Study on Road Network Extraction using High-Resolution Satellite Imagery

By

Lili Yun

Graduate School of Science and Technology

KUMAMOTO UNIVERSITY

Submitted for the degree of

Doctor of Philosophy

February 2007

主査　内村　圭一　教授

副査　川路　茂保　教授

副査　宇佐川　毅　教授

副査　北園　芳人　教授

副査　胡　振程　助教授

Acknowledgements

First, My thanks go to my research supervisor, Professor Keiichi Uchimura, for his open-minded and creative interaction throughout the course of this research at Kumamoto University.

The members of my (various) graduate committees provided me with helpful advice regarding the material in this thesis: Professor Shigeyasu Kawaji, Professor Yashito Kitasono, Professor Tsuyoshi Usakawa, and Professor Zhencheng Hu.

I also enjoyed the inspiring conversations with, and their assistance, the members of my laboratory, especially Kousuke Matsushima, Hiroyuki Tominaga, Tomoyuki Matsumura, Go Koutaki, and Francisco Lamosa.

Finally, I wish to thank the members of my family, my husband, Jiankun Wang, my 16 year old son, Honghang Wang, and my 7 year old daughter, Yingyu Wang, for their great support during the course of my Doctoral Studies.

Summary

The high-resolution IKONOS satellite has provided the world with its first source of commercially available satellite images, marking the beginning of a new era in earth observation. Geographic imagery is now widely used in Geographical Information Systems (GIS) applications worldwide. Decisions made using these GIS systems by national, regional and local governments, as well as commercial companies, affect millions of people, so it is critical that the information in the GIS is up to date. In most instances, aerial or satellite imagery provides the most up to date source of data available, helping to ensure accurate and reliable decisions. However, with technological advancements come new opportunities and challenges. The challenge now facing the geotechnology industry is twofold - how best to fully exploit high-resolution imagery and how to get access to it in a timely manner.

The higher the resolution of the imagery, the more easily that man-made objects can be identified. The human eye – the best image processor of all – can quickly detect and identify these objects. Therefore, if the application is one that simply requires an operator to identify objects and manually add them into the GIS database, then the imagery is already making a positive difference. It is adding a new data

source for the GIS Manager to use.

The high-resolution IKONOS satellite can provide accurate and up-to-date information for the extraction and maintenance of road databases. Automated extraction of road information provides the means for readily creating, maintaining, and updating the transportation network databases used in traffic management and automated vehicle navigation. These data are frequently used in making critical decisions in areas such as emergency response, evacuation, or incident management. Updating road map information is a time-consuming operation when performed manually. Due to the need for efficient acquisition and update of data used in GIS, the automatic extraction of road networks from high-resolution satellite images has been a major topic of interest for over twenty years. Automatic extraction methods hold potential for reducing database development time and cost. When factors such as image resolution, degradation of image quality, obstructions, and presence of linear but non-road features are taken into consideration, the task of road identification becomes an overwhelmingly complex task. The importance of the task and expected advantages of its automation provide a strong motivation for research in automatic interpretation of satellite and aerial images.

Chapter 2 summarizes the experience of previous road extraction efforts. The resolution of the images has a strong influence on the representation of roads and other features, which influences the methods of road recognition. Different approaches used for road recognition in images of low- and high-resolution are discussed. Another important factor for road extraction is the degree of interaction between the extraction algorithms and human operators.

Chapter 3 presents a semi-automated approach, based on a heuristic fuzzy logic system, to extracting the main suburban roads in IKONOS images. It describes how a system of different high-resolution satellite images, acquired under different conditions, can be used to detect and recognize road segments using fuzzy reasoning approaches. From the obtained extraction results, we find that the boundaries of forests and large area fields are clearly distinguished from roads, and only the main road candidate regions taken from the images were obtained. The results show that

the approach is valid for roads extracted from high-resolution satellite images.

In Chapter 4, another semi-automatic method for road network extraction from high-resolution satellite images is also developed. In this approach, the first step is detecting the seed points in candidate road regions using a Kohonen-type self-organizing map (SOM). The second step is searching for connected points in the direction and candidate domain of a road using a road tracking approach. With the goal of evaluating the method's potential in extracting road networks, many simulations were carried out using different types of images: small size (300×300 pixels) country road, overpass road images, middle size (500×500 pixels) country road images and large size (1000×1000 pixels) country road images. At the same time, for the purpose of choosing a learning rate factor $\alpha(t)$, many kinds of different simulations were also carried out using different types of images. The optimized initial learning rate factor α_0 is obtained, and good convergence onto the road candidate domain is shown. The results of the road network extraction can be used for Intelligent Transport Systems (ITS) applications.

Road network extraction has been one of the most challenging research topics in GIS. In addition to the afore-mentioned problems, special attention must be devoted to the flexibility of the heuristic system and road network extraction in urban areas. For example, there are perturbations, such as the shadows of buildings. The algorithm outlined in this thesis is merely a first step in this direction. Many further steps will have to be taken to develop the algorithm more fully. Therefore, future work will focus on the automatic road network completion methodologies. In addition, better road network descriptions can be achieved by performing more tests using different types of imagery.

Contents

Chapter 1
Preface

1.1 Satellite and Aerial Imagery

The recent generations of high-resolution satellite imaging systems, such as IKONOS, QuickBird and SPOT5 open a new era of earth observation and digital mapping. They provide not only for high-resolution and multi-spectral data, but also the capability for stereo mapping. The related sensors are all using linear array charge-coupled device (CCD) technology for image acquisition. The processing of high-resolution satellite images also provides an opportunity for algorithm redesign, thus opening the possibility of reconsidering and improving many photogrammetric processing components, such as image enhancement, multi-channel color processing, rectification, triangulation, orthophoto generation, Digital Terrain Model (DTM) generation, and feature extraction.

For example, since its launch in September 1999, Space Imaging's IKONOS Satellite (shown in Fig. 1.1) has provided a reliable stream of image data that has

become the standard for commercial high-resolution satellite data products. IKONOS produces 1-meter resolution black-and-white (panchromatic) and 4-meter resolution multi-spectral (red, blue, green, near infrared) images that can be combined in a variety of ways to accommodate a wide range of high-resolution imaging applications. High- resolution imagery will add an entirely new level of geographic knowledge and detail to the intelligent maps that we create from images.

Figure 1.1 Space imaging's IKONOS satellite

How are satellite image data and aerial image data different? The pixel size of IKONOS images corresponds to the information content of aerial images with a scale of 1:80,000; QuickBird images even correspond to aerial images at a scale of 1:50,000. Such images are used for topographic mapping up to a map scale 1: 10,000 and 1:60,000, respectively. Orthoimages should have 8 pixels /mm or more if the pixel structure is not to be visible. Based on this rule, IKONOS images can be used for generating Orthoimages with a scale of 1:8,000 and QuickBird images for orthoimages at a scale of 1:5,000. With this specification, there is a direct competition between very high-resolution satellite images and aerial photographs. In addition to the issue of information content, the geometric accuracy is important. This is depending upon the exact identification of the objects in the images and the image geometry itself together with an appropriate mathematical model.

An increasingly common source of image data is existing aerial survey photograph.

With the massive improvement in scanning technology and orthophoto production software, these old photo archives can be readily made available to Geographical Information Systems (GIS) users. Recently, the issue of digital versus analog is also being addressed. Old aerial photographs need to be scanned to convert them to a digital format. New digital aerial cameras get around this step, directly providing high quality aerial imagery at any user defined resolution. Depending upon the application and the level of accuracy needed, cameras ranging in price from a few hundreds to millions of dollars can be used. The drop in price and increased availability of Global Positioning System (GPS) units is also aiding the growth in the use of low cost digital cameras for GIS applications. Attached to remotely controlled aircraft or helicopters, they can provide very high-resolution, targeted aerial surveys for specific applications. Obviously, digital aerial cameras have a number of substantial advantages. At the same nominal image scale, the aerial images cover a larger ground area and, together with their superior geometrical resolution, contain much more information. Therefore, while both systems might complement each other in some applications, space images and images from digital aerial cameras are competing with each other in a more direct fashion.

Geographic imagery is now widely used in GIS applications worldwide. Decisions made using these GIS systems by national, regional, and local governments, as well as commercial companies, affect millions of people, so it is critical that the information in the GIS is up to date. In most instances, aerial or satellite imagery provides the most up to date source of data available, helping to ensure accurate and reliable decisions. However, with technological advancements come new opportunities and challenges. The challenge now facing the geotechnology industry is twofold - how best to fully exploit high-resolution imagery and how to get access to it in a timely manner.

The higher the resolution of the imagery, the more easily that man made objects can be identified. The human eye – the best image processor of all – can quickly detect and identify these features. Therefore, if the application is one that simply requires an operator to identify objects and manually add them into the GIS database,

then the imagery is making a positive difference. It is adding a new data source for the GIS Manager to use.

However, if the imagery requires information to be extracted from it in an automated or semi-automated fashion (for example, land cover classification, road network extraction), it is a different matter. If the techniques that were developed for earlier lower resolution satellite imaging are used on the high-resolution imagery, (such as maximum likelihood classification), the results can actually have a negative impact. Whilst lower resolution imagery isn't affected greatly by artifacts such as shadows, high-resolution data can be. Lower resolution data "smoothes" out variations across ranges of individual pixels, allowing statistical processing to create effective land cover maps. Higher resolution data doesn't do this – individual pixels can represent individual objects like manhole covers, puddles and bushes – and contiguous pixels in an image can vary dramatically, creating very mixed or "confused" classification results. There is also the issue of linear feature extraction. Lines of communication on a lower resolution image (such as roads) can be identified and extracted as a single line. However, on a high-resolution image, a road comprises the road markings, the road itself, the kerb (and its shadow) and the pavement (or sidewalk). A different method of feature extraction is therefore needed. Figure 1.2 shows the range and variety of information contained in a high-resolution image and the problems caused by shadows, overhanging trees and parked cars.

In addition, it's not just the spatial resolution that can affect the usage of the imagery. With 11 bit imaging becoming available, the ability of the GIS to work with high spectral content imaging becomes key. 11 bit data means that up to 2048 levels of gray can be stored and viewed. If the software being used to view the imagery assumes it is 8 bit (256 levels), then it will either a) display only the information below the 255 level (creating either a black or very poor image) or b) try to compress the 2048 levels into 256, also reducing the quality of the displayed image considerably. Having 2048 levels allows more information in shadowy areas to be extracted as well as enabling more precise spectral signatures to be defined, thus aiding in feature identification. However, without the correct software, this added

"bonus" can easily turn into a problem.

Figure 1.2 An example of the complex information contained in high-resolution imagery

Satellite or aerial imagery, by its very nature, is distorted in its raw form, either because of the movement of the sensor, the movement of the lens system, or terrain height variations of the area being imaged. Happily there are many software packages available which can correct for these distortions and make the images GIS ready. In some instances, where images are being used to look for relative changes or to identify simple features, geometric transformations are used. The users can use Image Analysis to simply warp the images to fit their vector database. The images can then be used to update the attribute information in the vector database. In this instance, it is the relative accuracy that is important.

For more precise mapping applications, corrections for the look angle and variations in terrain height are critical; in these instances, rigorous photogrammetric processes must be used.

Overall, the users, dependent upon their own application, must choose the correct level of accuracy. In general, most data providers now offer highly accurate, ortho rectified data as standard, with simple geometric rectification products becoming less common.

Whilst orthophotos can be used to digitize 2D features relatively easily, the ability to get accurate 3D data from imagery has been more difficult. If a very high accuracy Digital Elevation Model (DEM) is already available, then 3D coordinates can be applied to the 2D features captured from the orthophoto. It is rare, however, that such a DEM exists.

1.2 Intelligent Transport Systems

As mentioned above, high-resolution imagery from both satellite and aerial sources provides a challenge to the user community. While the human eye and brain can easily identify features objects in an image, the computer has a difficult time. If we cannot automate this process, then we will most certainly lose out on some of the major economic benefits of the imagery.

The brain can make a decision on an image very quickly by understand and using context. If we see grassland in the center of an urban development, we can easily decide that it is a park, as opposed to agricultural land. To make this decision we are using knowledge and experience as the basis of our expertise. Computer based expert systems are beginning to emerge that mimic this process. Actually, the computer can work well if it uses rule or knowledge-based processing, just as the human brain does.

For many years, expert systems have been used successfully for medical diagnoses

and various information technology (IT) applications but only recently they have been applied successfully to GIS applications.

Statistical image processing routines, such as maximum likelihood and Iterative Self-Organizing Data (ISODATA) classifiers, work extremely well at performing pixel-by-pixel analyses of images to identify land-cover types by common spectral signature. Expert-system technology takes the classification concept a giant step further by analyzing and identifying features based on spatial relationships with other features and their context within an image.

Expert systems contain sets of decision rules that examine spatial relationships and image context. These rules are structured like tree branches with questions, conditions and hypotheses that must be answered or satisfied. Each answer directs the analysis down a different branch to another set of questions.

In terms of satellite images, the knowledge base identifies features by applying questions and hypotheses that examine pixel values, relationships with other features and spatial conditions, such as altitude, slope, aspect and shape. Most importantly, the knowledge base can accept inputs of multiple data types, such as digital elevation models, digital maps, GIS layers and other pre-processed thematic satellite images, to make the necessary assessments.

In Intelligent Transport Systems (ITS), for example, an expert classification might identify and create ideal traffic conditions. That is to say, ITS system will reduce traffic accidents and congestion while saving energy and protecting the environment. ITS requires not only the roads to be intelligent but also the various modes of transportation, such as rail, aviation, and shipping, to cooperate with each other. ITS is a national project that will even change the structure of society. It has great potential to create new industries and markets. The objectives are to enhance traffic safety, to smooth traffic flow, and to improve the environment (shown in Fig. 1.3).

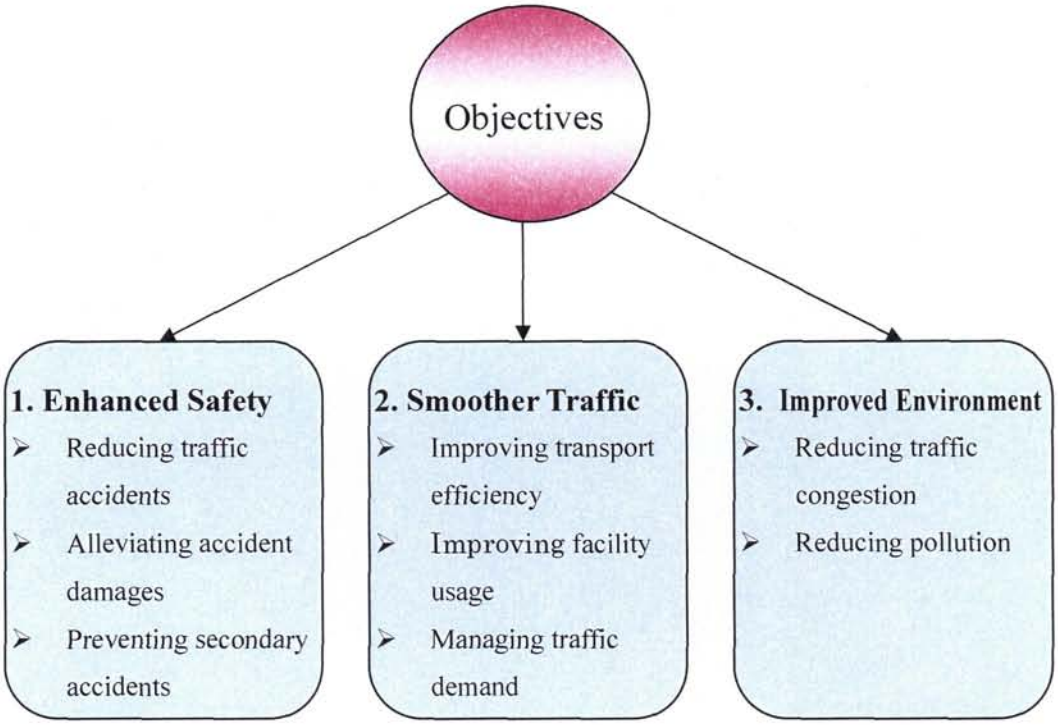


Figure 1.3 The objectives of ITS

ITS offers a fundamental solution to various issues concerning transportation, including safety, congestion and the environment. ITS addresses these issues using the most advanced communications and control technologies. ITS systems receive and transmit information on humans, roads and automobiles.

Road information is needed for a variety of applications such as the description of the basic topographic infrastructure, needed in transportation planning, traffic and fleet management and optimization. It also constitutes an important component of 3D City Models. National Mapping Agencies (NMAs) and National Road Administrations (NRAs) increasingly plan to update their GIS in shorter cycles. Many NMAs plan to add 3D information, especially for roads and buildings. Customers have increasing demands regarding the level of accuracy and modeling completeness, and often request additional attributes, e.g., the number of lanes.

While road extraction has been sometimes performed by digitizing maps, road update and refinement can mainly be performed only from satellite and aerial imagery, while some ground-based methods including mobile GPS are also important.

Mapping using high-resolution satellite images will be a hot issue in the field of computer vision, because most of the major interesting man-made objects can be easily identifiable. One of the automatically identifiable features of importance may be a road. However, with technological advances come new opportunities and challenges. Developments in technology and science constantly increase the number of solutions to problems. The interpretation of satellite images is a problem that now has the potential of being solved automatically.

1.3 Objectives

The high-resolution IKONOS satellite has provided the world with the first source of commercially available satellite images, marking the beginning of a new era in earth observation. It can provide accurate and up-to-date information for the extraction and maintenance of road databases. Automated extraction of road information provides the means for readily creating, maintaining, and updating the transportation network databases used in traffic management and automated vehicle navigation. These data are frequently used in making critical decisions in areas such as emergency response, evacuation, or incident management. Updating road map information is a time-consuming operation when performed manually. Due to the need for efficient acquisition and update of data used in Geographical Information Systems (GIS), the automatic extraction of road networks from high-resolution satellite images has been a major topic of interest for over twenty years. Automatic extraction methods hold potential for reducing database development time and cost. When factors such as image resolution, degradation of image quality, obstructions, and presence of linear

but non-road features are taken into consideration, the task of road identification becomes an overwhelmingly complex task. The importance of the task and the expected advantages of its automation provide a strong motivation for research in automatic interpretation of satellite and aerial images.

Roads are important features of infrastructure. The characteristic properties of roads, such as significant length, and network topology, can seldom be found for other features in satellite and aerial images. Thus, based on these properties, algorithms for road extraction can be constructed. At the same time the difference in properties of roads and other features makes the strategies and techniques used for their recognition different. This motivates the treatment of road extraction as a separate problem. In order to investigate the afore-mentioned problem, research on road network extraction from high-resolution satellite images was carried out. Initially, this work focuses on the extraction of a candidate road domain using fuzzy logic techniques. Subsequently, the work concentrates on road network extraction using self-organizing map (SOM) neural network and road tracking.

1.4 Thesis Overview

In this thesis a study of Geographical Information Systems (GIS) with high-resolution satellite images is presented. This thesis is structured as described below. First, an initial introduction to the subject and the objectives of the thesis are presented. Then, the experience of previous road extraction work is discussed in Chapter 2. Chapter 3 presents a semi-automated approach, based on a fuzzy logic system, to extracting the main suburban roads in IKONOS images, and shows how to implement structure extraction algorithms using fuzzy reasoning approaches. Later, a semi-automatic method for road network extraction from high-resolution satellite images is also developed in Chapter 4. The first step is detecting the seed points in candidate road regions using a Kohonen-type self-organizing map (SOM). The

second step is searching for connected points in the direction and candidate domain of a road using a road tracking approach. In Chapter 5, conclusions are presented and give possible directions for continuation of the topic.

Chapter 2

Survey on Road Extraction

Updating road maps (or other types of maps) is a time-consuming operation when performed manually. Unfortunately, no existing methods allow complete and reliable automation of this process. Due to its importance, much effort is devoted to finding a solution to this problem. During the last two decades many approaches for automatic road extraction have appeared in the literature. Most of them are quite different due to the differences in their goals, the available information, and their assumptions about roads. To classify and analyze these approaches several criteria can be chosen. However, two factors seem to have the greatest influence on the nature of the road extraction: the spatial resolution of the images and the requirement of an operator to guide the extraction process.

The resolution of the images has a strong influence on the representation of both roads and other objects. Digital images are surjective mappings of the infinite world into a finite number of image pixels. This implies that many details recognizable on the images of high resolution become unclear and finally disappear when the image resolution gets lower. Thus, road edges, vehicles, and even road markings, which can clearly be seen in images with ground resolution of 0.2m per pixel disappear when

the image resolution becomes lower than 2m per pixel. Whereas roads in high-resolution images can have a great number of variations, the road represented in low-resolution images of often appear as line structures which are often barely distinguishable from other linear features in images, buildings. This influences the methods needed for road recognition in each type of images. Some different approaches used for road recognition in images of low- and high-resolution are presented and discussed in Sections 2.2.

Another important factor for road extraction is about interaction between the extraction algorithms and human operators. Algorithms that require input from a human operator are called semi-automatic. In contrast to the fully-automatic methods they demand a number of seed points which are usually chosen by the operator in an interactive fashion. Given such seed points the semi-automatic algorithm connects them by a path which is most likely a road. The given seed points strongly reduce the problem-space of semi-automatic algorithms. Unlikely fully-automatic algorithms they do not have to consider the problem of road finding and often are implemented as so called road trackers. A comparison of semi- and fully-automatic methods is given in Sections 2.2.

2.1 Road Characteristics

In order to correctly detect roads, we must understand how the physical characteristics of a road influence its visual characteristics. These visual characteristics are used to identify roads in an image.

The physical characteristics are:

(i) The surface of a road is firm and smooth. Roads are built using materials such as concrete or asphalt; consequently they have a spectral signature corresponding to these materials. Most of the time, there are road-signs and markings.

(ii) The steepness of roads has an upper bound, which is why, in mountainous regions, roads tend to be winding.

(iii) The width of a road is almost constant and has an upper bound. This upper bound is highly dependent in the importance of the road. For instance, highways are much wider than rural or suburban road.

(iv) The local curvature of roads has an upper bound, which likewise depends on the importance of the road. For instance, the maximal curvature is lower for a highway than for a rural road. It should be noted that roads are straight locally but not globally.

(v) Roads are built to link certain places together and neighboring roads are connected to form networks. The density of roads and connections depends on the surroundings context. It is higher in urban areas than in rural areas.

We can divide these properties into four types. Surface characteristics are spectral properties, while steepness, width and curvature are geometric. Links and networking are topological properties and the fact that the type of road influences the maximum width and maximum curvature is a contextual property.

Of these characteristics, many can be determined from satellite or aerial images. In low-resolution images, roads appear as lines. In high-resolution images, roads appear as elongated regions with parallel borders. Generally, the intensity within the road is not constant because the presence of cars, shadowing by trees or building, and marking lines can produce rapid changes in image intensity. Steepness is not easily measurable unless we have access to stereoscopic pairs or other 3D information. Width is easier to determine from high-resolution images. Curvature and topological information are directly visible in low- and high- resolution images; however, this information may be distorted by the image acquisition system.

2.2 Road Extraction Methods

The classification of surveys and road extraction techniques is very difficult, given the variety of existing proposals in the literature. However, we can select the principal factors in order to approach this task. These factors considered in this classification are the following: the preset objective, the extraction technique applied, and the type of sensor utilized. The classification of methods and works on road extraction is described schematically as follows.

1. ***According to the preset objective:***
   - General road extraction methods
   - Road network reconstruction methods
   - General segmentation methods
   - Vectorization methods
   - Optimization methods
     - Neural networks
     - Genetic algorithms
     - Other optimization methods
   - Evaluation methods
   - Other objectives

2. ***According to the extraction technique applied:***

   1) Low and mid level methods
   - Road tracking methods
   - Morphology and filtrate
   - Dynamic programming and snakes
   - Segmentation and classification
   - Multi-scale and multi-resolution
   - Stereoscopic analysis
   - Multi-temporal analysis
   - Other techniques

2) Mid and high level methods

- Knowledge representation and fuzzy modeling
  - ✧ Logic systems
  - ✧ Rules based systems
  - ✧ Blackboard systems
  - ✧ Frames based systems
  - ✧ Semantic networks
  - ✧ Fuzzy logic based systems
- Other methods of spatial reasoning

3. ***According to the type of sensor utilized:***

- Monochromatic imagery
- Infrared band
- Color imagery (RGB)
- Multi- and hyper-spectral imagery (HYDICE)
- Synthetic aperture radar imagery (SAR)
- Laser imagery (LIDAR, . . .)
- Other type of sensor

Evidently, in using this structure, the possibility exists that the same work is included in several categories of the classification. Next, classification according to the preset objective and classification according to the extraction technique applied are used. The classification according to the type of sensor is not explicitly used since this is implicitly included in the two other classifications.

2.2.1 Road Extraction in Low Resolution

In low-resolution images, as noted above, roads appear as linear structures and often are indistinguishable from other linear features in images. Due to the suppression of

small details, roads are represented as lines with more or less homogeneous surface.

At first sight such simplification can be considered to be disadvantageous. In such a representation, road can easily be confused with other linear structures in the images [Fischler *et al.*, 1981]. Furthermore, the automatic algorithms will have difficulties in tracking roads through occluded and shadowed image parts since the latter do not contain enough evidence for roads. Moreover, images of low-resolution images do not contain some useful information about roads, such as their width, which is sometimes a partial goal of road extraction.

Nevertheless road extraction at low-resolution has some advantages. As argued in [Mayer *et al.*, 1996] the suppression of small road details can be interpreted as a sort of abstraction: abstraction occurs by means of the elimination of substructure (annihilation) and the merging of different image parts. By this means, abstraction helps road recognition since fewer special cases of roads have to be considered: the low-resolution road extraction problem largely reduces to the general problem of line extraction [Fischler *et al.*, 1981]. On the other hand, as noted in [McKeown *et al.*, 1988] although more information is available in higher resolution imagery, more processing is required to take full advantage of it. Thus road extraction from high-resolution images will generally require more complex algorithms and more time.

Approaches based on dynamic programming appeared in [Grun *et al.*, 1995] and [Merlet *et al.*, 1996]. They define a generic road model using contrast to the adjacent terrain, the homogeneity of the surface, connectivity, low curvature and constant width. The model is embedded into a dynamic programming algorithm and is controlled by stating constraints, e.g. road curvature. A different approach to line and road extraction is presented in [Steger, 1996]. It is argued that the methods described above use purely local criteria, e.g., local grey value differences. This implies the generation of a lot of false hypotheses for line points and computationally expensive schemes that have to be used to select salient lines in the images. Instead a differential geometric approach that can determine the positions of lines to sub-pixel accuracy is proposed.

As shown in [Grun *et al.*, 1995] and [Merlet *et al.*, 1996], their algorithms successfully follow the roads through forest and urban areas. However, this is possible only due to the semi-automatic nature of these algorithms. Given the knowledge about the presence of roads in form of seed points, the dynamic programming algorithms are forced to track a road between the given points. If the points do not belong to the roads or are located on different roads, the algorithms will track a false road between them. Furthermore, the algorithms do not give any qualitative information that can be used to accept or discard the detected road hypotheses. All they guarantee is a connecting path through the image which best fits all the points to their model. Thus, dynamic programming algorithms are potentially semi-automatic and are not suitable, their original form, for fully- automatic road extraction.

2.2.2 Road Extraction in High-Resolution

Similar to the situation with low-resolution imagery, road extraction with high-resolution images has both advantages and disadvantages. On the one hand, the presence of many characteristic road details provides more evidence for the road. Analyzing these details, one will, in general, find more reasons for accepting or rejecting a given hypothesis.

On the other hand, the details of the roads can greatly vary: roadsides can change the contrast; cars of different sizes can be present; the road surface can change material and width and so forth. Thus, the larger number of details will increase the complexity of road extraction algorithms. In order to keep the algorithm complexity low, the set of considered road details has to be chosen carefully.

Besides the choice of characteristic road details, there is a problem of weighting their contributions.

2.2.3 Road Extraction in Multi-Resolution

None of the methods presented above offers a reliable and fully-automatic road extraction algorithm. Low-resolution methods are able to locate many roads using simple road models but lack information and thus make mistakes. High-resolution methods have access too much more information and can investigate roads in great detail. However, since other features in the images are also represented with many details, the task of road finding requires complicated methods for the abstraction of these details.

In [Baumgarter *et al.*, 1997], the advantages of road extraction at both resolutions are combined in a multi-resolution approach. Road extraction is divided into two tasks: the detection of roads, and their verification and delineation. Road detection is made by line extraction at low-resolution. At high-resolution, parallel edges are extracted and grouped into rectangles whose surface is checked for homogeneity in the direction of the edges. Finally, the sets of near and collinear rectangles are connected into road segments. In order to verify these, the results of extraction of both resolutions are fused.

2.3 Outline of the Work

The difficulties of the previous work, discussed in mention above, clearly prove that road extraction is a highly non-trivial problem. This work aims to improve road extraction by solving some of its partial problem. A semi-automatically extracting the main suburban roads and road network are carried out using high-resolution satellite imagery. Initially, fuzzy logic is taken into consideration in extracting the main suburban roads, because fuzzy set theory and fuzzy logic offer us powerful tools to represent and process human knowledge in the form of fuzzy if-then rules. In addition, many difficulties in image processing arise because the data contains errors

and uncertainty. In these cases fuzzy logic can be applied. Subsequently, Self-Organizing Map (SOM) is taken into consideration in extracting road network, because the SOM can be viewed as a model of unsupervised (machine) learning, and as an adaptive knowledge representation scheme. The traditional knowledge representation formalisms - semantic networks, frame systems, predicate logic, to provide some examples, are static and the reference relations of the elements are determined by a human. Related to knowledge representation and learning, the cognitive and philosophical aspects are highly relevant. The detailed discussions are given in the following Chapter 3 and Chapter 4.

Chapter 3
Road Extraction

3.1 Introduction

Research on the extraction of man-made objects from aerial and satellite images has made significant progress in recent years. To perceive features in satellite images, it appears that humans are applying heuristic algorithms to understand such images. Although these algorithms may be "implemented" in the wetware of our visual systems, it is possible to characterize an equivalent process systematically. Fuzzy logic is a suitable framework for expressing the heuristic processes applied to the extraction of roads from satellite images.

Fuzzy logic is an alternative to the traditional notions of set membership and logic that has its origins in ancient Greek philosophy. It has applications at the leading edge of Artificial Intelligence. Fuzzy logic is a superset of conventional logic that has been extended to handle uncertainty in data. It was introduced by Dr. Lotfi A. Zadeh in the 1960s as a means to model the uncertainty of natural language [Zadeh,

1965]. Yet, despite its long-standing origins, it is a relatively new field, and, as such, leaves much room for development.

Fuzzy logic is useful in applications such as manufacturing, process control, and image processing, because of its ability to handle situations that traditional binary logic cannot adequately deal with. In many image processing applications, we have to use expert knowledge to overcome these difficulties (e.g. object recognition, scene analysis). Fuzzy set theory and fuzzy logic offer us powerful tools to represent and process human knowledge in the form of fuzzy if-then rules. In addition, many difficulties in image processing arise because the data contains errors and uncertainty In these cases fuzzy logic can be applied.

The cognitive activity of the brain, unlike the computational function of a binary computer, is based upon the relative grades of information acquired by the natural sensory system. Fortunately, since the methodology was first proposed by Zadeh [Zadeh, 1965], fuzzy logic has been applied to numerous problems and applications, and now incorporates the notion of inaccuracies and provides a framework through which non-deterministic problems can be addressed. Recently, a perception-based theory of probabilistic reasoning was presented by Zadeh [Zadeh, 2002]. Fuzzy logic has been also successfully adopted for feature detection and matching. Agouris et al. demonstrate the use of a fuzzy supervisor for the extraction and tracking of objects from aerial images [Agouris *et al.*, 1998]. Steger et al. present the use of a fuzzy approach to road network extraction by applying it to the segment-linking problem in low-resolution aerial images [Baumgartner *et al.*, 1999a], [Steger *et al.*, 1997]. Road extraction processes have many different forms according to the images' spatial resolution. In low-resolution imagery (i.e., ground resolution more than 2m per pixel) roads correspond to lines whereas in a resolution of 0.2m – 0.5m they can be described as elongated homogeneous areas. In low-resolution images, the global network structure of the road can be seen clearly and small disturbances like shadows and cars can be easily avoided. The operator makes use of line detection algorithms because roads are "line-like" structures [Fischler *et al.*, 1981]. The work in [Agouris *et al.*, 2001a] focuses on road extraction from small scale aerial images using

deformable contour models (snakes). A modification of the classic concept of snakes to perform automated road extraction is presented in [Agouris *et al.*, 2001b]. In high-resolution images, the accuracy of geometrical properties like width and curvature is much better. A multi-resolution approach for the automatic extraction of roads from digital aerial images is presented in [Baumgartner *et al.*, 1997a], [Laptev *et al.*, 2000]. They claim that if relations between roads and other objects (cars, buildings or trees) are neglected, reliable extraction is often difficult. So they use either GIS information or texture analysis to partition images into regions (e.g. urban, forest, rural). This approach to road detection is based on the extraction of edges in a high-resolution image and the extraction of lines in an image at reduced resolution.
Besides the classification of road extraction schemes into low- and high-resolution image extraction, the strategies used in addressing this problem can also be classified into two broad categories: semi- and fully-automatic. Semi-automatic approaches require an operator to interactively provide some information to control the extraction. According to Park and Kim, the road segments are extracted using template matching based on an adaptive least square matching algorithm [Park & Kim, 2001]. A semi-automated approach using elongated region analysis for 2D road extraction from high-resolution imagery is presented in [Doucette *et al.*, 2001]. In fully automatic schemes, knowledge sources such as GIS or geographical databases [Gibson, 2003], [Zhang & Baltsavias, 2002], heuristics [Hinz *et al.*, 2001], [Laptev *et al.*, 2000], [Steger *et al.*, 1997] and stochastic methods [Barzohar & Cooper, 1996] are used to initialize the road tracking. Also, in [Hu & Tao, 2003] a particular perceptual grouping strategy - hierarchical grouping- is proposed to automatically extract main-road centerlines from high-resolution satellite imagery.

Traditionally, the algorithms dealing with this topic are mostly semi-automatic, which means they need human operators to provide the starting points of the roads. In [Agouris *et al.*, 1998], the limitation was that it required the initial seed points and road width estimates for road extraction to be successful. Our work in this chapter focuses on semi-automatically extracting the main suburban roads in IKONOS images using a fuzzy logic system. This chapter proposes an algorithm which

analyzes high-resolution satellite images and extracts the main roads after first using some pre-processing operations. This algorithm is implemented in three-stages. First, the color profiles taken from the satellite image are analyzed to search for grayscale values similar to those resulting from the radiometric properties of the road. Next, the algorithm recognizes the potential geometric shapes of the road using a straight-line Hough transform. The road segments are determined extracted by means of fuzzy logic rules. Finally, successive image processing and analysis is able to exploit the fuzzy reasoning and to identify false positives.

3.2 Outline of Algorithm

3.2.1 Fuzzy Logic Algorithm

The following sets of linguistic values are considered in our application:

Quality A = {small, large},
Quality B = {low, high}, and
Result C = {fast, slow}.

The fuzzy membership values, μ_i, for Quality A and Quality B are determined as shown graphically in Fig. 1.

Each membership function is determined from the statistical properties of the corresponding input variable. The fuzzy rules that will produce the uncertainty values have the following form:

IF (Quality A is *small*) AND (Quality B is *low*), THEN Result C is *fast*.

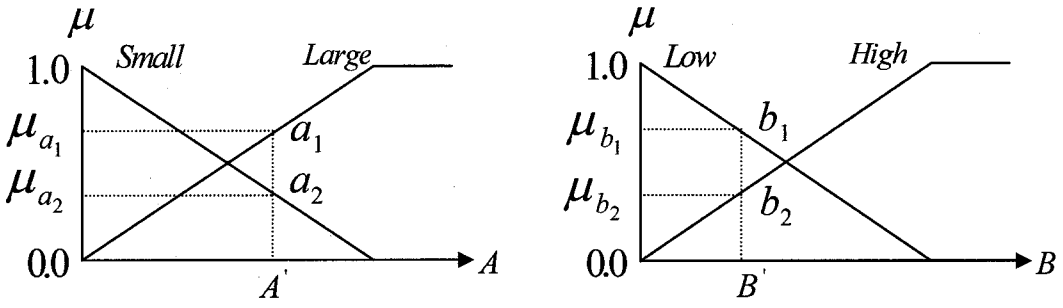
and are displayed in Table 3.1. Let us consider an example:

The rule states that *small* Quality A and *low* Quality B in the solution implies *fast* Result C.

For a given value of A, denoted A', there is both a *small* component a_2 and a *large* component a_1, which correspond to fuzzy memberships μ_{a_2} and μ_{a_1} for each of the *small* and *large* fuzzy sets respectively. Likewise for a given value of B, we obtain fuzzy memberships μ_{b_1} and μ_{b_2} for each of the *low* and *high* fuzzy sets respectively, in the case of B'.

Table 3.1 Fuzzy rules for quality evaluation

Quality A	Quality B	Result C
Small	Low	Fast
Small	High	Fast
Large	Low	Slow
Large	High	Slow



(a) Quality A (b) Quality B

Figure 3.1 Determining fuzzy membership values

The Result C is determined by Eq. (3.1),

$$z_0 = \frac{\sum z_i \cdot \mu_{C'}(z_i)}{\sum \mu_{C'}(z_i)} \tag{3.1}$$

Where $\mu_{C'}$ are the minima of the fuzzy subsets of z_i, the fuzzy member values for each of the two categories of Quality A and Quality B, and z_0 is the resulting σ. Equation (3.1) is a simplified version of the min-max centroid method of simplified fuzzy reasoning. Applying Eq. (3.1) to this example gives

$$\sigma = \frac{\min(\mu_{a_2}, \mu_{b_1}) \cdot F + \min(\mu_{a_2}, \mu_{b_2}) \cdot F + \min(\mu_{a_1}, \mu_{b_1}) \cdot S + \min(\mu_{a_1}, \mu_{b_2}) \cdot S}{\min(\mu_{a_2}, \mu_{b_1}) + \min(\mu_{a_2}, \mu_{b_2}) + \min(\mu_{a_1}, \mu_{b_1}) + \min(\mu_{a_1}, \mu_{b_2})} \tag{3.2}$$

which is a simultaneous version of the four rules in Table 1, producing a final value for Result C, where S and F are constants (determined empirically) for *Slow* C and *Fast* C respectively, and each term in the numerator corresponds to a rule in Table 3.1.

3.2.2 Algorithm of Road Extraction Using Fuzzy Logic

Effective road extraction methods must not only use radiometric and geometric properties, but must reflect structural constraints. Detected road segments appear not only as a local area, but also as part of larger structures that can also be characterized heuristically. So it is possible to express structural constraints heuristically in a systematic way. In order to extract features from satellite imagery, a feature extractor using fuzzy logic is shown in Eq. (3.3). Different feature categories correspond to different properties. These can be divided into geometric and radiometric properties and used together as follows:

IF (Radiometry) AND (Geometry) THEN Feature (3.3)

Radiometric properties include rules on absolute gray values, hue values, saturation values, and contrast information. At the same time geometric properties may include linearity, parallelism, and width information, among other attributes. The fuzzy rules are applied to analyze these properties for specific objects such as, for example, a road. The implementation of the rule can be described as shown in Eq. (3.4):

IF (Road radiometric properties) AND (Road geometric properties)
THEN Road (3.4)

3.3 Road Extraction Approach

This chapter describes a system for detection and recognition of road segments from different high-resolution satellite images taken under different conditions. A fuzzy filtering of each input RGB image is carried out. Subsequent processing then considers the radiometric properties using fuzzy logic; after this, the geometric features are detected by applying a Hough transform. Finally, integration of the radiometric and geometric properties is implemented, using fuzzy logic for the elimination of false positives. This results in the extraction of only the main suburban road segments.

In order to extract roads from a satellite image, it is necessary to have an explicit idea about the concept "road". For the proposed approach, the model comprises explicit knowledge about radiometry (reflectance properties) and geometry (road width, parallelism of road both sides). The models are described below the content. The road extraction approach diagram is shown in Fig.3.2.

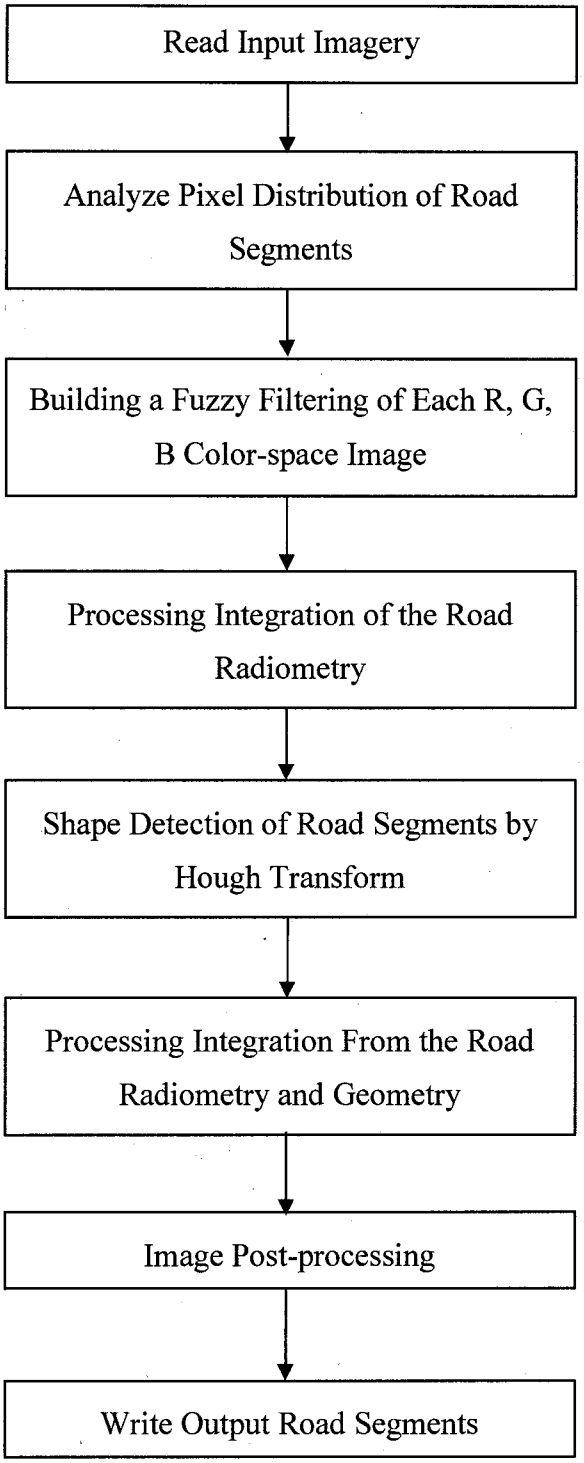


Figure 3.2 The road extraction approach diagram

3.3.1 Road Model

In order to propose the technique of extracting the road by fuzzy reasoning, the reflective color (Gray level) of the road, the width of the road, and parallel lines of road both sides can be considered the membership function. Therefore, we build the geometric-stochastic road model based on the following assumptions:

(i) The gray level of the road local average only changes slowly.

(ii) The road width change is small. The road both-sides borderline becomes parallel lines.

(iii) The road direction changes are likely to be slow. The road neighborhood becomes a straight line approximately.

(iv) Roads are unlikely to be short.

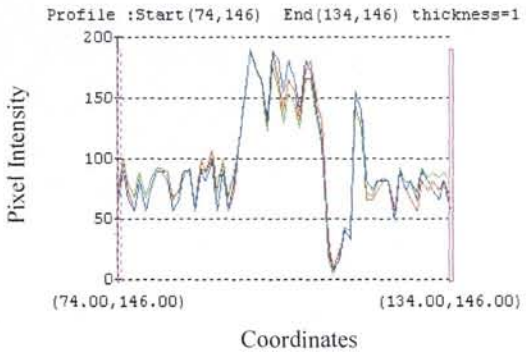
3.3.2 Image Pre-Processing

The road detection problem requires an adequate characterization of a road. Specifically, a road intuitively corresponds to suitable reflectance properties in satellite imagery, leading to attempts to detect roads by locating appropriate local gray levels. In this chapter different high-resolution IKONOS satellite images (300×300 pixels, 500×500 pixels, 1000×1000 pixels, 1m/pixel) acquired under different conditions were processed using fuzzy reasoning to identify road candidates.

The pixels of a color image using the RGB model have three components: Red, Green, and Blue, so that an original satellite image (shown in Figs. 3.3 (a), 3.4 (a)) can be divided into the red, green and blue color-components. In fact they are shown as "gray-level" color images, as they are composed of a single element (shown in Figs. 3.5, 3.6).

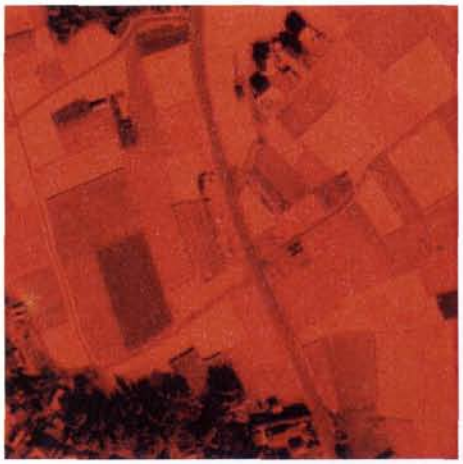
(a) Original image of example 1 (b)

Figure 3.3 (a) Original image (300×300 pixels, 1m/pixel) (b) Road pixel intensity

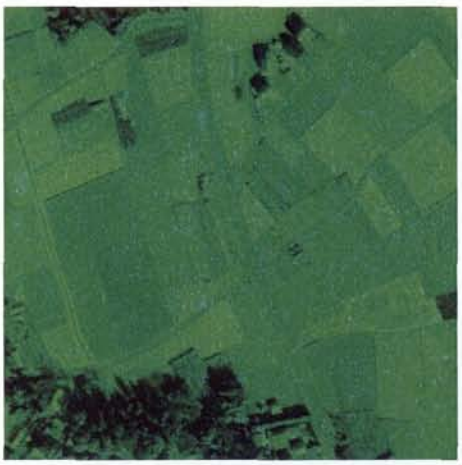


(a) Original Image of example 2 (b)

Figure 3.4 (a) Original image (300×300 pixels, 1m/pixel) (b) Road pixel intensity



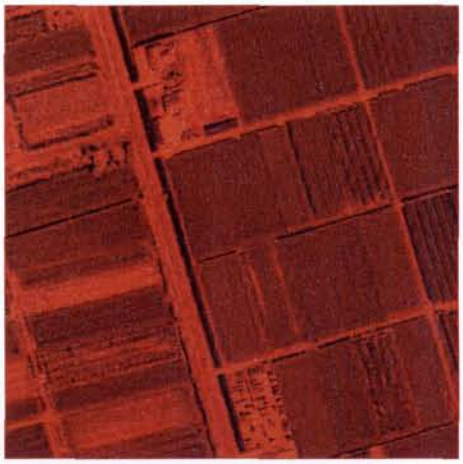
(a) Red color-component



(b) Green color-component

(c) Blue color-component

Figure 3.5 Image of example 1



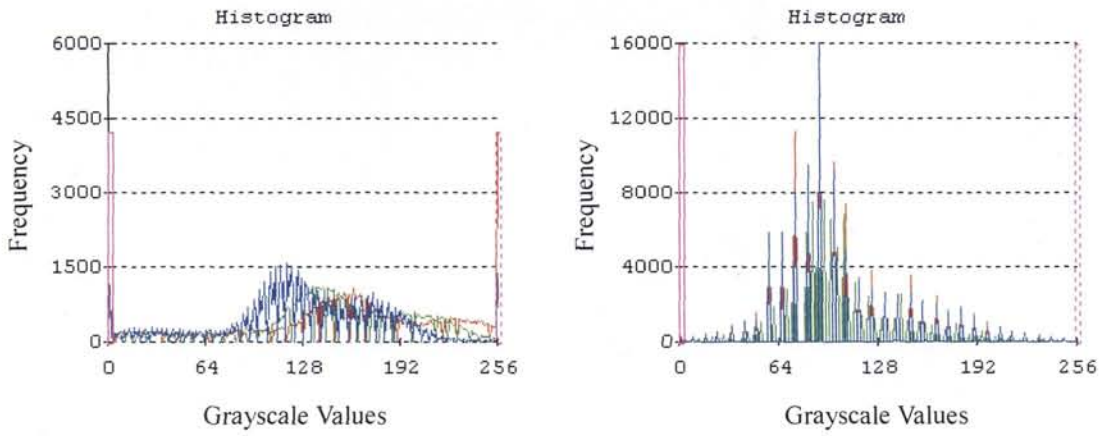
(a) Red color-component

(b) Green color-component

(c) Blue color-component

Figure 3.6 Image of example 2

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values. Histograms can be also taken of color images, which either individual histogram of red, green and blue channels can be produced, with the horizontal axes representing the red, blue and green channels, and brightness at each point representing the pixel count (shown in Fig. 3.7). Analysis of the histogram gives useful information about image contrast in image processing. The pixel values of an image are manipulated in such a way as to distribute the pixel values in an approximately even distribution over the color range of the image. An important class of point operations is based upon the manipulation of an image histogram or a region histogram. The pixel distribution of road segment has been seen in the middle gray level approximately.



(a) Image of example 1 (b) Image of example 2

Figure 3.7 Histogram

A pixel intensity profile of a road segment shows that the road pixels are in the mid-gray level range approximately, allowing it to be neatly segmented from its background (Figs. 3.3 (b), 3.4 (b)).

Using the grayscale values obtained from the profiles, a probability histogram was constructed to assign probabilities for a given pixel belonging to a road. A triangular probability histogram, as shown in Fig. 3.8, was constructed.

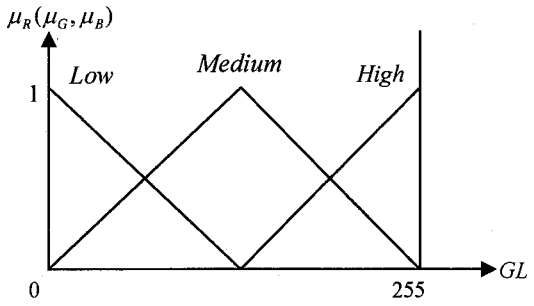


Figure 3.8 Input membership function for each gray level

Table 3.2 Fuzzy rules for the road possibility of R (or G, B) image

Input Gray Level of R (or G, B) Image μ_R (or μ_G, μ_B)	Road Possibility for R (or G, B) Gray Levels μ_{RG} (or μ_{GG}, μ_{BG})
Low	Low
Medium	High
High	Low

3.3.3 Image Processing of Radiometric Data Using Fuzzy Logic

Generation of fuzzy rules from R, G, and B images

In this stage, the fuzzy rules used for the extracting of features, based on the color reflectance of the road surface are generated as described below.

From the satellite images shown in Figs. 3.3 (a), 3.4 (a), a fuzzy filtering of each R, G, B image is carried out. The input membership functions for each color-band are shown in Fig. 3.8. Three classes have been chosen, according to the grayscale values of the pixels: low, medium and high. All the membership functions are considered triangular functions. From the previous color profile analysis (shown in Figs. 3.3 (b), 3.4 (b)), it can be seen that the pixels for the road segments generally exist in the central region of the color profile. In the fuzzy rules of Table 3.2, it is assumed that the road spectral reflectance is high in comparison with the medium grayscale values of the each image. Table 3.2 shows the fuzzy rules of each input R, G and B image.

This fuzzy reasoning logic was then implemented in a program.

In this program, the processed image was searched to identify pixels that match road extraction rules for each R, G and B image. The candidate pixels are considered as inputs to a filter function. If the output of the filter is between the specified limits, then the pixel value remains 0 (black), otherwise it becomes 255 (white). The application of the fuzzy reasoning to the original images of Figs. 3.3 (a), and 3.4 (a) yields the reflectance values of the R, G and B components respectively. The processing results of Figs. 3.5 (a)-(c), and 3.6 (a)-(c) by the fuzzy reasoning to the road surface reflective colors each R, G and B value are shown in Figs. 3.9 (a) - (c), and Figs. 3.10 (a) - (c) respectively. Although this method does find the valid road segments, it also is subject to false positives. To eliminate the incorrectly found road segments, it was necessary to introduce a second step.

(a) Processing result for R image

(b) Processing result for G image

(c) Processing result for B image

(d) Extraction result

Figure 3.9 Results of example 1 based on fuzzy reasoning for radiometry

(a) Processing result for R image

(b) Processing result for G image

(c) Processing result for B image

(d) Extraction result

Figure 3.10 Results of example 2 based on fuzzy reasoning for radiometry

Processing using road radiometry

In this stage, fuzzy reasoning was used to extract features from the previous processing results having similar color characteristics to a theoretical road surface.

Table 3.3 Fuzzy rules for the discrimination of road surface reflective colors

Road Possibility for R Gray Levels μ_{RG}	**Road Possibility for G Gray Levels μ_{GG}**	**Road Possibility for B Gray Levels μ_{BG}**	**Road Possibility for radiometry μ_{radio}**
Low	Low	Low	Low
Low	Low	High	Low
Low	High	Low	Low
Low	High	High	Low
High	Low	Low	Low
High	Low	High	Low
High	High	Low	Low
High	High	High	High

The fuzzy membership function was developed by specifying that

(i) When the grayscale values of all three components, (R Gray Levels μ_{RG}, G Gray Levels μ_{GG}, and B Gray Levels μ_{BG}), are high at the same time, the road spectral reflectance was considered high.

(ii) Conversely, when the grayscale values of one or more of the components are low, the total spectral reflectance was considered low.

Using these rules, the fuzzy rules shown in Table 3.3 were obtained. The result of the fuzzy reasoning of the road surface color reflectance is shown in Figs. 3.9 (d), 3.10 (d).

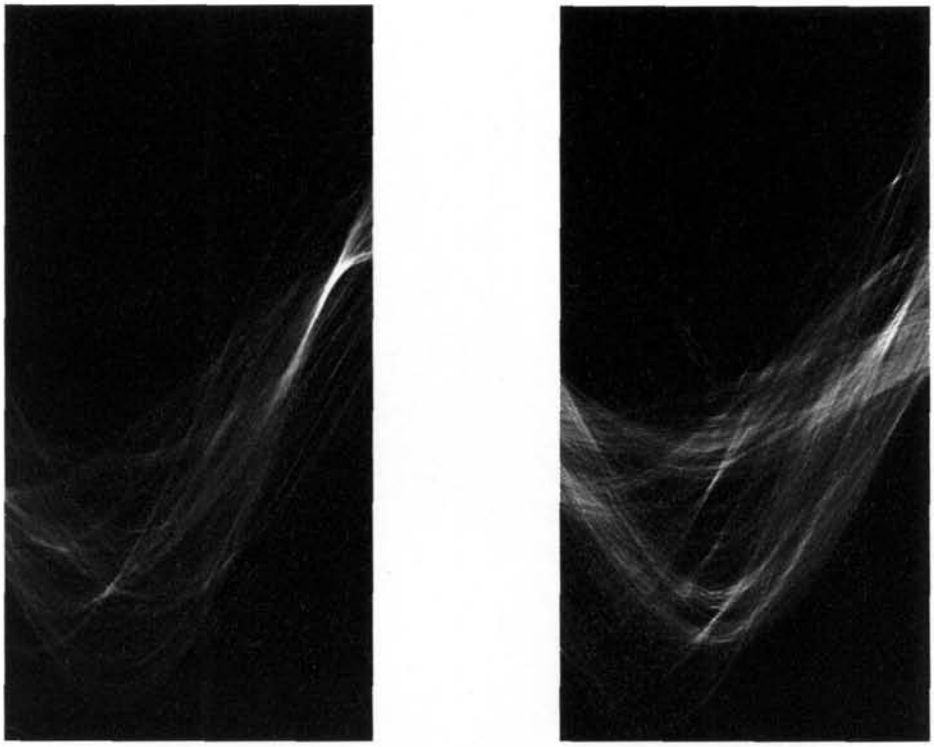
3.3.4 Image Processing Using Fuzzy Logic from Geometry

Shape detection of road segments by Hough Transform

From the previous results, the potential geometric shapes of the road are detected with a Hough transform. The main advantage of the Hough transform technique is that it is tolerant of gaps in linear features and is relatively unaffected by image noise. It is a method widely used for the detection of lines. The transformation is based on the collection of evidence regarding the existence of lines in image space using a line parameter space. By considering road pixels as evidence, each given pixel (x_0, y_0) is considered as belonging to a set of lines that may pass through it, hence each pixel provides evidence to the existence of a set of lines. These set of lines are presented in a parameter space (ρ, θ) ("Hough space" shown in Figs. 3.11 (a), (b)) by a sinusoid-like curve of the form:

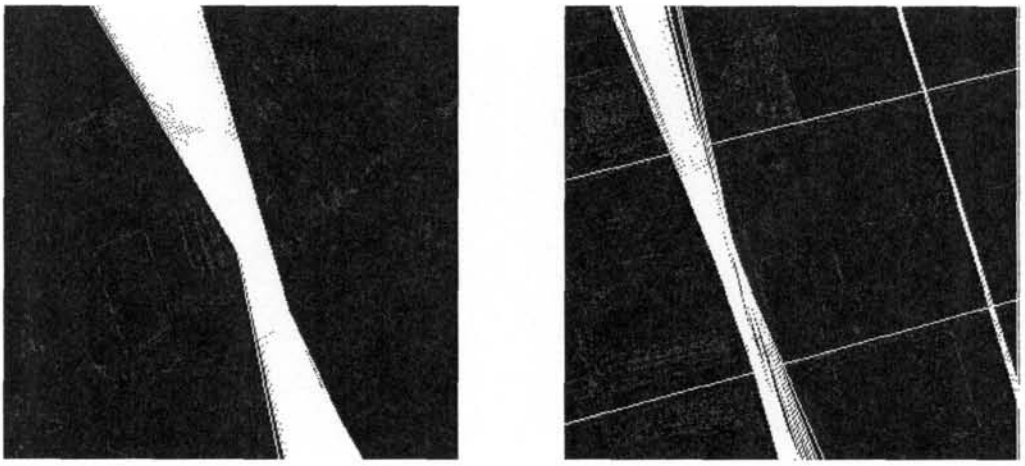
$$\rho = x_0 \cos\theta + y_0 \sin\theta \tag{3.5}$$

where ρ is the perpendicular distance of the line from the origin and θ is the angle of the line normal. A set of collinear pixels will generate several corresponding curves of Eq. (3.5). Local maxima of the pixel intensity identify straight-line segments in the image space (shown in Figs. 3.9 (d), 3.10 (d)). Ideally, the Hough domain has to be searched for a maximum only once. In situations where a picture contains many patterns of different size, it may, however, be necessary to first take out those patterns that correspond to clearly identifiable peaks in the Hough domain and to repeat the process. As a result, we know the roads are existed in the somewhere of the images (shown in Figs. 3.12 (a), (b)). The place where there are a number of collinear bundles imply that the roads are wide, while the place where there are a few collinear bundles imply that the roads are narrow.



(a) Result of example 1 (b) Result of example 2

Figure 3.11 The middle results of ρ - θ Hough transform



(a) Result of example 1 (b) Result of example 2

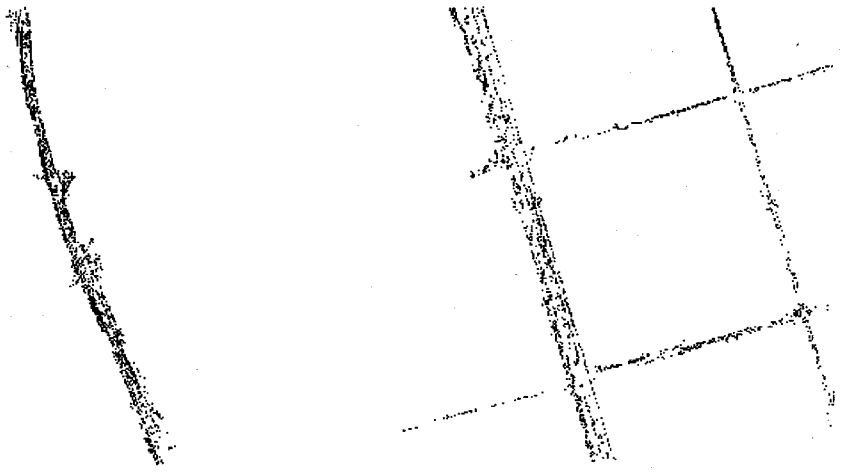
Figure 3.12 Results based on Hough transform

Integration of road radiometry and geometry

As mentioned above, it was explained that the road segments are described by their radiometric and geometric properties. The ideal road candidate regions will have both good radiometric and geometric properties in a local neighborhood. According to the previous satellite image processing and analysis, it is assumed that possibility of a real road can be evaluated using the fuzzy rules (shown in Table 3.4) that integrate the radiometric and geometric properties. Consequently, the steps below will focus on integrating the results of Figs. 3.9 (d) and 3.10 (d) (radiometric properties) and results of Fig. 3.12 (geometric properties) more tightly. The final results of road extraction are shown in the Fig. 3.13.

Table 3.4 Fuzzy rules for the road extraction

Road Possibility for radiometry μ_{radio}	Road Possibility for geometry μ_{geo}	Road Candidate Regions μ_{road}
Low	Low	Low
Low	High	Low
High	Low	Low
High	High	High

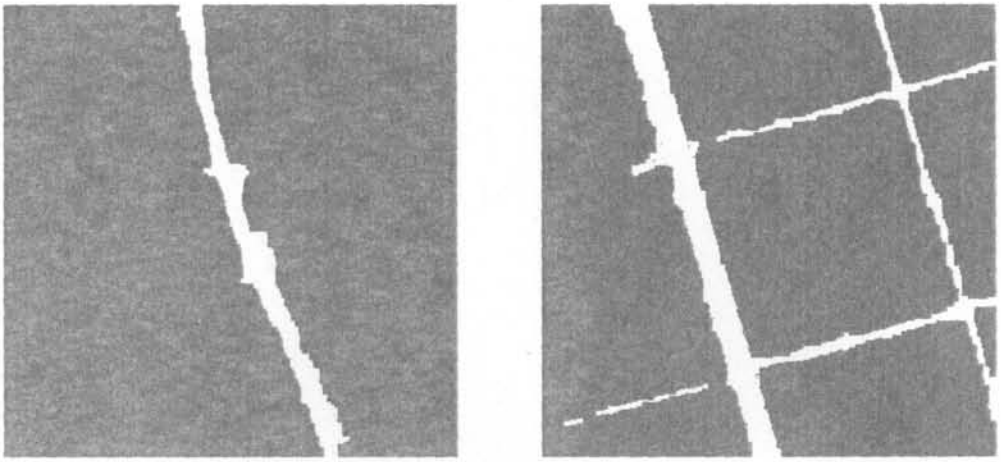


(a) Result of example 1 (b) Result of example 2

Figure 3.13 Results of road extraction from radiometry and geometry

3.3.5 Image Post-Processing

To modify the shape of road segments, morphological operations such as thinning, thickening, erosion, dilation, opening and closing are very effective. Erosion and dilation are neighborhood operations that determine each pixel's value according to its geometric relationship with neighboring pixels. Erosion peels layers from objects or particles, removing extraneous pixels and small particles from the image. Dilation adds layers to objects or particles. Then the final results of the extracted main road segments are clearly shown in Fig. 3.14.



(a) Result of example 1 (b) Result of example 2

Figure 3.14 Final results of road segments extraction

3.4 Experiment with Large Images

We provide some additional examples for the proposed fuzzy logic based road extraction algorithms for large satellite images. Figure 3.15 is a 1.0 meter high-resolution IKONOS image (1000×1000 pixels, 1m/pixel). It is clear that the main suburban road segments are accurately extracted from high-resolution satellite imagery as shown in Fig. 3.16. Initial seed points, used by Agouris *et al.* [Agouris *et al.*, 1998], need not be selected and there is no need for a road width estimate in the road extraction process.

Figure 3.15 Original image of example 3 (1000×1000 pixels, 1m/pixel)

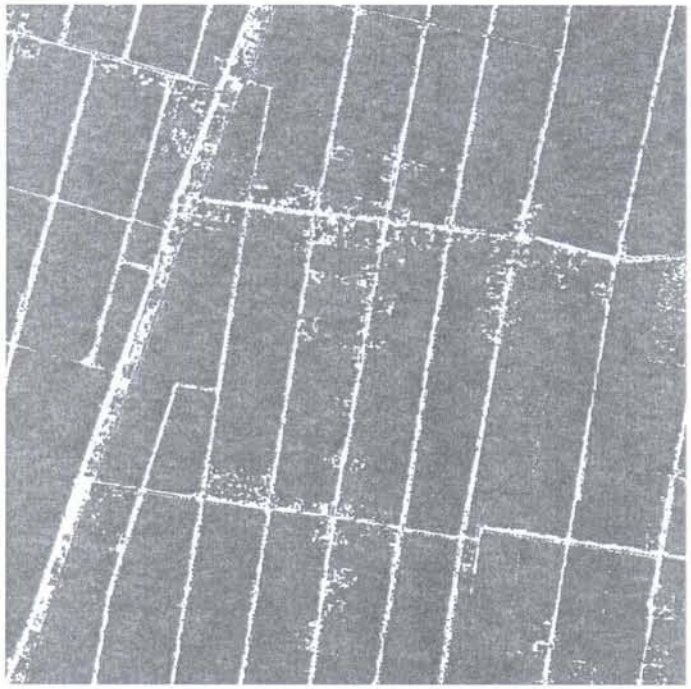


Figure 3.16 Extracted result of the main road segments

3.5 Discussion and Conclusions

The contribution of this chapter is that it proposes an effective method for road extraction from satellite images based on a heuristic fuzzy logic approach. We have described how a system of different high-resolution satellite images, acquired under different conditions, can be used to detect and recognize road segments.

Compared with [Agouris *et al.*, 1998], initial seed points need not be selected and there is no need for a road width estimate in the road extraction process. The algorithms have been successfully implemented and tested on different high-resolution IKONOS satellite images, and the major suburban roads were extracted (shown in Fig. 3.14 and Fig. 3.16). From the extraction results obtained using fuzzy logic, we find that forests and large area fields are clearly distinguished from roads, and only the main road candidate regions taken from the images were obtained. The results show that the approach is valid for roads extracted from high-resolution satellite images. We have also shown how to implement structure extraction algorithms using fuzzy logic approaches.

In addition, because imaging conditions can vary considerably, special attention must be devoted to the flexibility of the heuristic system. For example, there are perturbations, such as the shadows of buildings, in satellite images. The algorithm outlined in this chapter is merely a first step in this direction. Many further steps will have to be taken to develop the algorithm more fully. The same approach will be used in the future for the construction of a heuristic system, which is able to extract road segments from satellite images of urban roads. This work is intended to be part of such a system.

Chapter 4
Road Network Extraction

4.1 Introduction

Neural network is one of the important components in Artificial Intelligence (AI). It has been studied for many years in the hope of achieving human-like performance in many fields, such as speech and image recognition as well as information retrieval. Kohonen's Self-Organizing Map (SOM) is one of the most popular artificial neural networks [Kohonen, 2001]. The SOM is a two-layer network with an input layer and a Kohonen layer (shown in Fig. 4.1). The input layer acts as a "fan-out" or distribution layer, sending the complete input pattern to each neuron in the Kohonen layer across a set of adaptive weighted connections. The Kohonen layer neurons make up a competitive assembly. That is, the neurons within the layer compete for the privilege of producing the layer's single output, a +1 signal from the most competitive neuron within the layer. The competition is mediated by the interneuron connections among the neurons in the layer. Typically, each neuroid has strong

positively weighted connections to its immediate physical neighbors in the layer, and somewhat negatively weighted connections to neurons farther away.

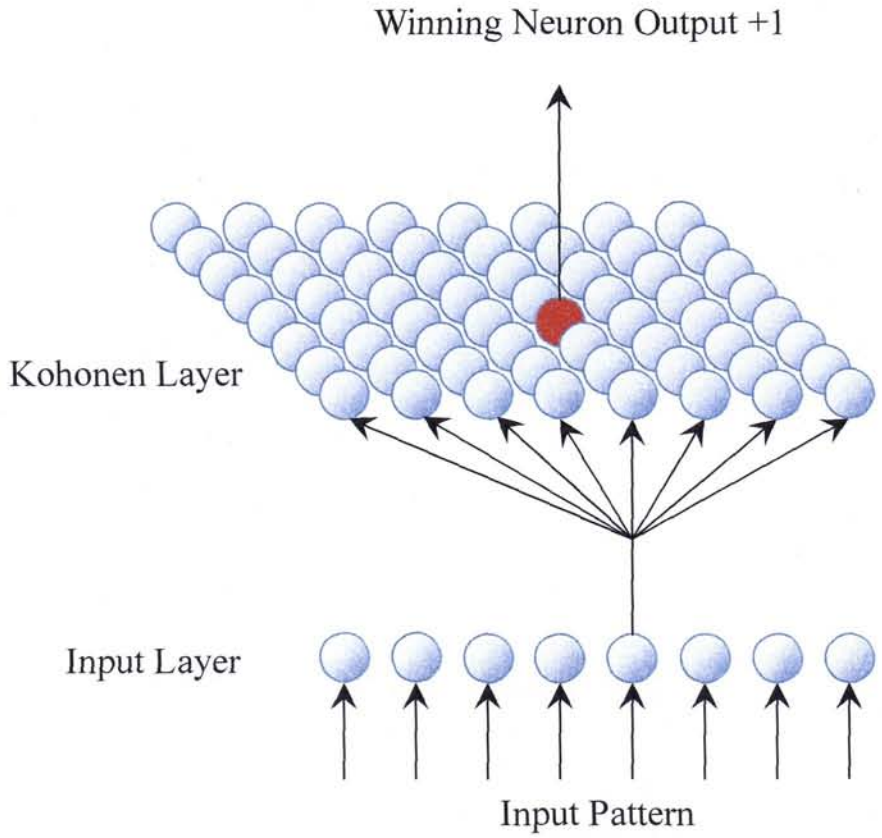


Figure 4.1 Kohonen layer

The self-organizing map (SOM) [Kohonen, 2001] is a neural network algorithm that has been used for a wide variety of applications, mostly for engineering problems but also for data analysis. There are four particular views of the SOM that are selected to serve as a basis for further inspection.

(1) The SOM is a model of specific aspects of biological neural nets.

The most typical notion of the SOM is to consider it as an artificial neural network model of the brain, especially of the experimentally found ordered "maps" in the

cortex. There exists a lot of neuron physiological evidence to support the idea that the SOM captures some of the fundamental processing principles of the brain.

(2) The SOM constitutes a representative of a new paradigm in artificial intelligence and cognitive modeling.

The SOM can be viewed as a model of unsupervised (machine) learning, and as an adaptive knowledge representation scheme. The traditional knowledge representation formalisms - semantic networks, frame systems, predicate logic, to provide some examples, are static and the reference relations of the elements are determined by a human. Related to knowledge representation and learning, the cognitive and philosophical aspects are highly relevant.

(3) The SOM is a tool for statistical analysis and visualization.

The SOM is nowadays often used as a statistical tool for multivariate analysis. The SOM is both a projection method which maps high-dimensional data space into low-dimensional space, and a clustering method so that similar data samples tend to be mapped to nearby neurons. From the methodological and computational point of view the mathematical and statistical properties of the algorithm can be considered (for instance, the time and space complexity, the convergence properties), as well as the nature of the input data (signals, continuous statistical indicators, symbol strings) and their preprocessing. There exist a number of variants of the SOM in which, e.g., the adaptation rules are different, various distance measures can be used, and the structure of the map interconnections is variable.

(4) The SOM is a tool for the development of complex applications

The SOM is widely used as a data mining and visualization method for complex data sets. Application areas include, for instance, image processing and speech recognition, process control, economical analysis, and diagnostics in industry and in medicine.

Road category maps are SOM that have been organized according to road similarities, measured by the similarity of the short contexts of the road segments.

Conceptually interrelated road segments tend to fall into the same or neighboring map nodes. Nodes may thus be viewed as road categories. Although no a priori information about classes is given, during the self-organizing process a model of the road classes emerges.

The extraction of road networks from aerial and satellite imagery is a fundamental image analysis operation. The automatic completion of road networks based on the generation and verification of link hypotheses from incompletely extracted road networks is described in [Wiedemann *et al.*, 2000]. The proposed link hypotheses are verified using actual image data. In [Zhang *et al.*, 2002], a concept for road network reconstruction from aerial images using knowledge based image analysis is developed. The proposed approach uses multiple cues about the object existence, employing existing knowledge, rules and models, and treats each road subclass differently to increase success rate and reliability of the results. In [Doucette *et al.*, 2001], Doucette *et al.* present an approach to road extraction in classified imagery using self-organized clustering, combining a straight K-medians spatial clustering approach with a post-convergence node linking minimum spanning tree (MST) algorithm. Most of the existing road extraction methods for multi-spectral imagery rely heavily on automatic and reliable classification of road surfaces. The limitation of node linking was found in the opened loop candidate domain. In contrast to other approaches, our work focuses on road network extraction from grayscale images, first preprocessed with fuzzy logic techniques [Yun *et al.*, 2006], and then using neural networks for seed points extraction, followed by linking the seed points by tracking. With the goal of evaluating the method's potential in extracting the road network, many simulations have been carried out using different types of images. First, in this chapter we propose to use a self-organizing map, inspired from a specialized variation of Kohonen's self-organizing map (SOM) neural network, in order to find the seeds of the road network. The seed points converge onto the road candidate domain by a learning process using grayscale images. During the learning process, the learning rate factor $\alpha(t)$ that is important parameter in the convergence is discussed in detail. Next, our approach to road tracking is presented. This

algorithm searches for connected points in the direction and the candidate domain of a road. The definition of the connection rules is described between the nodes. Finally, results of simulations are shown for a road network extracted using this method. The central topic of the research is the use of the SOM in natural road network processing. Tests with high-resolution satellite images are presented in this Chapter.

4.2 Methodology for Road Network Extraction

4.2.1 Interpretation of Self-Organizing Maps

The self-organizing map is a competitive network with the ability to form topology preserving mappings between its input and output spaces. In this program the network learns to balance a pole by applying forces at the base of the pole. The behavior of the pole is simulated by numerically integrating the differential equations for its law of motion using Euler's method. The task of the network is to establish a mapping between the state variables of the pole and the optimal force to keep it balanced. This is done using a reinforcement learning approach: for any given state of the pole the network tries a slight variation of the mapped force. If the new force results in better control, the map is modified, using the pole's current state variables and the new force as a training vector.

Basically, the proposed neural network consists of two layers (shown in Fig. 4.2). The basic SOM can be visualized as a sheet-like neural-network array, the cells (or nodes) of which become specifically tuned to various input signal patterns or classes of patterns in an orderly fashion. The learning process is competitive and unsupervised, meaning that no teacher is needed to define the correct output for a given input. In the basic version, only one output node (winner) is activated for each input. The locations of the responses in the array reflect the statistical distribution of the input features.

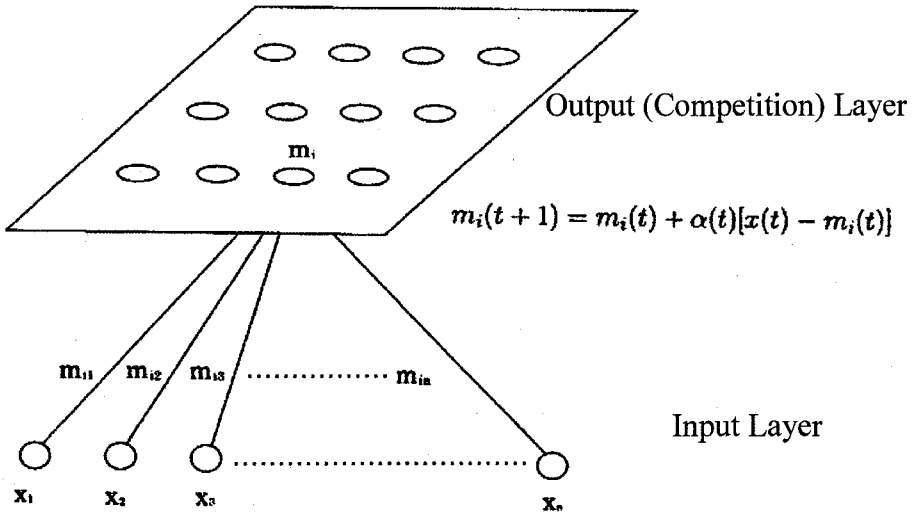


Figure 4.2 The structure of the proposed neural network

Here, we proceed with a brief description of its behavior. Assume that some sample data sets are mapped onto the array depicted. The set of input samples is described by a real vector $\mathbf{x}(t) \in R^n$ where t is the index of the sample, i.e., the discrete-time coordinate. Each node i in the map contains a model vector $\mathbf{m}_i(t) \in R^n$, which has the same number of elements as the input vector $\mathbf{x}(t)$.

The stochastic SOM algorithm essentially performs a regression analysis. The initial values of the components of the model vector, $\mathbf{m}_i(t)$, can be selected at random. In practical applications, however, the model vectors are more profitably initialized in some orderly fashion, e.g., along a two-dimensional subspace spanned by the two principal eigenvectors of the input data vectors.

Any input is mapped into the location for which $\mathbf{m}_i(t)$ of which most closely matches $\mathbf{x}(t)$, according to some predefined metric. The self-organizing network creates the mapping of the inputs by using the following of steps:

(i) An input vector $\mathbf{x}(t)$ is compared with all the model vectors $\mathbf{m}_i(t)$. The best-matching unit (node) on the map, i.e., the node where the model vector is most similar to the input vector in some metric (e.g.

Euclidean distance) is identified. This best matching unit is often called the winner.

(ii) The model vectors of the winner and a number of its neighboring nodes are moved toward the input vector according to the learning principle described below.

The basic idea in the SOM learning process is that, for each sample input vector $\mathbf{x}(t)$, the winner and the nodes in its neighborhood are moved closer to $\mathbf{x}(t)$ in the input data space. During the learning process, individual changes may be contradictory, but the net outcome in the process is that ordered values for the $\mathbf{m}_i(t)$ emerge over the array. If the number of available input samples is restricted, the samples must be presented reiteratively to the SOM process.

Fitting of the model vectors $\mathbf{m}_i(t)$ is usually carried out by a sequential regression process, where t is the step index. Each input vector $\mathbf{x}(t)$ may be compared with all the $\mathbf{m}_i(t)$ in any metric; in many practical applications, the smallest of the *Euclidean distances* $\|\mathbf{x}-\mathbf{m}_i\|$ can be made to define the *best-matching node*, and the presented input vector is selected as the winner. Assuming a total of K neurons, the node c that best represents the input is then searched for using the *Euclidean distances* to define the quality of the representation:

$$\text{node } c = \arg\min_i \{\|\mathbf{x}-\mathbf{m}_i\|\}, \quad \text{which means the same as}$$

$$\|\mathbf{x}-\mathbf{m}_c\| = \min_i \{\|\mathbf{x}-\mathbf{m}_i\|\} \quad \text{for } i = 1, 2, \ldots, K \tag{4.1}$$

where c indicates the best-matching unit for $\mathbf{x}$. During training, the best-matching unit in the map is sought, for each input sample vector, and the average of the respective quantization errors is obtained.

Next, unit c as well as neighboring units learns in the learning process to represent the data sample more accurately. The model vector of unit $\mathbf{m}_i(t)$ is updated according to the following equations:

$$\begin{cases} \mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] & \text{for each } i \in N_c(t) \\ \mathbf{m}_i(t+1) = \mathbf{m}_i(t) & \text{otherwise} \end{cases} \tag{4.2}$$

where t is the discrete-time index of the variables, the factor $\alpha(t) \in [0,1]$ is a scalar that defines the relative size of the learning step, and $N_c(t)$ specifies the *neighborhood* around the winner in the map array.

At the beginning of the learning process the radius of the neighborhood is fairly large, but it shrinks during learning. This ensures that the global order is obtained at the beginning. As the radius gets smaller, the local corrections of the model vectors become more specific. A learning rate factor $\alpha(t)$ should be taken into account. The factors depend on the parameter t, which is referred to as the epoch. The learning rate factor $\alpha(t)$ also decreases monotonically during learning. Here $\alpha(t)$ can be defined to be linear with respect to t.

$$\alpha(t) = \alpha_0(1 - t/T) \tag{4.3}$$

where α_0 is the initial learning rate factor, and T is the total number of steps.

4.2.2 Approach to Road Tracking

The road candidate domain is tracked in order to form the road centerlines of the road network. An approach to road tracking involves linking two corresponding seed points in the road network. It assumes that road direction changes are likely to be slow and that roads are unlikely to be short. First, starting from an image edge, find an unrestricted seed point. Then, the algorithm searches for the two closest connected seed points in the direction of the road according to the white candidate domain. Finally, the two seed points are connected. A polygon, consisting of a series of

centerline segments in the candidate domain of the road, is created. Their topology describes the continuous road network.

Suppose the initial seed points are near the side of the image. The order for matching seed points is from left to right and top to bottom. The rules used to build the road network are as described below:

(i) The path vector direction is given priority. Namely, assuming that the initial direction is for vector $\mathbf{A}$, the next tracking direction vector between related seed points in the candidate domain is for vector $\mathbf{B}$. The deflection angle θ between the vectors $\mathbf{A}$ and $\mathbf{B}$ must be minimal.

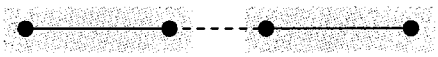
$$\min\theta = \min\arccos\left(\frac{\mathbf{A}\bullet\mathbf{B}}{|\mathbf{A}||\mathbf{B}|}\right) \tag{4.4}$$

(ii) The linked lines should fall in the candidate domain of the road.

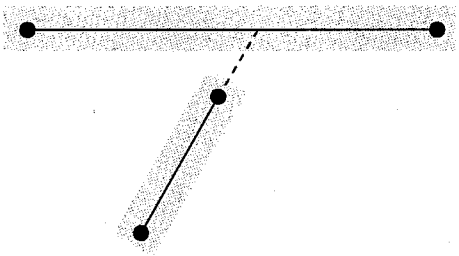
(iii) Roads are unlikely to be short, and are, moreover, continuous.

The process is repeated until all relevant seed points have been tracked. Due to disconnected candidate domains in the grayscale images, the tracking of some seed points is terminated prematurely. There are two types of breaks (the straight candidate domain, at T-junctions) in grayscale images as shown in Figs. 4.3 (a), (b). In such cases, the tracking approach will find other seed points to form another centerline that will terminate close to the prematurely terminated centerline end point. The domains are close if they are within the same path vector direction. A check is made to ensure that the seed points of the link actually lie in a grayscale image region to ensure that dead end streets do not erroneously get connected. Connection of road centerline end points is performed by concatenating chains as described in Fig. 4.3 (a). Once two seed points have been connected, the process is repeated until no more concatenation can occur. During concatenation, all road junctions are found

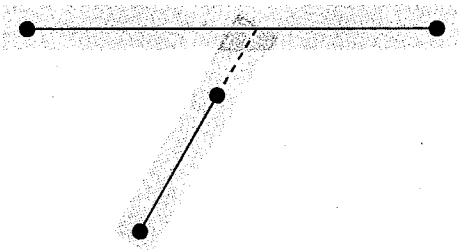
by identifying joining road candidate domain. These candidate domains are broken at the T-junctions and the road concatenating chain is created as shown in Fig. 4.3 (b).



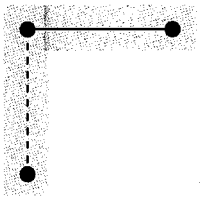
(a) Break in the straight candidate domain



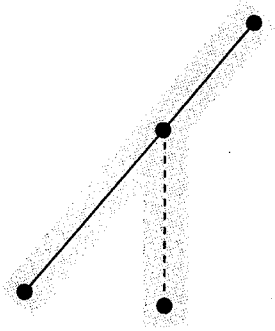
(b) Break at a T-junctions



(c) T-junctions



(d) Corner-junction



(e) Branch-junction

Figure 4.3 Seed points connected completely

In order to maintain the topology of the road network, a list of all intersections and

T-junctions is kept in conjunction with the connecting seed points on the candidate domain of road. In the case of T-junctions, this requires that the end of each connected line be checked against all other connected lines. If the direction to the closest adjacent connected line position matches the current line direction, the line of connected seed points is extended provided the concatenating chain is created as shown in Figs. 4.3 (b), (c). The process is repeated until all T-junctions have been modeled. Where a corner-junction is encountered, the seed points are tracked by turning round. The concatenating line is created from the road candidate domain (as shown in Fig. 4.3 (d)). Where a branch-junction is encountered, the unlinked seed points should be checked; starting from this point, these are once again linked by road a tracking approach (as shown in Fig. 4.3 (e)).

4.2.3 Simulated Image

Figure 4.4 demonstrates the process using a simulated image (200×200 pixels) with some noise added. The samples shown in Fig. 4.4 represent pixels along a road segment. The input to the network is the (x, y) coordinate of each sample. In this case, the 25 nodes (forming an initialization grid with points at 50 pixel intervals) in the input space are shown. In Eqs. (4.1) and (4.2), the vector $\mathbf{x}(t)$ contains the image coordinates of the gray pixels, while the vector $\mathbf{m}_i(t)$ contains the image coordinates of the nodes. The Euclidean distances between an input vector $\mathbf{x}(t)$ and the model vector $\mathbf{m}_i(t)$ is signified by the subscript E_i:

$$E_i = (x_1 - m_1^i)^2 + (x_2 - m_2^i)^2 \tag{4.5}$$

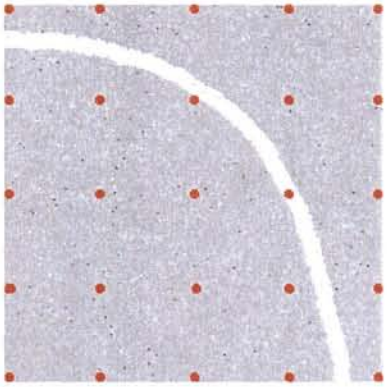
where x_1, m_1^i correspond to x coordinate component of an input vector $\mathbf{x}(t)$ and the model vector $\mathbf{m}_i(t)$, respectively. At the same time, x_2, m_2^i correspond to y coordinate component of an input vector $\mathbf{x}(t)$ and the model vector $\mathbf{m}_i(t)$, respectively.

The smallest of the Euclidean distances E_{ic} can be made to define the best-matching node.

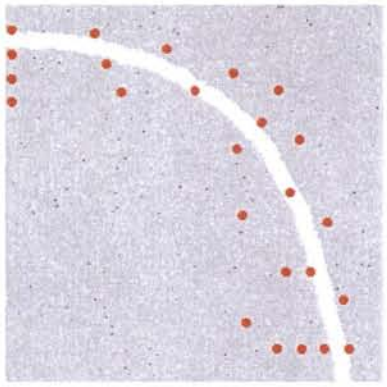
$$E_{ic} = \min E_i \tag{4.6}$$

The initial nodes move toward the gray pixels. The size of the neighborhood kernel is typically set up to be a function of time (epoch). Initially it is large, and a single node carries along with it many nodes (this is the ordering stage of the SOM), but it progressively shrinks until only the winning node is updated in the last iteration Similarly, the learning rate factor approaches 0 as the number of epochs increases, in order to facilitate the stabilization of the solution. Finally, the winning nodes are connected using the tracking rules as shown in the Fig. 4.4 (f).

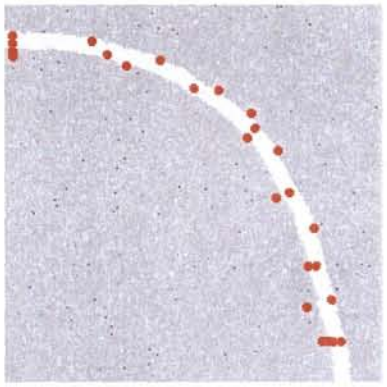
(a) Sample image (add some noise)



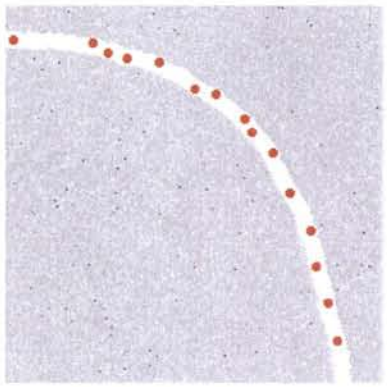
(b) Initialization grid (50 pixel intervals)



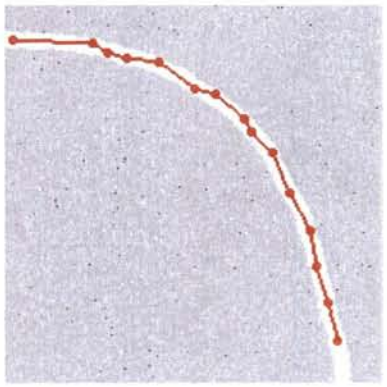
(c) After the second iteration



(d) After the 4th iteration



(e) After the 10th iteration



(f) Points linked

Figure 4.4 Simulation process of a sample image with some noise

4.3 Simulations of Road Network Extraction

4.3.1 Outline of Approach

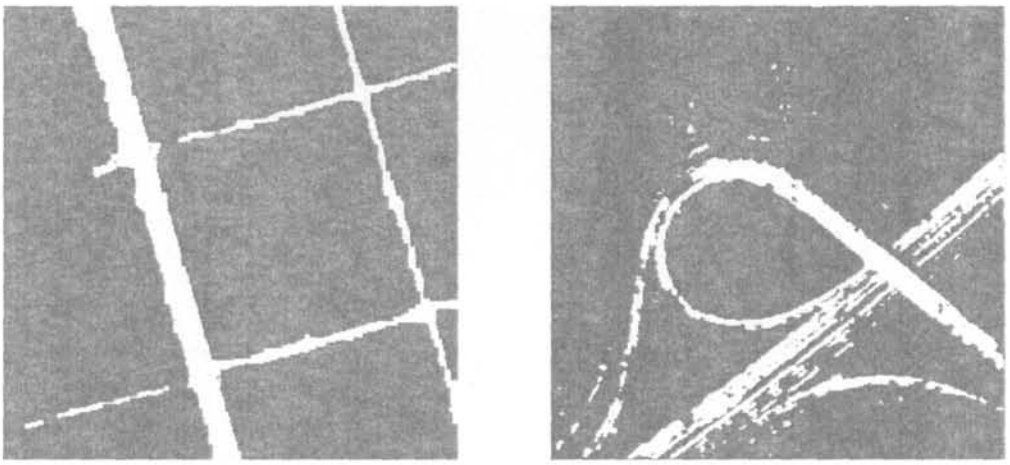
In order to evaluate the potential of this methodology for road network extraction, experiments with three kinds of IKONOS images were carried out. First of all, preprocessing of the original satellite images is carried out using fuzzy logic techniques [Yun *et al.*, 2006] in order to find the candidate domain of a road. The preprocessing results are shown in Fig. 4.6, Fig. 4.11 (b), and Fig. 4.13 (b). Road category maps are SOM that have been organized according to road similarities, where the measure is the similarity of short road segments. Interconnected road segments tend to converge to the same or neighboring map nodes. Nodes may thus be viewed as defining separate roads. Although no a priori road information is given, the self-organizing process yields a model of the roads. As mentioned in the introduction, the road seed points are first found using a self-organizing map. Then the points are connected using direction information and the road candidate domain, according to the road tracking rules.

(a) Image of country road (b) Image of overpass

Figure 4.5 Original satellite imagery (300×300 pixels, 1m/pixel)

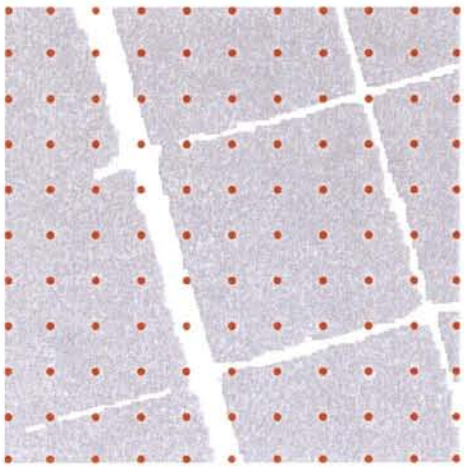
4.3.2 Small Image (300×300 pixels)

Here, the simulations are carried out on two kinds of small images: country road and overpass road images (300×300 pixels, 1m/pixel, shown in Figs. 4.5 (a), (b)). First, we give an initialization grid interval of 30 pixels for these small images (shown in Figs. 4.7 (a), 4.8 (a)). The results of SOM process can be seen in Figs. 4.7 (a)-(d) and Fig. 4.8 (a)-(d) respectively. It was observed that road seed points are shrunk onto the candidate domain of a road. In the second step, the road tracking process is implemented for linking seed points to obtain a road network (shown in Figs. 4.7 (e), (f), and Figs. 4.8 (e), (f)). In general, reducing the number of computations and raising the computing speed should allow, when using initialization grid interval of 50 pixels, to apply the approach to larger images. The processing results are shown in Figs. 4.9 and 4.10, respectively. In the same way, the road seed points are shrunk onto the candidate domain of a road. As a result, because the road seed points that lay on the candidate domain of the curved road are a few, it causes additional difficulties in the reconstruction of the curved road network using the road tracking approach. In fact, roads are continuous, so that we finally concatenate the chains obtained in the road network extraction (shown in Figs. 4.10 (e) (f)).

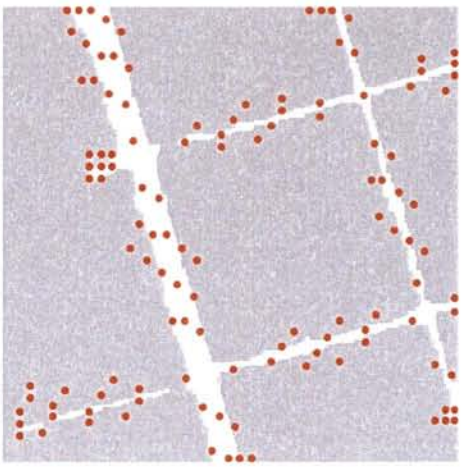


(a) Country road (b) Overpass

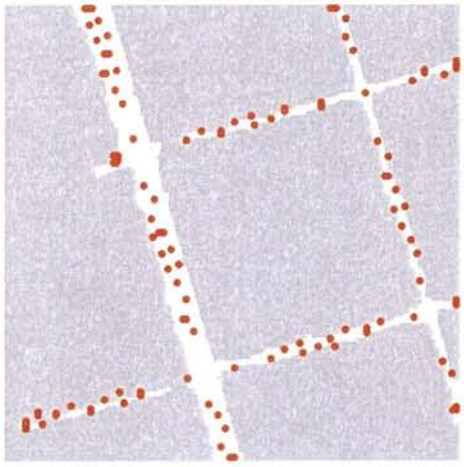
Figure 4.6 Preprocess results for original satellite imagery (300×300 pixels)



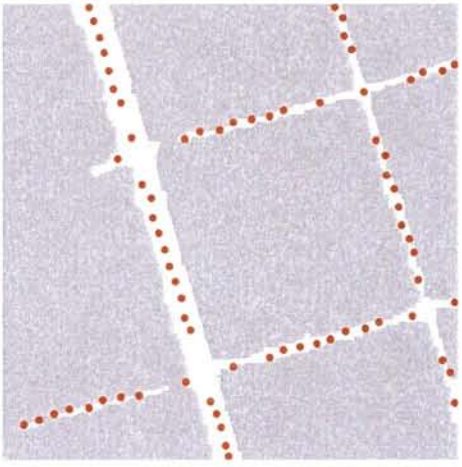
(a)



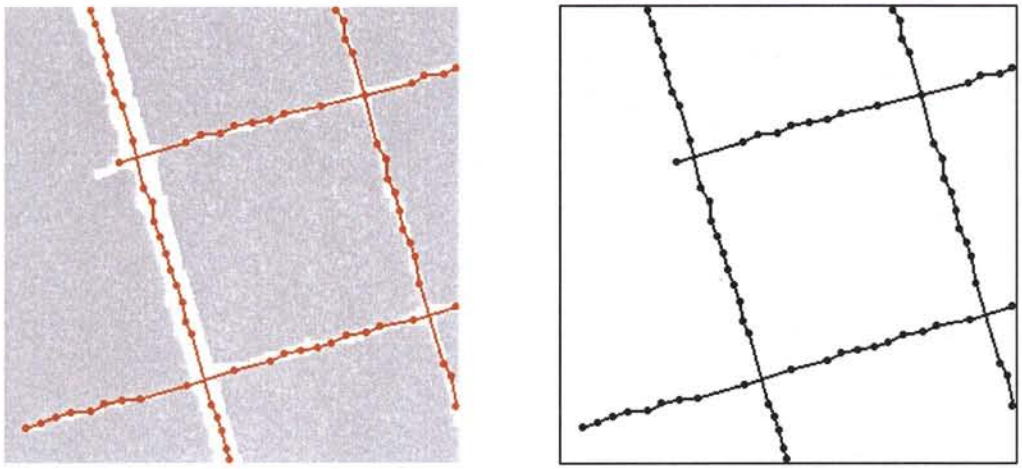
(b)



(c)

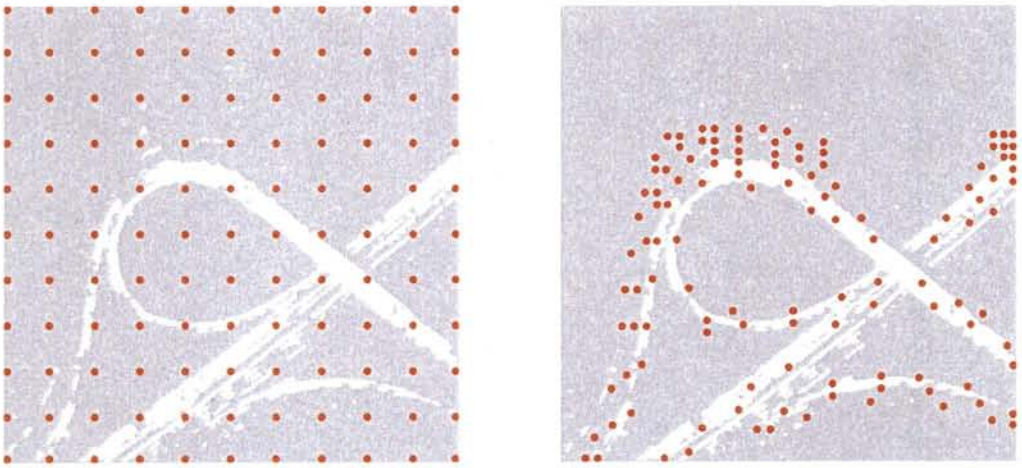


(d)

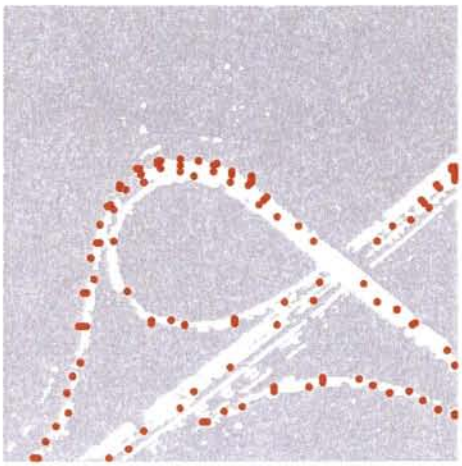


(e) (f)

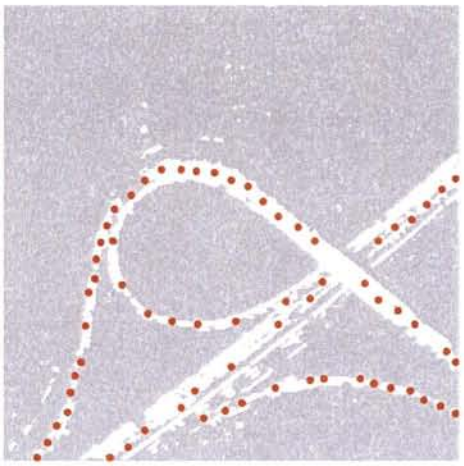
Figure 4.7 The results for the small country road image (Initialization grid interval of 30 pixels) (a) Initialization grid interval of 30 pixels (b) After the second iteration (c) After the 4th iteration (d) After the 10th iteration (e) Linking based on road tracking (f) Road network extraction



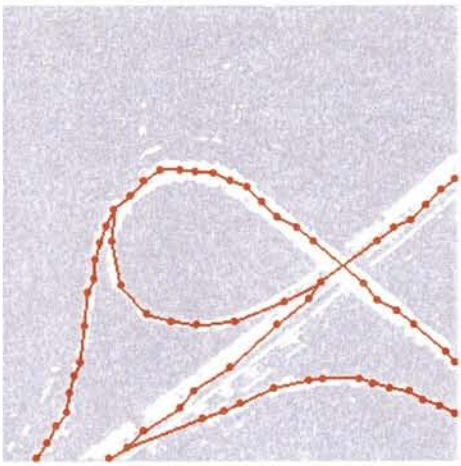
(a) (b)



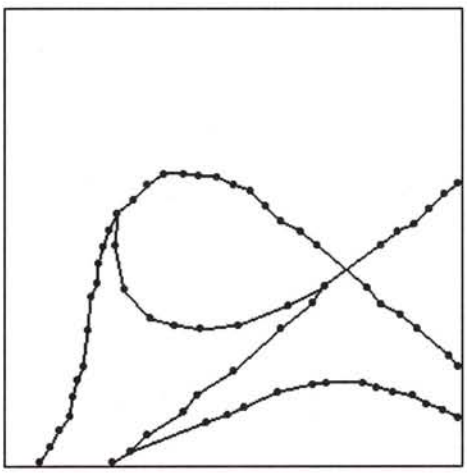
(c)



(d)

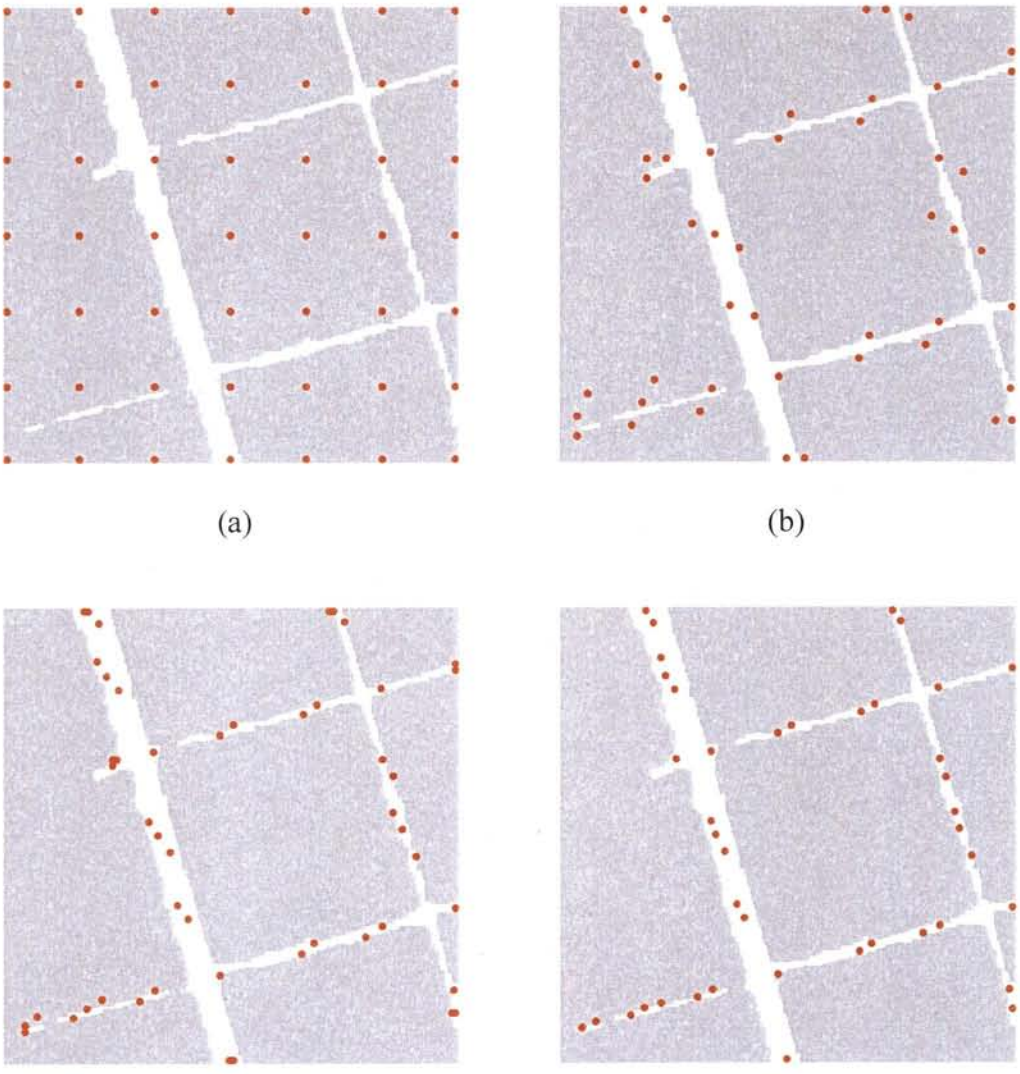


(e)



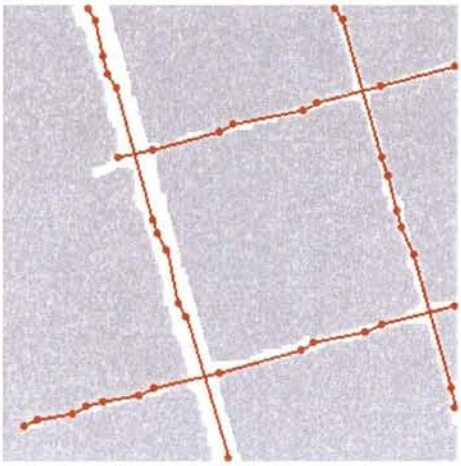
(f)

Figure 4.8 The Results for the small overpass image (Initialization grid interval of 30 pixels) (a) Initialization grid interval of 30 pixels (b) After the second iteration (c) After the 4th iteration (d) After the 10th iteration (e) Linking based on road tracking (f) Road network extraction

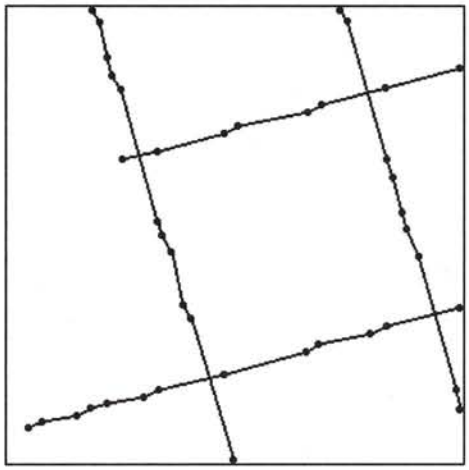


(a) (b)

(c) (d)

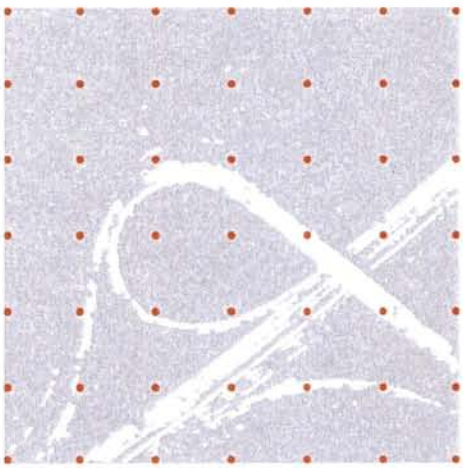


(e)

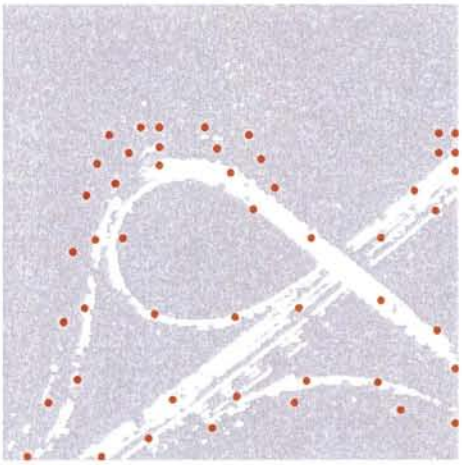


(f)

Figure 4.9 The results for the small country road image (Initialization grid interval of 50 pixels) (a) Initialization grid interval of 50 pixels (b) After the second iteration (c) After the 4th iteration (d) After the 10th iteration (e) Linking based on road tracking (f) Road network extraction



(a)



(b)

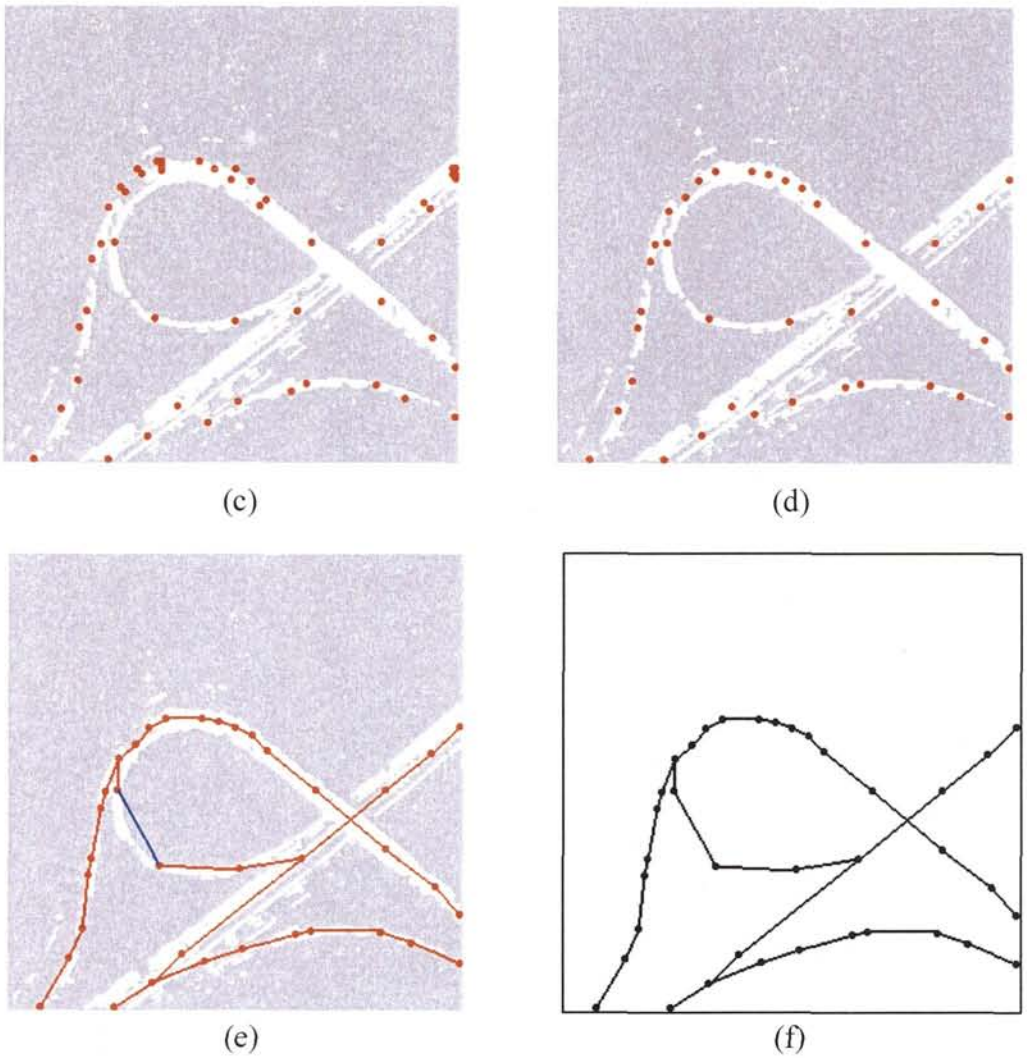
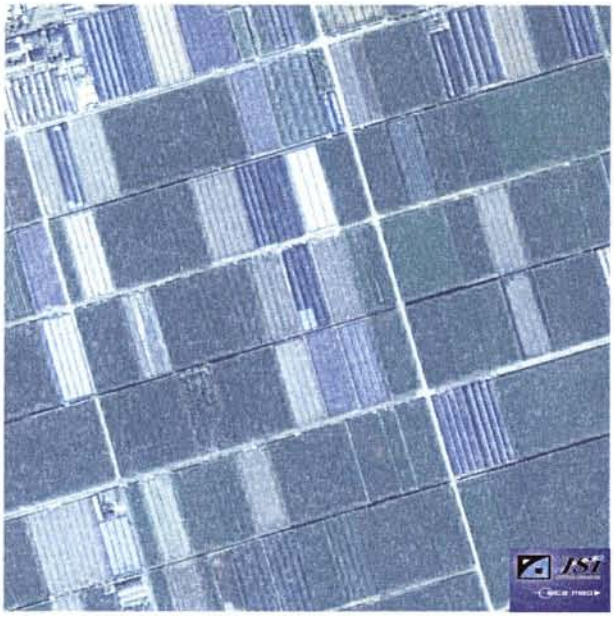


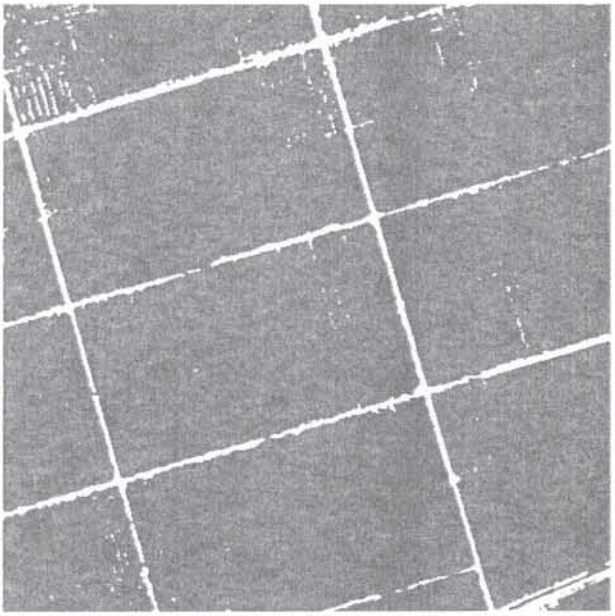
Figure 4.10 The Results for the small overpass image (Initialization grid interval of 50 pixels) (a) Initialization grid interval of 50 pixels (b) After the second iteration (c) After the 4th iteration (d) After the 10th iteration (e) Linking based on road tracking (f) Road network extraction

4.3.3 Middle Country Image (500×500 pixels)

Here, although we present the experimental results of small images (300×300 pixels), but for the purpose of evaluating performance of our approach, we implement the algorithm in the middle country road images (500×500 pixels, 1m/pixel, shown in Fig. 4.11 (a)). From the previous results shown in Figs. 4.9 and 4.10, we find that when the initialization grid intervals are increased, there is no effect in processing the images containing straight line roads. However, there is some effect in the processing of images containing curved roads. Therefore, we can use previous processing results to establish an initialization grid interval of 50 pixels for the processing of middle country road images (shown in Fig. 4.12 (a)). After many iterations, we can see that the road seed points are also shrunk onto the candidate domain of a road (shown in Fig. 4.12 (d)). Then, when all the relevant seed points have been tracked according to road tracking approach, the road network is obtained from the large country road image (shown in Figs. 4.12 (e) (f)).

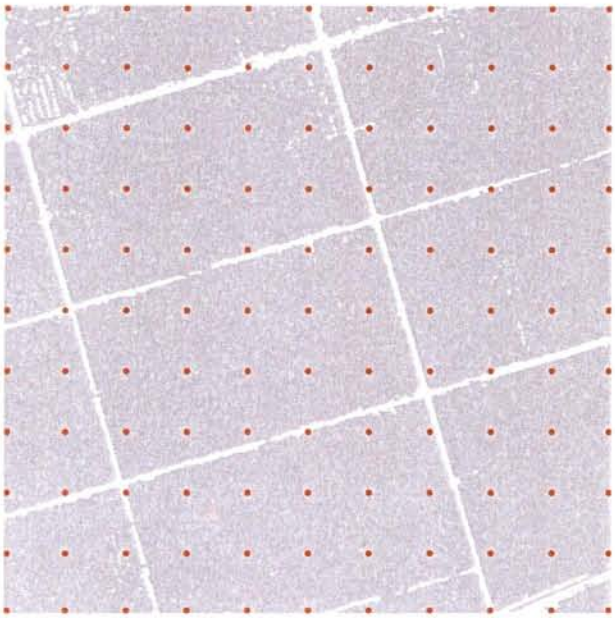


(a) The middle image with country road

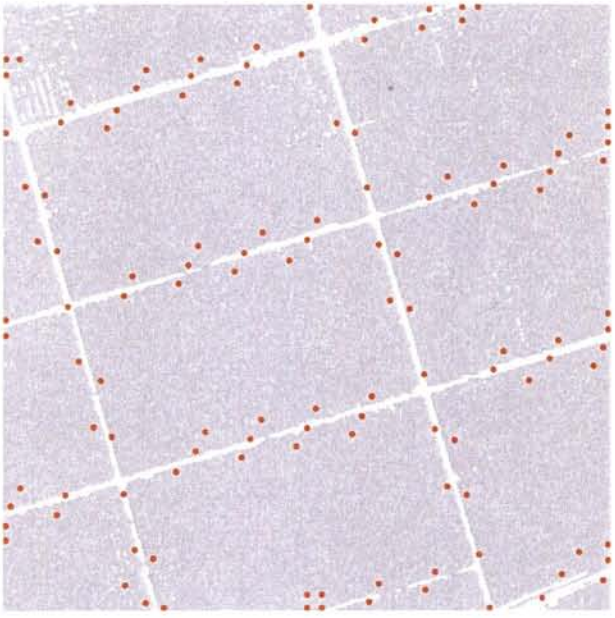


(b) Preprocess result for middle image

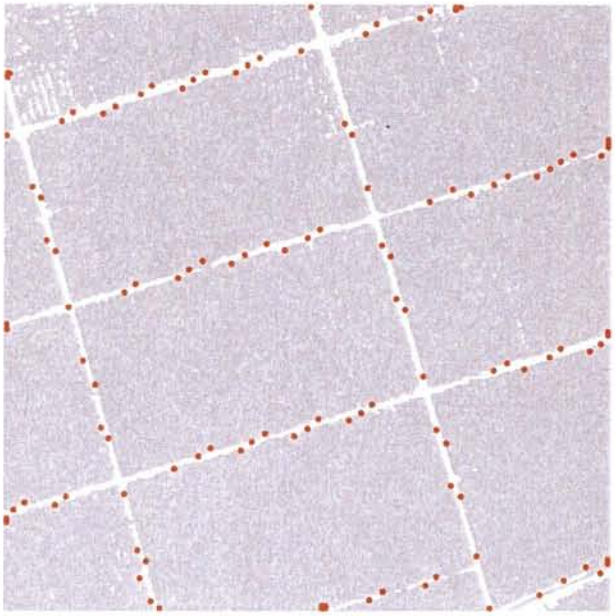
Figure 4.11 Middle original satellite imagery (500×500 pixels, 1m/pixel)



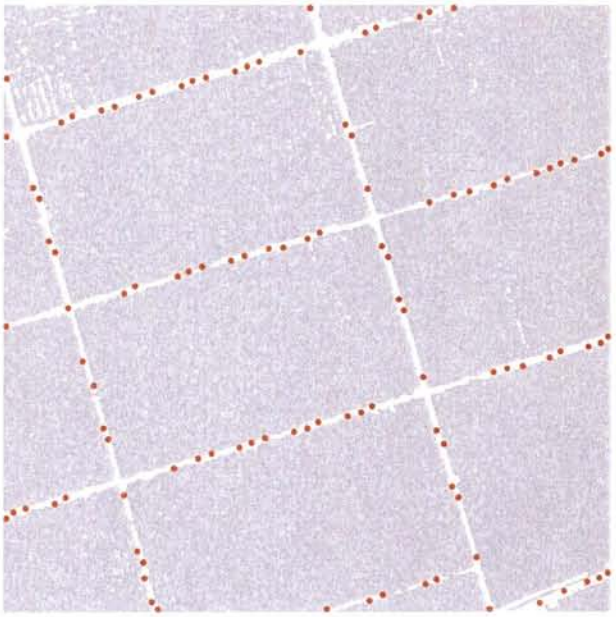
(a) Initialization grid interval of 50 pixels



(b) After the second iteration

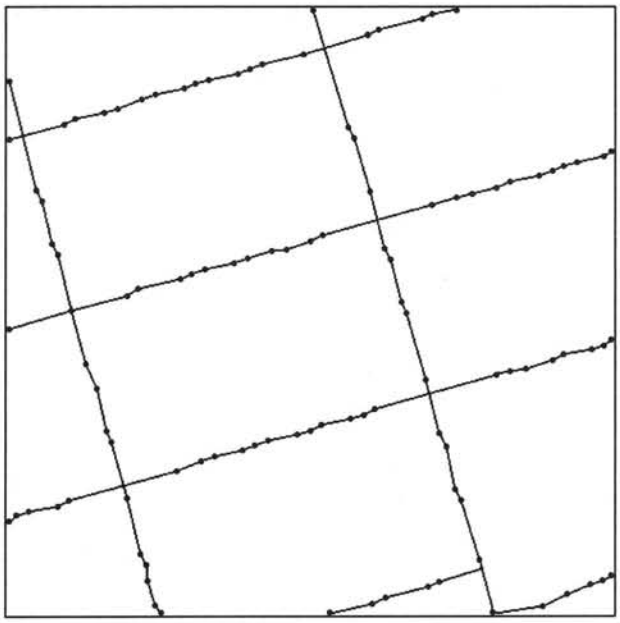


(c) After the 4th iteration



(d) After the 10th iteration

(e) Linking based on road tracking



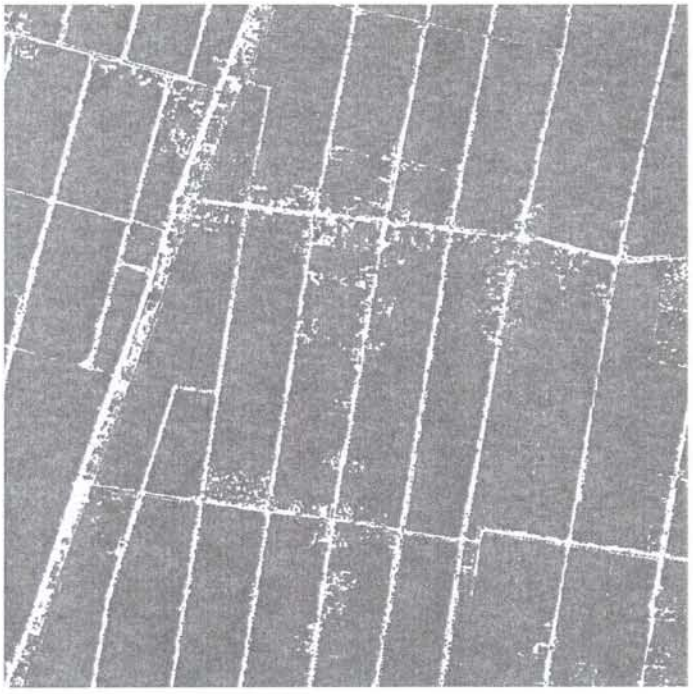
(f) Road network extraction

Figure 4.12 The results of SOM and road tracked process for the middle country road image

4.3.4 Large Country Image (1000×1000 pixels)

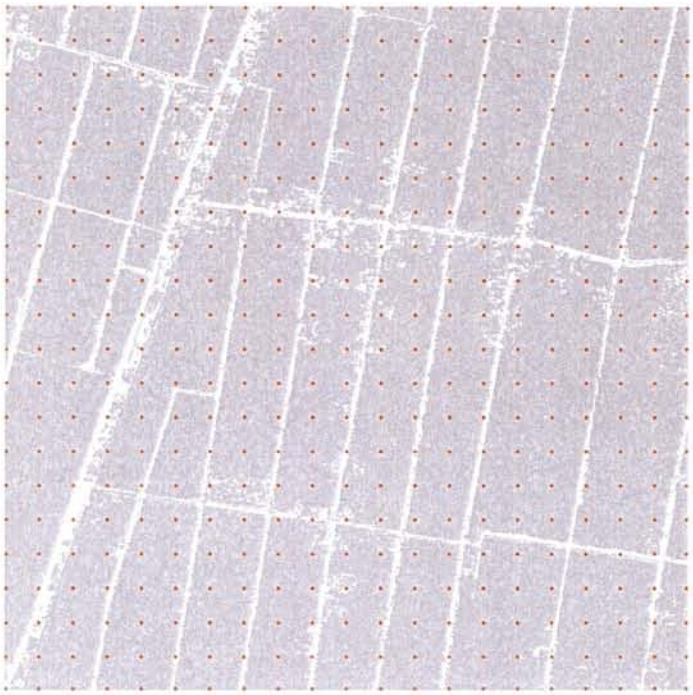
As mentioned above, the experimental results of small images (300×300 pixels, 500 ×500 pixels) are presented, but for the purpose of evaluating performance of the approach, a further large country road image (1000×1000 pixels, 1m/pixel, shown in Fig. 4.13 (a)) used the algorithm is also implemented. From the previous results shown in Figs. 4.7, 4.9 and 4.12, we find that when the initialization grid intervals are increased, there is no effect in processing the images containing straight line roads. Therefore, we also establish an initialization grid interval of 50 pixels for the processing of large country road images (shown in Fig. 4.14 (a)). After many iterations, we can see that the road seed points are also shrunk onto the candidate domain of a road (shown in Fig. 4.14 (d)). Then, when all the relevant seed points have been tracked according to road tracking approach, the road network is obtained from the large country road image (shown in Figs. 4.14 (e) (f)).

(a) The large image with country road

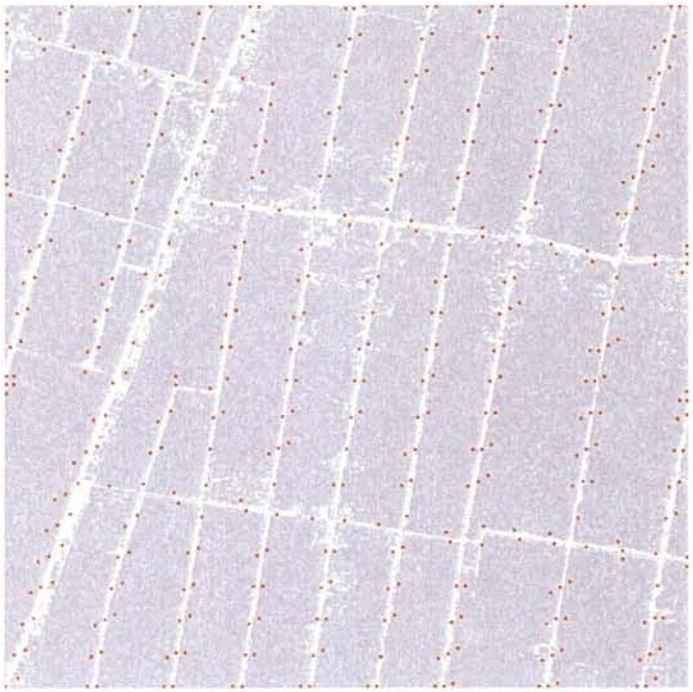


(b) Preprocess result for large image

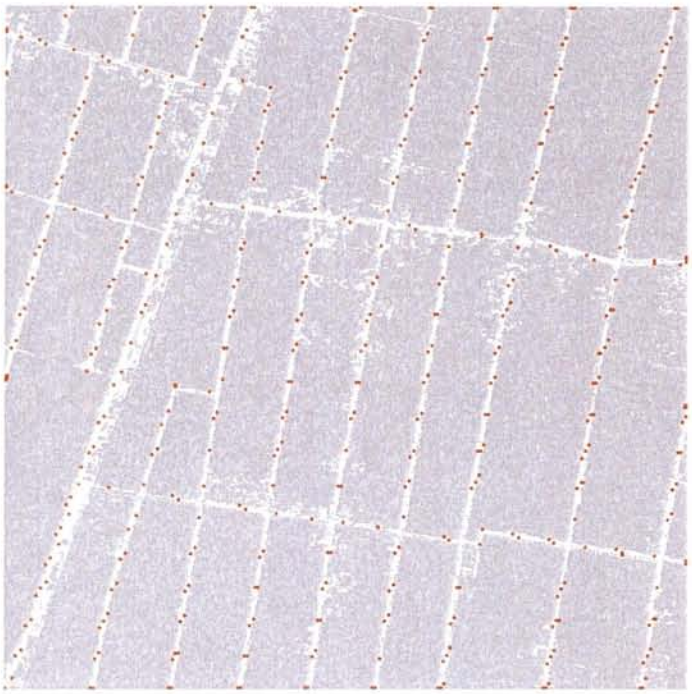
Figure 4.13 Large original satellite imagery (1000 × 1000 pixels)



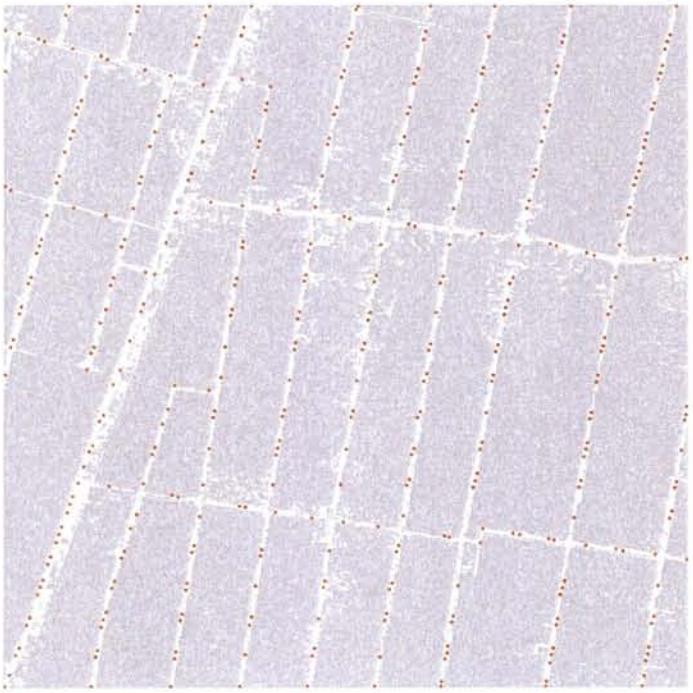
(a) Initialization grid interval of 50 pixels



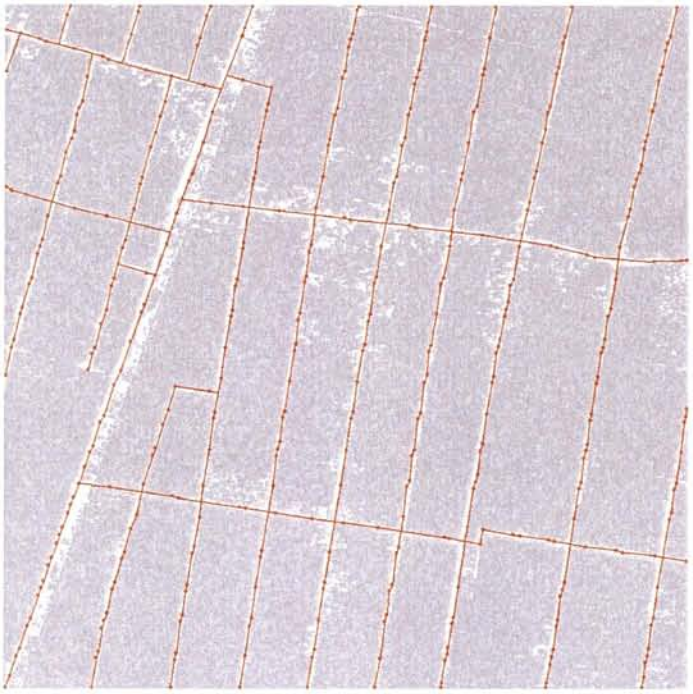
(b) After the second iteration



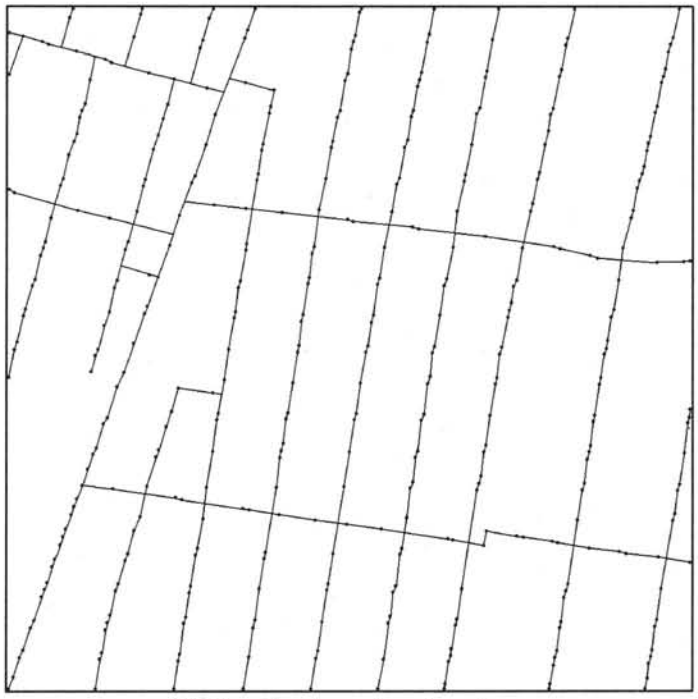
(c) After the 4th iteration



(d) After the 10th iteration



(e) Linking based on road tracking

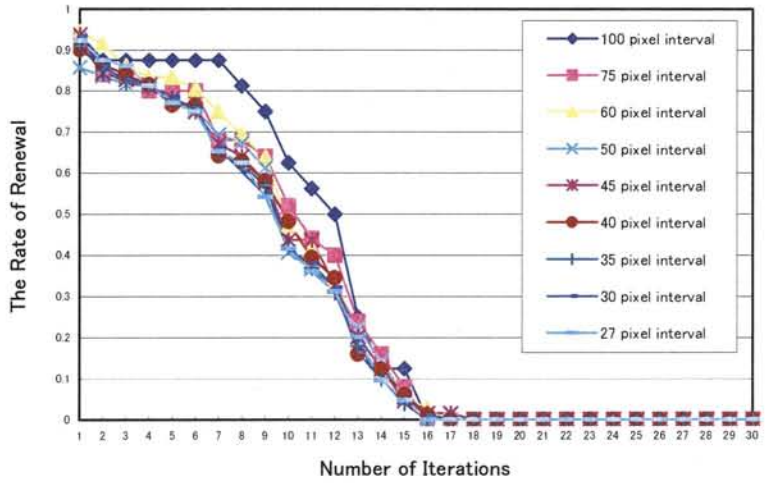


(f) Road network extraction

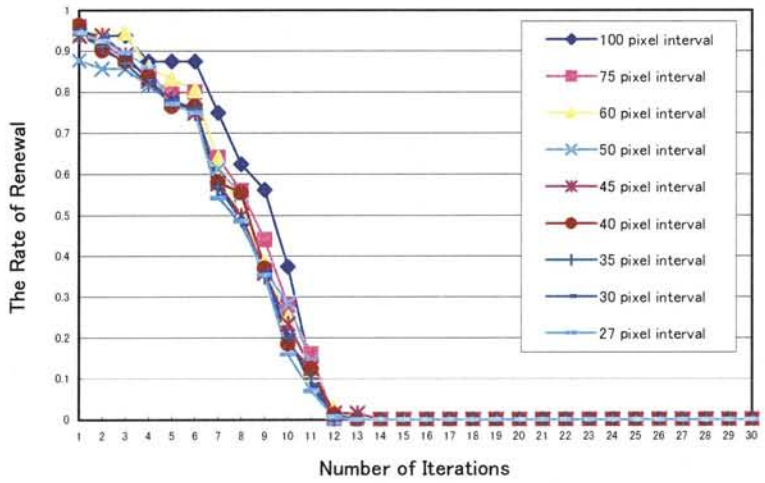
Figure 4.14 The results of SOM and road tracked process for the large country road image

4.4 Learning Rate Factor

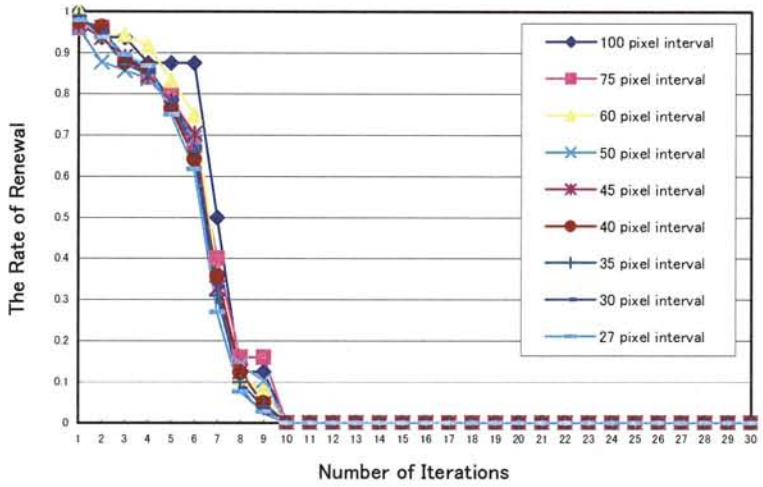
With large maps, however, it may be important to minimize the total learning time. Here, we choose a learning rate factor $\alpha(t)$ that is linear in t (shown in Eq. (4.3)). Usually $\alpha(t)$ is a small number to ensure that the network will converge to a solution, although at the cost of requiring many iterations. The size of $\alpha(t)$ can be increased as learning proceeds. Increasing $\alpha(t)$ as the network error decreases helps speed convergence by increasing the step size as the error approach a minimum However, with this approach, this is the risk that the network may oscillate around the minimum, rather than converging, if $\alpha(t)$ becomes too large.



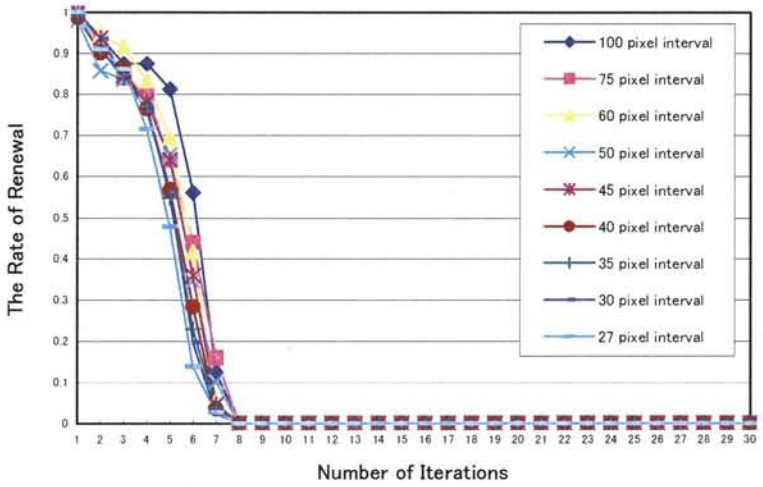
(a) α_0 =0.2



(b) α_0 =0.3



(c) α_0 =0.4



(d) α_0 =0.5

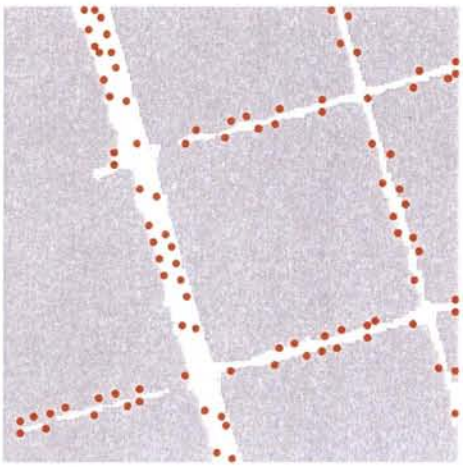
Figure 4.15 The rate of renewal

In the Eq. (4.3), the initial learning rate factor α_0 is usually kept on the order of 0.2 to 0.5, although values of $\alpha(t)$ in the 0.0 to 1.0 range are possible. In Fig. 4.15, the rate of renewal R expresses the updated state of the model vectors, $\mathbf{m}_i(t)$, for various initialization grid intervals. The rate of renewal R is defined as:

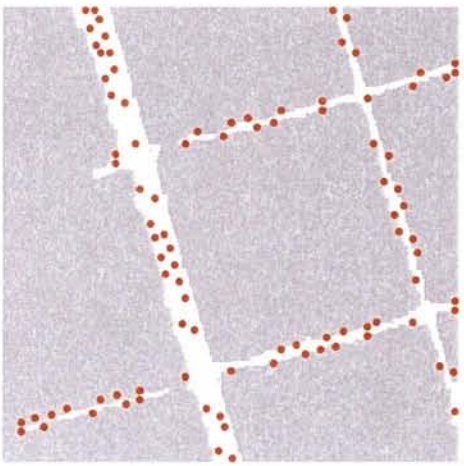
$$R = \frac{\Delta P_i}{P_0} = \frac{P_{i-1} - N_i}{P_0} \tag{4.4}$$

where P_0 is the total number of points corresponding to given initialization grid interval. ΔP_i is the number of renewal points in the ith iteration. Namely, it is different between the numbers of learning points in the i-1th iteration (P_{i-1}) and the numbers of no learning points in the ith iteration (N_i).

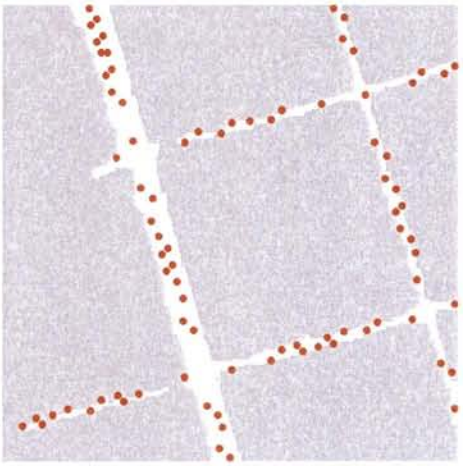
Figure 4.15 give us an overview of the relationship between the initial learning rate factor α_0, the initialization grid interval, and the number of iterations needed. We can see that no matter how many pixel initial intervals are determined, the number of iterations is the same corresponds to an initial learning rate factor α_0, when the learning process has reached stability. As α_0 becomes smaller, the number of iterations needed to achieve stability is greater. Moreover, the road convergence in the candidate domain is not very good (shown in Figs. 4.16 (a)-(c) and 4.17 (a)-(c)), because renewal points which are needed to the self-organizing are superposed in the first iteration, causing the number of the model vectors $\mathbf{m}_i(t)$ to be reduced. That is to say, the rate of renewal is not equal to 1 in the first iteration (shown in Figs. 4.15 (a)-(c)), this means the renewal information has been lost as soon as started iteration. However, when the initial learning rate factor α_0 is given 0.5, all the model vectors $\mathbf{m}_i(t)$ are updated, and there is no superposition in the first iteration. Namely, the rate of renewal is equal to 1 at that time (shown in Fig. 4.15 (d)) and convergence at road candidate domain is obtained (shown in Figs. 4.7 (d) and 4.8 (d), Figs. 4.9 (d) and 4.10 (d), Fig. 4.12 (d), Fig. 4.14 (d)). Figure 4.15 (d) shows that the rate of renewal is equal to zero approximately after the 8th iteration, which means that the learning process has reached stability. Moreover it is more quickly to achieve stability than the other situations.



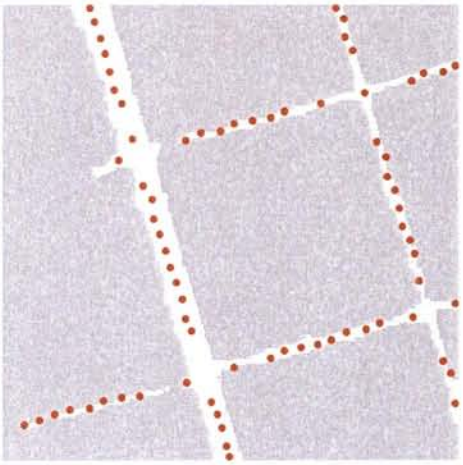
(a) α_0=0.2



(b) α_0=0.3

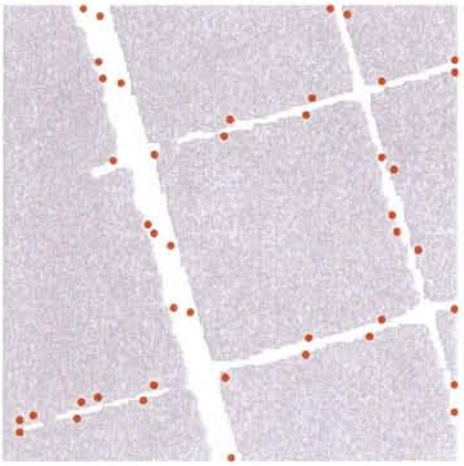


(c) α_0=0.4

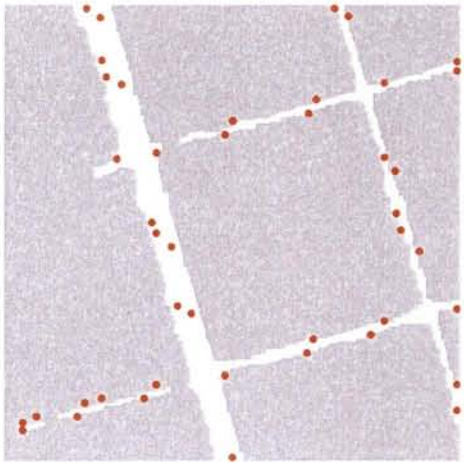


(d) α_0=0.5

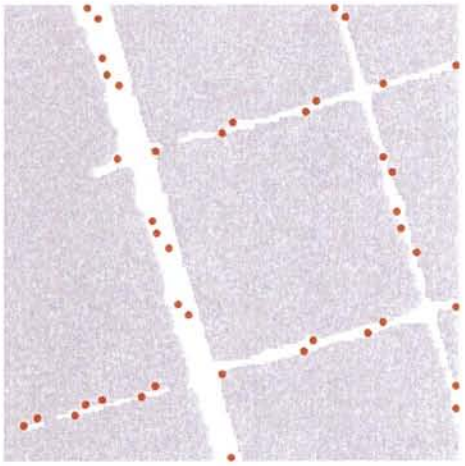
Figure 4.16 The situation of converge onto the road candidate domain (Initialization grid interval of 30 pixels)



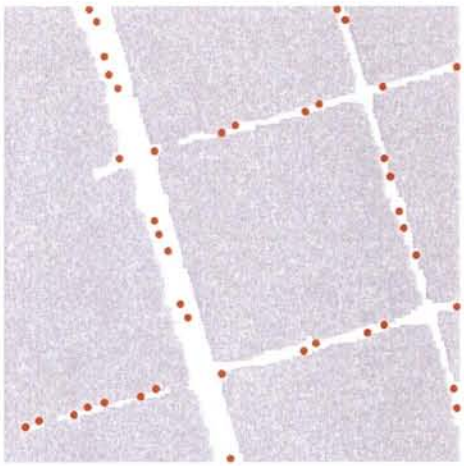
(a) α_0 =0.2



(b) α_0 =0.3



(c) α_0 =0.4



(d) α_0 =0.5

Figure 4.17 The situation of converge onto the road candidate domain (Initialization grid interval of 50 pixels)

4.5 Discussion and Conclusions

This chapter presents a semi-automatic method for road network extraction from high-resolution satellite images. The innovation in the proposed methodology is that the seed points converge onto the road candidate domain by a learning process using grayscale images, and the definition of the connection rules between the nodes. Finally, during the learning process, the learning rate factor $\alpha(t)$ is discussed in detail. The goal of the SOM process is to provide an approach to road centerline delineation from high-resolution imagery that is independent of conventional edge defection. The results of the road network extraction can be used for Geographical Information Systems (GIS) applications.

With the goal of evaluating the method's potential in extracting the road network, many simulations have been carried out using different types of images: small size (300×300 pixels) country road, overpass road images, middle size (500×500 pixels) country road images and large size (1000×1000 pixels) country road images. We can see that the seed points are shrunk onto the candidate domain of a road by a self-organizing map. At the same time, for the purpose of choosing a learning rate factor $\alpha(t)$, many kinds of different simulations have been also carried out using different types of images. The optimized initial learning rate factor α_0 is obtained, and good convergence onto the road candidate domain is shown. Then the seed points are connected using direction information and the road candidate domain according to the road tracking rules. However, when the initialization grid intervals are increased too much, there is an effect in the processing of images containing curved roads. As a result, because the seed points that lay on the candidate domain of the curved road are a few, it causes some difficulties in the reconstruction of the curved road network. However, in general, the results obtained are generally satisfactory in all cases.

But when there are missing road candidate domains, a part of the road network may be missed. As one of the image understanding tasks, road network extraction has been one of the most challenging research topics in GIS. Future developments

will focus on the automatic road network completion methodologies. Special attention must be devoted to road network extraction in urban area from satellite and aerial imagery. Moreover, better road network descriptions can be achieved by performing more tests using different types of imagery.

Chapter 5
Conclusions and Future

The purpose of the present thesis is to demonstrate that road network extraction system can also be regarded as application in Intelligent Transport Systems (ITS). In conclusion, a system capable of performing the task of road extraction and road network extraction from satellite images have been designed. Here outlined a large research field about road network extraction, which be wholly covered in a single dissertation. In this thesis, a study of Geographical Information Systems (GIS) with high-resolution satellite images is presented.

Chapter 2 summarizes the experience of previous road extraction. The resolution of the images has a strong influence on the representation of both roads and other objects, which influences the methods for road recognition. Different approaches used for road recognition in images of low- and high-resolution are discussed. Another important factor for road extraction is the interaction between the extraction algorithms and human operators.

Chapter 3 presents a semi-automated approach, based on a heuristic fuzzy logic system, to extracting the main suburban roads in IKONOS images, and describes how a system of different high-resolution satellite images, acquired under different

conditions, can be used to detect and recognize road segments based on fuzzy reasoning approaches. From the obtained extraction results, we find that forests and large area fields are clearly distinguished from roads, and only the main road candidate regions taken from the images were obtained. The results show that the approach is valid for roads extracted from high-resolution satellite images.

Subsequently, a semi-automatic method for road network extraction from high-resolution satellite images is also developed in Chapter 4. The first step is detecting the seed points in candidate road regions using a Kohonen-type self-organizing map (SOM). The second step is searching for connected points in the direction and candidate domain of a road by an approach to road tracking. With the goal of evaluating the method's potential in extracting the road network, many simulations have been carried out using different types of images: small size ($300\times$ 300 pixels) country road, overpass road images, middle size (500×500 pixels) country road image and large size (1000×1000 pixels) country road image. At the same time, for the purpose of choosing a learning rate factor $\alpha(t)$, many kinds of different simulations have been also carried out using different types of images. The optimized initial learning rate factor α_0 is obtained, and good convergence onto the road candidate domain is shown. The results of the road network extraction can be used for Intelligent Transport Systems (ITS) applications.

As one of the image understanding tasks, road network extraction has been one of the most challenging research topics in GIS. Moreover, special attention must be devoted to the flexibility of the heuristic system and road network extraction in urban area from satellite and aerial imagery. For example, there are perturbations, such as the shadows of buildings in the satellite and aerial images (shown in Fig. 1.2). The algorithm outlined in this thesis is merely a first step in this direction. Many further steps will have to be taken to develop the algorithm more fully. Therefore, future works will focus on the automatic road network completion methodologies. In addition, better road network descriptions can be achieved by performing more tests using different types of imagery.

Bibliography

1. Agouris, P., Gyftakis, S., and Stefanidis, A. [1998]: "Using a Fuzzy Supervisor for Object Extraction Within an Integrated Geospatial Environment." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part III/1. pp.191-195.

2. Agouris, P., Gyftakis, S., and Stefanidis, A. [2001a]: "Dynamic Node Distribution in Adaptive Snakes for Road Extraction." In: *Vision Interface, Ottawa, Canada.* pp.134-140.

3. Agouris, P., Stefanidis, A., and Gyftakis, S. [2001b]: "Differential Snakes for Change Detection in Road Segments." *Photogrammetric Engineering & Remote Sensing*, Vol. 67, No. 12. pp.1391-1399.

4. Bacher, U., and Mayer, H. [2005]: "Automatic Road Extraction from Multispectral High Resolution Satellite Images." In: Stilla, U., Rottensteiner, F., and Hinz, S., (Eds.): *CMRT05. IAPRS*, Vol. XXXVI, Part 3/W24, *Vienna, Austria.* pp.29-34.

5. Barzohar, M., and Cooper, D. B. [1996]: "Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation." *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 7. pp.707-721.

6. Baumgartner, A., Steger, C., Wiedemann, C., Mayer, H., Eckstein, W., and Ebner, H. [1996]: "Update of Roads in GIS from Aerial Imagery: Verification and Multi-resolution Extraction." In: *International Archives of Photogrammetry and Remote Sensing*, XXXI B3/III. pp.53-58.

7. Baumgartner, A., Steger, C., Mayer, H., and Eckstein, W. [1997a]: "Semantic Objects, and Context for Finding Roads." In: David M. McKeown, Jr., J. Chris McGlone, and Olivier Jamet. (Eds.): *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III, Proceeding of SPIE 3072.* pp.98-109.

8. Baumgartner, A., Steger, C., Mayer, H., and Eckstein, W. [1997b]: "Multi-Resolution, Semantic Objects, and Context for Road Extraction." In: Wolfgang, F., and Lutz, P. (Eds.): *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Basel, Switzerland.* pp.140-156.

9. Baumgartner, A. [1998]: "Extraction of Roads from Aerial Imagery Based on Grouping and Local Context." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. (32) 3/1. pp.196-201.

10. Baumgartner, A., Eckstein, W., Heipke, C., Hinz, S., Mayer, H., Radig, B., Steger, C., and Wiedemann, C. [1999a]: "T-REX: TUM Research on Road Extraction." In: *Proceeding of ISPRS Conference, Automatic Extraction of GIS Objects from Digital Imagery, Munich, Germany.* pp.37-58.

11. Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., and Ebner, H. [1999b]: "Automatic Road Extraction Based on Multi-Scale, Grouping, and Context." *Photogrammetric Engineering and Remote Sensing*, Vol. 65, No. 7. pp.777-785.

12. Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., and Ebner, H. [1999c]: "Automatic Road Extraction in Rural Areas." In: *International Archives of Photogrammetry and Remote Sensing.* pp.107-112.

13. Doucette, P., Agouris, P., Stefanidis, A., and Musavi, M. [2001]: "Self-organised Clustering for Road Extraction in Classified Imagery." *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 55. pp.347- 358.

14. Fischler, M., Tenenbaum, J., and Wolf, H. [1981]: "Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique." *Computer Graphics and Image Processing*, Vol. 15. pp.201-223.

15. Geman, D., and Jedynak, B. [1996]: "An Active Testing Model for Tracking Roads in Satellite Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 1. pp.1-14.

16. Gerke, M., Butenuth, M. and Heipke, C. [2003]: "Automated Update of Road Databases Using Aerial Imagery and Road Construction Data." In: *ISPRS Archives*, Vol. XXXIV, Part 3/W8, *Munich, Germany*. pp.99-104.

17. Gibson, L. [2003]: "Finding Road Networks in IKONOS Satellite Imagery." In: *ASPRS 2003 Annual Conference Proceedings, Anchorage, Alaska, USA*. (CD-ROM).

18. Gruen, A., and Li, H. [1995]: "Road Extraction from Aerial and Satellite Images by Dynamic Programming." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, No. 4. pp.11-20.

19. Gruen, A., and Li, H. [1997]: "Semi-Automatic Linear Feature Extraction by Dynamic Programming and LSB-Snakes." *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 8. pp.985-995.

20. Hahn, M. and Baltsavias, E. [1998]: "Cooperative Algorithms and Techniques of Image Analysis and GIS." In: *ISPRS Symposium on GIS – Between Visions and Applications, Stuttgart, Germany.*

21. Heipke, C., Steger, C., and Multhammer, R. [1995]: "A Hierarchical Approach to Automatic Road Extraction from Aerial Imagery." In: *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II, Proceeding of SPIE 2486*. pp.222-231.

22. Hinz, S., Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H., and Radig B. [1999]: "Road Extraction in Rural and Urban Areas." In: *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, SMATI'99, Munich, German.* pp.7-27.

23. Hinz, S., Baumgartner, A., Mayer, H., Wiedemann, C., and Ebner, H. [2001]: "Road Extraction Focussing on Urban Areas." In: Baltsavias, E., Gruen, A., Van Gool, L. (Eds.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Balkema Publishers, Rotterdam.* pp.255-265.

24. Hu, X., and Tao, C. V. [2003]: "Automatic Extraction of Main-Road Centerlines from High Resolution Satellite Imagery Based on Perceptual Grouping." In: *ASPRS 2003 Annual Conference Proceedings, Anchorage, Alaska, USA.* (CD-ROM).

25. Katartzis, A., Sahli, H., Pizurica, V., and Cornelis, J. [2001]: "A Model-Based Approach to the Automatic Extraction of Linear Features from Airborne Images." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 9. pp.2073-2079.

26. Klang, D. [1998]: "Automatic Detection of Changes in Road Databases Using Satellite Imagery." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32/4. pp.293-298.

27. Kohonen, T. [2001]: "Self-Organizing Maps." Third Edition, Springer, Berlin.

28. Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., and Saarela, A. [2000]: "Self organization of a massive document collection." *IEEE Transactions on Neural Networks*, Vol. 11, No. 3. pp.1-10.

29. Kumagai, J., Zhao, H., Nakagawa, M., and Shibasaki, R. [2001]: "Road Extraction from High-resolution Commercial Satellite Data." In: *the 22nd Asian Conference on Remote Sensing, Singapore.*

30. Lanser, S., and Eckstein, W. [1992]: "A Modification of Deriche's Approach to Edge Detection." In: *11th International Conference on Pattern Recognition*, Vol. III. pp.633-637.

31. 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 Applications*, Vol. 12, No. 1. pp.23-31.

32. Mayer, H., Laptev, I., Baumgartner, A., and Steger, C. [1997]: "Automatic Road Extraction Based on Multi-scale Modeling, Context, and Snakes." In: *International Archives of Photogrammetry and Remote Sensing,* Vol. XXXII, Part 3-2W3. pp.106-113.

33. Mayer, H., Laptev, I., and Baumgartner, A. [1998]: "Multi-Scale and Snakes for Automatic Road Extraction." In: *5th European Conference on Computer Vision, Vol. I, Freiburg, German.* pp.720-733.

34. McKeown Jr., D. M., and Denlinger, J. L. [1988]: "Cooperative Methods for Road Tracking in Aerial Imagery." In: *Proceedings of IEEE, International Conference on Computer Vision and Pattern Recognition.* pp.662-672.

35. Park, S. R., and Kim, T. [2001]: "Semi-Automatic Road Extraction Algorithm from IKONOS Images Using Template Matching." In: *22nd Asian Conference on Remote Sensing, Singapore.* pp.1209-1213.

36. Price, K. [2000]: "Urban Street Grid Description and Verification." In: *Processing of 5th IEEE Workshop on Application of Computer Vision.* pp.148-154.

37. Rosenfeld, A., and Kak, A. C. [1982]: "Digital Picture Processing." Volume 1, Second Edition, Rheinboldt, Werner (Eds.): *Computer Science and Applied Mathematics, A Series of Monographs and Textbooks*, ACADEMIC PRESS, INC. San Diego, California 92101.

38. Rosenfeld, A., and Kak, A. C. [1982]: "Digital Picture Processing." Volume 2, Second Edition, Rheinboldt, Werner (Eds.): *Computer Science and Applied Mathematics, A Series of Monographs and Textbooks*, ACADEMIC PRESS, INC. San Diego, California 92101.

39. Steger, C., Glock, C., Eckstein, W., Mayer, H., and Radig. B. [1995]: "Model-Based Road Extraction from Images." In: *Automatic Extraction of*

Man-Made Objects from Aerial and Space Images, Birkhäuser Verlag, Basel, Switzerland. pp.275-284.

40. Steger, C., Mayer, H., and Radig, B. [1997]: "The Role of Grouping for Road Extraction." In: Gruen, A., Baltsavias, E. and Henricsson, O. (Eds.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images II, Basel, Switzerland.* pp.245-256.

41. Vosselman, G., and De Knecht, J. [1995]: "Road Tracking by Profile Matching and Kalman Filtering." In: Gruen, Kuebler, Agouris, (Eds.): *Birkhäuser, Birkhäuser Verlag, Basel, Switzerland.* pp.265-274.

42. Vesanto, J., and Alhoniemi, E. [2000]: "Clustering of the Self-Organizing Map." *IEEE Transactions on Neural Networks*, Vol. 11, No. 3. pp.586-600.

43. Wiedemmann, C., Heipke, C., Mayer, H., and Jamet, O. [1998]: "Empirical Evaluation of Automatically Extracted Road Axes." In: Bowyer, K.J., Jonathon Phillips, P. (Eds.): *Empirical Evaluation Methods in Computer Vision. IEEE Computer Society Press, Silver Spring, MD,* pp.172-187. Also in: *Proceedings of 9th Australasian Remote Sensing Photogrammetry Conference., The University of New South Wales, Sydney,* Paper No. 239. (CD-ROM).

44. Wiedemann, C., Heipke, C., Mayer, H., and Hinz, S. [1998]: "Automatic Extraction and Evaluation of Road Networks from MOMS-2P Imagery." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32 (3/1). pp.285-291.

45. Wiedemann, C. [1999]: "Automatic Completion of Road Networks." In: *Proceedings of the ISPRS Joint Workshop "Sensors and Mapping from Space 1999". International Society for Photogrammetry and Remote Sensing.*

46. Wiedemann, C., and Hinz, S. [1999]: "Automatic Extraction and Evaluation of Road Networks from Satellite Imagery." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 3-2W5. pp.95-100.

47. Wiedemann, C., and Ebner, H. [2000]: "Automatic Completion and Evaluation of Road Networks." In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33, Part B3/2, *Amsterdam, Netherlands.* pp.979-986.

48. Zadeh, L. A. [1965]: "Fuzzy set." *Information and Control*, Vol. 8. pp.338-353.

49. Zadeh, L. A. [2002]: "Toward a Perception-based Theory of Probabilistic Reasoning with Imprecise Probabilities." *Journal of Statistical Planning and Inference*, Vol. 105. pp.233-264.

50. Zhang, C., Murai, S., and Baltsavias, E. [1999]: "Road Network Detection by Mathematical Morphology." In: *Processing of ISPRS Workshop on 3D Geospatial Data Production: Meeting Application Requirements, Paris, France.* pp.185-200.

51. Zhang C., Baltsavias, E., and Gruen, A. [2001]: "Updating of Cartographic Road Databases by Image Analysis." In: Baltsavias, E.P., Gruen, A., & Van Gool, L. (Eds.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Zurich, Switzerland.*

52. Zhang, C., Baltsavias, E., and Gruen, A. [2001]: "Knowledge-Based Image Analysis for 3D Road Reconstruction." *Asian Journal of Geoinfomatics*, Vol. 1, No. 4. pp.1-14.

53. Zhang, C., and Baltsavias, E. [2002]: "Edge Matching and 3D Road Reconstruction Using Knowledge Based Methods." In: *Institute of Geodesy and Photogrammetry, ETH Honggerberg, CH-8093 Zurich, Switzerland.*

54. Zhang, Q. [2006]: "Automated Road Network Extraction from High Resolution Multi-Spectral Imagery." In: *ASPRS 2006 Annual Conference, Reno, Nevada, USA.*

55. Zhao, H., Kumagai, J., Nakagawa, M., and Shibasaki, R. [2002]: "Semi Automatic Road Extraction from High Resolution Satellite Image." In: *ISPRS. Photogrammetric Computer Vision, Graz, Austria.* pp.406-411.

56. Zlotnick, A., and Carnine, P. D. [1993]: "Finding Road Seeds in Aerial Images." *Computer Vision, Graphics, and Image Processing: Image Understanding*, Vol. 57, No. 2. pp.243-260.

57. Yun, L., Uchimura, K., and Wakisaka, S. [2003]: "Road Extracted Using Fuzzy Reasoning From Satellite Images." In: *ASPRS 2003 Annual Conference Proceedings, Anchorage, Alaska, USA*. No. 0241. (CD-ROM).

58. Yun, L., Uchimura, K., and Wakisaka, S. [2004]: "Automatic Extraction of Main Road from IKONOS Satellite Imagery Based on Fuzzy Reasoning." In: Kumares C. Sinha, T. F. Fwa, Ruey L. Cheu, and Der-Horng Lee, (Eds.): *Proceedings of the 8th International Conference on Applications of Advanced Technologies in Transportation Engineering, Beijing, the P. R. of China*. No. 325, pp.641-646.

59. Yun, L., Uchimura, K., and Hu, Z. [2005]: "Application of Rule-based Fuzzy Logic in Road Extraction from IKONOS Imagery." In: *International Conference on Intelligent Computing, Hefei, the P. R. of China*. No. 1984, pp.443-452.

60. Yun, L., Uchimura, K., and Hu, Z. [2006]: "Application of Rule-based Fuzzy Logic for Road Extraction in IKONOS Imagery." *International Journal of Information Acquisition*, Vol. 3, No. 3. pp.1-10.

61. Yun, L., and Uchimura, K. [2007]: "Using Self-Organizing Map for Road Network Extraction from IKONOS Imagery." *International Journal of Innovative Computing, Information and Control*, Vol. 3, No. 3. pp.1-16.

62. Yun, L., Uchimura, K., and Hu, Z. [2006]: "Road Network Extraction Based on Self-organizing Map from IKONOS Imagery." 情報処理学会 高度交通システム研究会, No.2006-ITS-25 (3), pp.15-22.

63. 惲俐麗, 内村圭一, 脇阪信治 [2003]: "ファジィ推論を用いた衛星画像からの道路抽出に関する研究." 第 17 回熊本県産学官技術交流会講演論文集, pp.24-25.

64. 惲俐麗, 胡振程, 内村圭一 [2004]: "自己組織化地図を用いた道路ネットワークの抽出." 第 57 回電気関係学会九州支部連合大会, No.13-2P-20. (CD-ROM).

65. T.コホネン 著, 徳高平蔵, 大藪又茂, 堀尾恵一, 藤村喜久郎, 大北正昭 監修, [2005]: "自己組織化マップ." 改訂版, シュプリンガー・フェアラーク東京(Springer-Verlag Tokyo).

66. 田村秀行, [1985]: “コンピュータ画像処理入門.” 総研出版.

67. 長尾 真 監訳, [1978]: “ディジタル画像処理.” 近代科学社.

68. 日本ファジィ学会編集, [2000]: “ファジィとソフトコンピューティング ハンドブック.” 共立出版.