the reference roads that are out of the extraction buffer. In this study the width of the road is selected as the tolerance T.

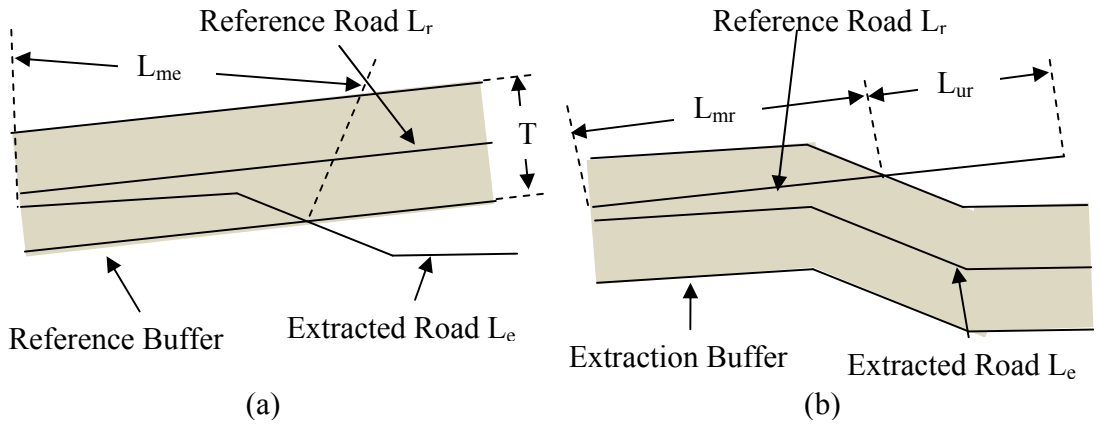


Figure 5.1 Buffers for the extracted roads and the reference roads

5.2 Evaluation Results

Figures 5.2 (a), (c) and 5.4 (a), (c) show the reference road centrelines for Figures 4.1 (a), (b), (e) and (f), respectively. These reference data are generated by on-screen manual digitizing. Figures 5.3 (a) and (c) show the reference road centrelines for Figures 4.1 (c) and (d), respectively. These reference data refer to the GIS database (obtained from the Map Library of the University of Waterloo). Figure 5.5 (a) are the reference road centrelines which refer to a GIS database and (b) is the extracted centreline by using the method of Hu et al. (2007). These reference and extracted data are provided by Hu et al. (2007). The roads extracted by RoadModerler are input into the thinning algorithm to obtain the road centrelines. Figures 5.2 (b), (d) and 5.3(b), (d) and 5.4 (b), (d) and 5.4(c) present the extracted road centrelines by the thinning algorithm.

Figure 5.2 Reference vs. extracted road centerlines of GeoEye-1 satellite images: (a) and (c) Reference centerlines generated by on-screen manual digitizing, (b) and (d) centrelines extracted by the RoadModeler

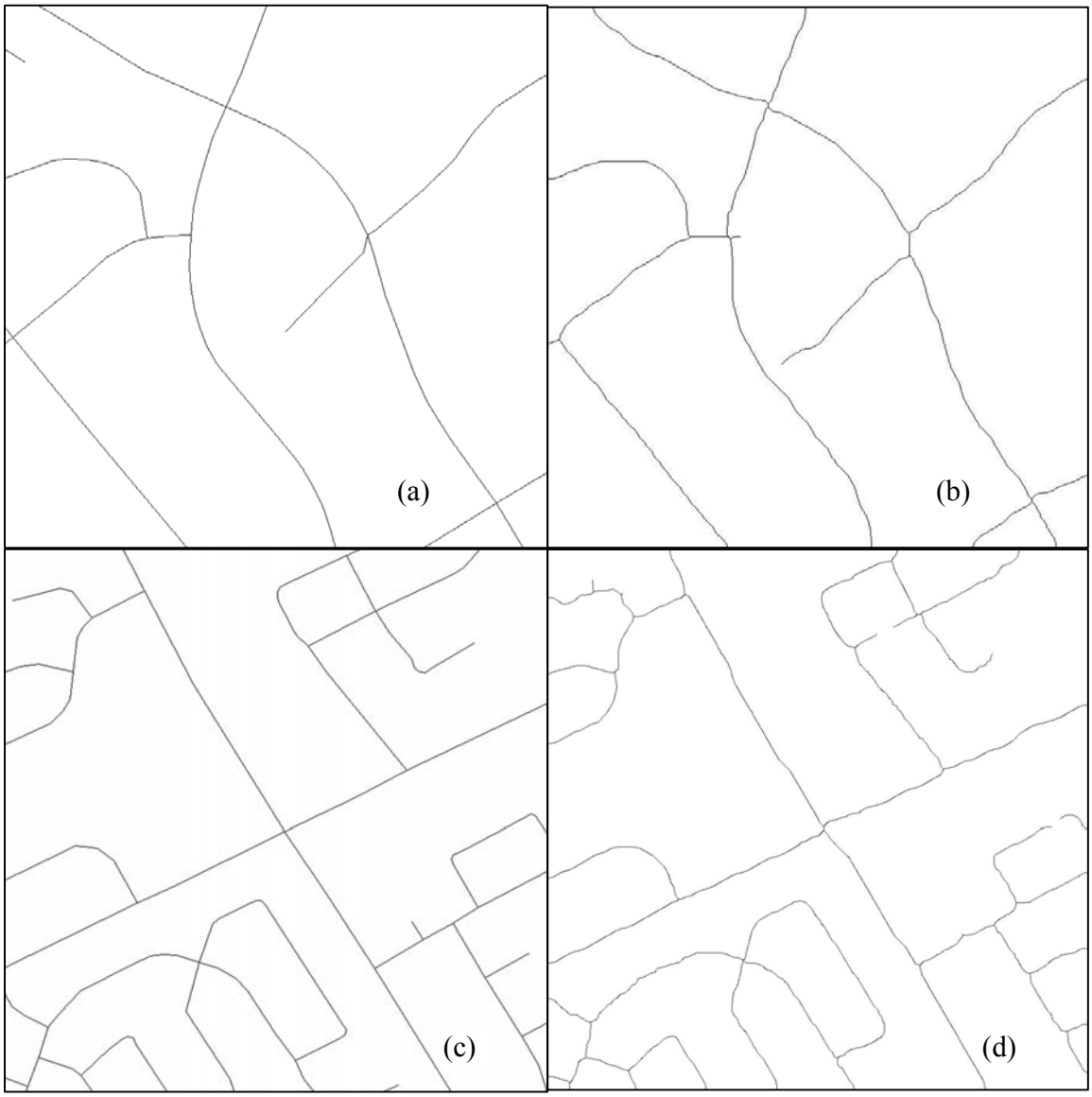


Figure 5.3 Reference vs. extracted road centerlines of aerial image for city of Waterloo: (a) and (c): Reference centerlines obtained from GIS database, (b) and (d) centrelines extracted by the RoadModeler

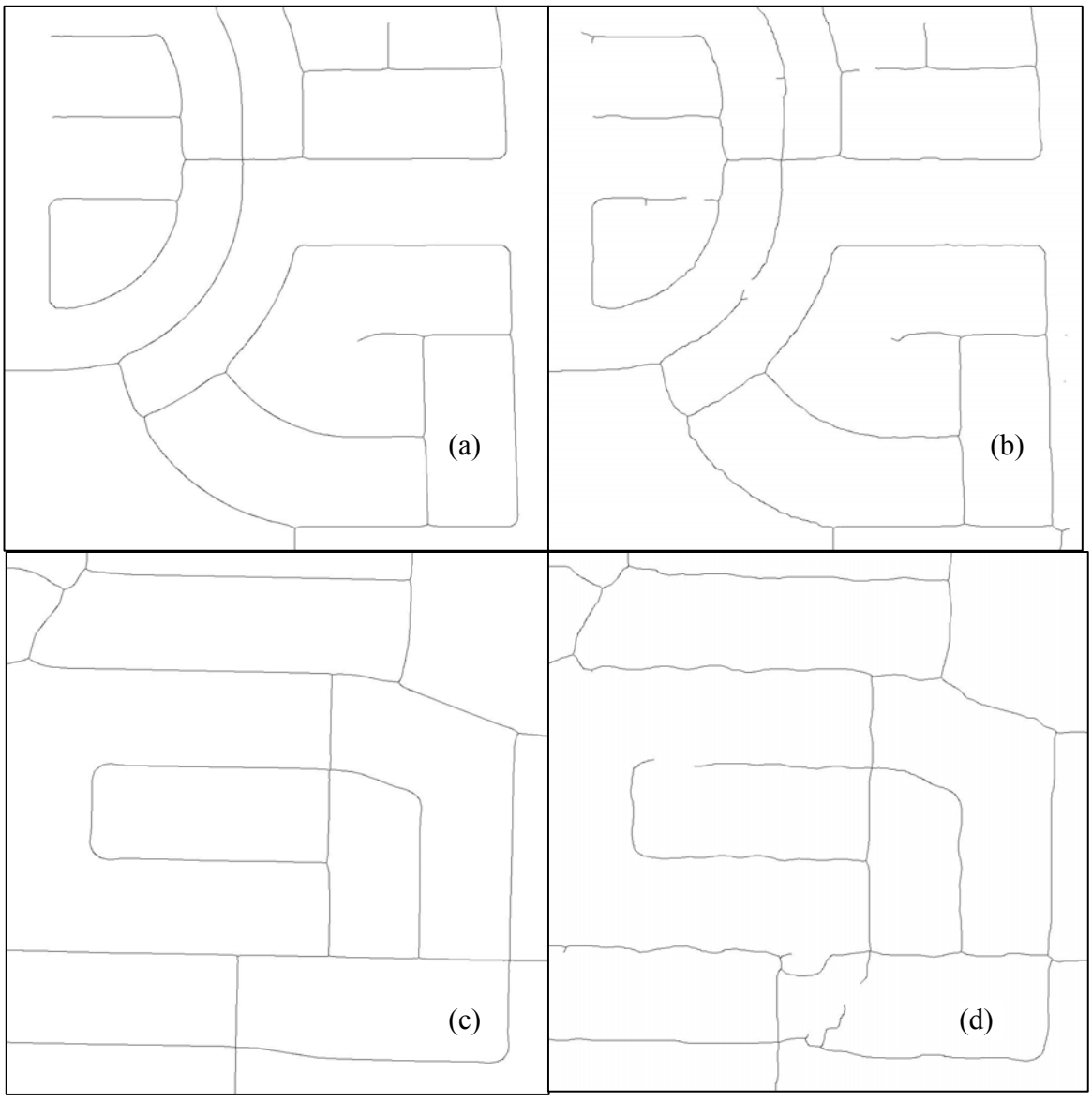


Figure 5.4 Reference vs. extracted road centerlines of aerial image for city of Davis: (a) and (c) Reference centerlines generated by on-screen manual digitizing, (b) and (d) centrelines extracted by the RoadModeler

Figure 5.5 Reference vs. extracted road centerlines of black and white aerial image: (a) Reference centrelines, (b) centrelines extracted by Hu et al (2007), (c) centrelines extracted by the RoadModeler

All the evaluation values for these seven scenes are shown in Table 5.1. Most of them are higher than 90%. The average values for completeness, correctness, and quality and the number of initial circle seeds in the seven scenes are 94%, 97%, 94%, 6, respectively. The

RoadModeler has higher average completeness, correctness, and quality than Hu et al. (2007) whose corresponding evaluation results are 91%, 90%, and 85%, respectively.

Table 5.1 Evaluation results with standard thresholds

Image ID Figure	Buffer Tolerance (pixels)	Completeness (%)	Correctness (%)	Quality (%)	Number of Initial Seeds
4-1a	10	97	100	100	2
4-1b	20	82	93	83	12
4-1c	40	99	98	97	7
4-1d	32	94	100	99	9
4-1e	40	97	96	93	4
4-1f	40	97	94	91	4
4-2	16	95	98	93	6
Average		94	97	94	6

The images used by Hu et al. (2007) were black-and-white aerial images with relatively low spatial resolution, where the widths of roads are less than 8 pixels on the images. The RoadModeler is appropriate for road extraction from VHR image. Therefore, the image with a relative large spatial resolution is used for the comparison. For the same scene, completeness, correctness, and quality of Hu et al. (2007) are 93%, 96%, 92%, respectively. Compared with this results, the corresponding evaluation values of the RoadModeler which are 95%, 98%, and 93% are better.

Additionally, the images used by Hu are black and white images, while the RoadModeler can be used on not only black and white images but also colour images so that the RoadModeler

can integrate more information for road extraction, which increases the extraction performance.

Furthermore, the scenes presented in Hu et al. (2007) are simple, because there are no parking lots and trees beside the roads, and no shadows, vehicles, and marking lines shown inside the roads. All these situations are shown on the images mentioned in Chapter 4.

Chapter 6
Conclusions and Recommendations

This chapter presents conclusions based on the findings of this study, and recommendations for further research.

6.1 Conclusions

In this thesis, a semi-automatic method called RoadModeler has been developed for road extraction from the VHR imagery. This approach brought significant improvement in completeness, correctness, and quality.

The developed RoadModeler integrates five types of detection algorithms, including general similarity, narrow similarity, grey similarity, jump similarity, and single-direction detection algorithms into two approaches. They are the multidirectional detection approach and the single-direction approach. Once an initial circle seed is placed and the approach type is selected by the human operator, a road could be detected by using different detection algorithms. These algorithms help extract the roads with different widths, different pavement materials, various lane markings, and disturbed by trees, buildings, and vehicles.

The computing time of the RoadModeler is highly dependent on the complexity of the scenes, instead of that of the algorithms. The five types of detection algorithms are operated

sequentially. If the first detection algorithm (e.g., general similarity) can find a road, the second detection algorithm (e.g., narrow similarity) will not be needed and thereafter. Conversely, if the first detection algorithm cannot find a road, then the second detection algorithm will be required. Regarding a simple case, the computing time comes from only one of the detection algorithms. For a complicated case, a large amount of time must be spent on the selection of detection algorithms. In addition, the only area calculated by the RoadModeler is the road and its surrounding areas. . The computing time of the RoadModeler is proportional to the road areas, rather than the entire image. Therefore, the running time is dependent on the complexity and the area of road network.

The placement of the circle seed and the selection of the road detection approaches are based on the decisions made by the human operator. Correct use of single-direction approach can extract roads from linked parking lots, while correct use of multidirectional approach can extract most roads based on one single circle seed. The RoadModeler requires a human operator to manually place an initial circle seed, to select different models, and to restrict the parameters. Most of the cases, the RoadModeler with standard thresholds can generate desired results; accordingly it is easy to use.

The RoadModeler has been used to extract different types of roads in rural areas, downtown areas, and residential areas from satellite and aerial images, respectively. The new strategy for aerial images with 0.15m and 0.1m spatial resolution is to degrade the resolution of the image, extract the road network from the low resolution image, and to upgrade the result to

the original image. This developed strategy requires less computing time while still producing a satisfied result. Most importantly, the developed RoadModeler has successfully resolved the blockage and leakage problems which had puzzled previous researchers.

The results produced by the RoadModeler shows that the entire road has been detected with less mistakes. From the visual evaluation, the proposed approach was much better than the object-oriented classification method. From a more accurate evaluation through the use of standard thresholds, the average completeness, correctness, and quality have achieved 94%, 97%, and 94%, respectively. They are higher than those of Hu et al. (2007) which are 91%, 90%, and 85%, respectively. The same image provided by Hu et al. (2007) is used to extract roads by the RoadModeler and obtained the corresponding evaluation results which are 95%, 98%, and 93%. The results is also better than that of Hu et al. (2007), which are 93%, 96%, and 92%, respectively.

Overall, this study has successfully developed an effective RoadModeler for semi-automatic extraction of road networks form the VHR imagery. Noticeably, the proposed strategy has higher extraction accuracy than the object-oriented classification method.

6.2 Recommendations for Future Research

Future improvements can be made towards each individual component of the RoadModeler. An interesting challenge will be to integrate more automatic processes into the RoadModeler.

Firstly, initializing a circle seed requires a human operator to point out the centre of the road and the nearest boundary. If the circle seed can be generated automatically by specifying the centre point of the road, the operation can be carried out. The width of the road can be automatically calculated by analyzing the background of that centre point.

Secondly, the results of the RoadModeler are used to extract the centreline of the road network. However the sidelines cannot be extracted precisely. The methods such as snakes may be applied here to obtain more accurate sidelines of a road.

Thirdly, the generated circle seeds in a road can be considered a whole to be validated by using road-shape information. A good validation of the generated circle seed can guarantee more tolerant and flexible thresholds, which reduces unnecessary work load of the human operator required for adjusting the thresholds.

In addition, morphological operations can be applied on a specific region, instead of the whole image. This helps the human operator to achieve better results.

Moreover, algorithms needs to be improved to be able to removed some branches which are derived from the thinning process.

Furthermore, the RoadModeler was developed by using MatLab which has low computation efficiency. The implementation of the RoadModeler using C++ programming may improve that.

In summary, based on the semi-automatic road extraction strategy, the RoadModeler has demonstrated to be efficient, accurate and reliable. With modifications suggested above, the RoadModeler may become of commercial value.

References

Agouris, P., Doucette, P., and Stefanidis, A., 2001. Spatiospectral cluster analysis of elongated regions in aerial imagery, *Proceedings of the IEEE International Conference on Image Processing*, 07–10 Oct, Thessaloniki, Greece, pp. 789–792.

Amini, J., Saradjian, M.R., Blais, J.A.R., Lucas, C., and Azizi, A., 2002. Automatic road side extraction from large scale image maps. Internat. *J. Appl. Earth Observat. Geoinformat.* 4, 95–107.

Amo, M., Martinez, F., and Torre, M., 2006. Road extraction from aerial images using a region competition algorithm. *IEEE Transactions on Image Processing*, 15(5):1192-1201.

Asano, T., Chen, D.Z., Katoh, N., and Tokuyama, T., 1996. Polynomial-time solutions to image segmentation. *www. Portal.acm.org*.

Bajcsy, R., Tavaloki, M., 1976. Computer recognition of roads from satellite picture. *IEEE Trans. Systems Man Cybernet*. 6, 76–84.

Baraldi A., and Parmiggiani, F., 1994. A Nagao-Matsuyama approach to high-resolution satellite image classification, *IEEE Transactions on Geoscience and Remote Sensing*, 32(4):749–758.

Baumgartner, A., Hinz, S., and Wiedemann, C., 2002. Efficient methods and interfaces for road tracking," *International Archives of Photogrammetry and Remote Sensing*, 34(3B): 309-312

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., and Ebner, H., 1999. Automatic road extraction based on multi-scale, grouping, and context, *Photogrammetric Engineering & Remote Sensing,* 65(7): 777-785.

Bonnefon, R., Dhérété, P., and Desachy, J., 2002. Geographic information system updating using remote sensing images, *Pattern Recognition Letters*, 23(9):1073–1083.

Chen, A., Donovan, G., Sowmya, A., and Trinder, J., 2002. Inductive clustering: Automating low level segmentation in high resolution images. *In: ISPRS Photogrammet. Comput*. Vision, 9–13 September, Graz, Austria, p. A-73.

Chen, K.M., and Chen, S.Y., 2002. Colour texture segmentation using feature distributions. *Pattern Recognition Letters*, 23: 755-771.

Couloigner, I., and Ranchin, T., 2000. Mapping of urban areas: A multiresolution modeling approach for semi-automatic extraction of streets, *Photogrammetric Engineering & Remote Sensing*, 66(7): 867–874.

Doucette, P., Agouris, P., Musavi, M., and Stefanidis, A., 1999. Automated extraction of linear features from aerial imagery using Kohonen learning and GIS data, *Lecturer Notes in Computer Science: Integrated Spatial Databases Workshop—Digital Images and GIS*, 1737: 20–33.

Definiens, 2008. Definiens eCognition homepage, URL: http://www.definiens.com/ (last date accessed: 22 April 2009)

Gibdaughter, L., 2003. Finding road networks in IKONOS satellite imagery, *Proceedings of the ASPRS Annual Conference, 05-09 May, Anchorage, Alaska (American Society for Photogrammetry and Remote Sensing), unpaginated CD-ROM*.

Gorin, B., 2005. Performance metrics for pan sharpening methods and comparison with ground truth, *ASPRS Annual Conference*, March 7-11, Baltimore, Maryland.

Gruen, A. and Li, H., 1997. Semi-automatic linear feature extraction by dynamic programming and LSB-snakes. *Photogrammetric Engineering and Remote Sensing,* 63 (8): 985–995.

Gruen, A., 1985. Adaptive least squares correlation - A powerful image matching technique, *South African Journal of Photogrammetry*, *Remote Sensing and Cartography*, 14(3):175-187.

Guindon, B., 1998. Application of spatial reasoning methods to the extraction of roads from high resolution satellite imagery, *Proceedings of International Geoscience and Remote Sensing Symposium*, 6-10 July, Seattle, Washington, pp. 1076-1078.

Hinz, S., Baumgartner, A., Mayer, H., Wiedemann, C., and Ebner, H., 2001. Automated Extraction of Man-Made Objects from Aerial and Space Images (E. Baltsavias, A. Gruen and L. Van Gool, editors), A. A. Balkema Publishers, Lisse, The Netherlands, pp. 255-265.

Haverkamp, D., 2002. Extracting straight road structure in urban environments using IKONOS satellite imagery, *Optical Engineering*, 41(9): 2107-2110.

Hu, J., Razdan, A., and Femiani J.C., 2007. Road network extraction and intersection detection from aerial images by tracking road footprints, *IEEE Transaction on Geoscience and Remote Sensing*, 45(12): 4144-4157.

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

Jensen, J.R., 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3/e, Prentice Hall, Upper Saddle River, pp. 420-430.

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: Active contour models. *Internat. J. Comput. Vis*., 321–331.

Kim, T., Park, S. R., Kim, M. G., Jeong, S. and Kim, K. O., 2004. "Tracking road centerlines from high resolution remote sensing images by least squares correlation matching," *Photogrammetric Engineering and Remote Sensing*, 70(12): 1417-1422.

Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., and Baumgartner, A., 2000. Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision Applicat*. 12 (1), 23–31.

Long, H. and Zhao, Z., 2005. Urban road extraction from high resolution optical satellite images, *International Journal of Remote Sensing*, 26(22): 4907– 4921.

Matheron, G., 1975. *Random Sets and Integral Geometry*, John Wiley and Daughters, New York, USA.

Mayer, H., Laptev, I., Baumgartner, A., 1998. Multi-scale and snakes for automatic road extraction. *In: 5th Europ. Conf. on Comput. Vision*, I: 720–733.

Mckeown, D. M. and Denlinger, J. L., 1988. Cooperative methods for road tracking in aerial imagery, in *Proceedings of the IEEE Conference in Computer Vision and Pattern Recognition*, pp. 662–672.

Mena, J. B., Malpica, J. A., 2003. Color image segmentation using the Dempster Shafer theory of evidence for the fusion of texture. *Pattern Recognition. Internat. Arch. Photogrammet. Remote Sensing* 34, Part 3/W8.

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

Mohammadzadeh, A., Tavakoli, A., and Valadanzoej, M., 2006. Road extraction based on fuzzy logic and mathematical morphology from pansharpened IKONOS images, *The Photogrammetric Record*, 21(113): 44–60.

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

Neuenschwander, W., Fua, P., Iverdaughter, L., Szekely, G., and Kubler, O., 1997. Ziplock snakes. *International Journal of Computer Vision* 25 (3): 191–201.

Niu, X., 2006. A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model. *ISPRS Journal of Photogrammetry and Remote Sensing* 61: 170–186.

Park, S. R. and Kim, T., 2001, Semi-automatic road extraction algorithm from IKONOS images using template matching, in *Proceedings of 22nd Asian Conference on Remote Sensing*, Singapore, pp. 1209-1213.

PCI, 2004. Pan Sharpening in Geomatica 9, available at http://www. Pcigeomatics.com/products/viewlets/pansharpening/pansharp_viewlet_final.swf

Pigeon, L., Moulin, B., Solaiman, B., Toutin, T., and Thomdaughter, K. P. B., 1999. Human-experts features extraction strategies for topographical map production and updating, *Proceedings of International Geoscience and Remote Sensing Symposium*, 28 June–02 July, Hamburg, Germany, pp. 296–298.

Quackenbush, L.J., 2004. A review of techniques for extracting linear features from imagery. *Photogrammetric Engineering & Remote Sensing*, 70(12): 1383–1392.

Serra, J., 1982. *Image Analysis and Mathematical Morphology*, Academic Press, London, England.

Shneier, M. O., 1982. Extracting linear features from images using pyramids, *IEEE Transactions on Systems, Man and Cybernetics,* July–Aug: 569–572.

Shukla,V., Chandrakanth, R., and Ramachandran, R., 2002. Semi-automatic road extraction algorithm for high resolution images using path following approach, *in ICVGIP02*, Ahmadabad, Vol. 6, pp. 31-236,

Song, M. and Civco, D., 2004, Road extraction using SVM and image segmentation. *Photogrammetric Engineering and Remote Sensing,* 70(12): 1365–1372.

Trinder, J.C. and Wang, Y., 1998. Automatic road extraction from aerial images. *Digital Signal Processing* 8 (4): 215–224.

Trinder, J. and Li, H., 1995. Semi-automatic feature extraction by snakes, Automatic Extraction of Man-Made Objects from Aerial and Space Images, *Birkhaeuser Verlag*, pp. 95-104.

Vosselman, G., De Knecht, J., 1995. Road tracking by profile matching and Kalman filtering. In: Gruen, Kuebler, Agouris, (Eds.), *Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkauser, Basel, pp. 265–274.

Wiedemmann, C., Heipke, C., and Mayer, H., 1998. Empirical evaluation of automatically extracted road axes, in *Proc. CVPR Workshop Empirical Eval. Methods Comput. Vis.*, Los Alamitos, CA, pp. 172–187.

Wyszecki, G. and Stiles, W. S. 1967. *Color Science, Concepts and Methods, Quantitative Data and Formulas*, Wiley, New York, USA.

Zhang, Q. and Couloigner, I., 2006. Benefit of the Angular Texture Signature for the Separation of Parking Lots and Roads on High Resolution Multi-spectral Imagery. *Pattern Recognition Letters* 27(9), pp. 937–946.

Zhang, T. Y. and Suen, C. Y. 1984. A fats parallel algorithm for thinning digital patterns. *Communications of the ACM*, 27(3), 236-239.

Zhao, H., Kumagai, J., Nakagawa, M., and Shibasaki, R., 2002, Semi-automatic road extraction from high resolution satellite images, *ISPRS Photogrammet. Comput. Vision*, Graz, Ustrailia, September, pp. A-406.

Zhou, J., Bischof, W. F., and Caelli, T., 2006 Road tracking in aerial images based on human-computer interaction and Bayesian filtering, *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 61: 108-124.

Zlatanova, S. and Li, J. (eds.), 2008. Geospatial Information Technology for Emergency Response, *ISPRS Book Series*, Vol. 6, Taylor & Francis, London, 382pp.