TECHNISCHE UNIVERSITÄT MÜNCHEN

Institut für Photogrammetrie und Kartographie

Lehrstuhl für Kartographie

A Congruent Hybrid Model for Conflation of Satellite Image and Road Database

Jiantong Zhang

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:

Univ.-Prof. Dr.-Ing. Uwe Stilla

Prüfer der Dissertation:

1. Univ.- Prof. Dr.-Ing. Liqiu Meng
2. Univ.-Prof. Dr. -Ing. habil. Dieter K. Fritsch
   Universität Stuttgart

Die Dissertation wurde am 04.02.2013 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 22.04.2013 angenommen.

To my parents, brother, sister, especially my wife and little daughter

whose love and support sustained me throughout

高山仰止,景行行止

–诗经

The mountain is so high that we look up with respects; the virtue is so great that we always take it as our example.

-The Book of Songs

Abstract

Spatial data conflation plays a fundamental role in many aspects of modern Geographic Information Systems (GIS) research and development such as geospatial data visualization, incremental updating of cartographic databases and disaster evaluation. The primary aim of conflation is to derive valuable information based on the comparison of multiple spatial data sources of homogenous or heterogeneous nature. This thesis is devoted to the conflation of two heterogeneous data sources - road network of a reference database and geo-referenced images where the systematic errors have been corrected in advance.

Due to various reasons, the spatial datasets of the same object type over the same geographic region reveal non-systematic misalignments or even different geometries; moreover, the object in one dataset may have no partner in the other dataset. The essential task involved in the conflation process is the correct identification of the inherent correspondences of objects between the two data sources which may have different cardinalities ranging from *1-1*, *1-0*, *1-n* to *n-m*. We present a Congruent Hybrid Model (CHM) to rectify the misalignments which could amount up to 200 meters when two datasets are overlaid. Theoretically, the optimal correspondence can be obtained by minimizing the matching cost between the centrelines of road features extracted from the image and the road network from the reference database. In the CHM which is a line-based conflation model, we also choose to optimize the matching cost at road intersections. The matched road network is then snatched to their counterparts in the image.

The main contributions of the CHM model include: 1) a linear feature extraction approach which consists of an Elastic Circular Mask (ECM) algorithm for feature extraction and a Genetic Algorithm (GA)-based grouping approach for the extracted road candidates; 2) a novel Sparse Matching Algorithm (SMA) approach which is especially useful to handle the problem of multiple correspondences; and 3) a performance evaluation of two popular transformation functions - the piecewise Rubber-Sheeting (RUBS) approach and the Thin Plate Splines (TPS) approach.

Through extensive tests which characterize various patterns of the road network in different scenarios, the CHM reveals its strength to deal with noise: redundancy can be reduced by 90% in ECM; by chaining the centrelines into longer features, GA can remove around 40% of the remaining redundancy in the results of ECM; SMA can handle the matching datasets with a sparse ratio of up to 14. In feature extraction, the correctness is improved around 34% from ECM to GA, and the completeness varies from 97%~193% in the experiments because of the over-extraction of the road features. The SMA reaches an effectiveness of at least 72.6% and a false ratio of at most 19.5% in our experiments, and the invisibility ratio of the roads between 1.1%~50.1% has been considered in the experiments.

The conflation quality relies very much on the importance of the transformation function as well, therefore, the author compares two popular transformation functions, which have slightly different performance in the experiments. Both functions have the ability to transform the feature to its counterpart properly. Therefore, the topology and geometry of the data sources can be well preserved in the conflated results. A main difference between the two functions lies in their behavior of handling context information: RUBS transforms the feature using the neighbors in the triangulation tessellation, whereas TPS considers the global influence of all the matched points in the test region. Another difference is caused by the ability to deal with the noise. The transformed results from RUBS are closer to the ground truth if the matching

quality is high, whereas the results from TPS have better performance if there are inaccurate matched points.

The experiments with synthesized data as well as the real spatial data have verified the time efficiency of the SMA, which has a stable performance; whereas the required time of ECM and GA is relatively sensitive to the amount of dataset. Comparing to the classical point-based conflation model, CHM can generate better results, especially with regard to the geometry, moreover, the inconsistency in both datasets can be also recognized and rectified when the accurate correspondence is derived from SMA.

The snake-based approach can be used to post process the results of the CHM, for instance, to rectify the remaining gentle misalignments. The CHM can be also directly employed for geospatial visualization applications.

Kurzfassung

Die Verschmelzung räumlicher Daten spielt eine wichtige Rolle in vielen Aspekten der Forschung und Entwicklung moderner Geoinformationssysteme (GIS), wie Geodaten-Visualisierung, inkrementelle Aktualisierung kartographischer Datenbanken und Katastrophenevaluierung. Das primäre Ziel der Verschmelzung ist die Ableitung wertvoller Informationen aus mehreren homogenen oder heterogenen räumlichen Datenquellen, Diese Arbeit ist der Verschmelzung von heterogenen Datenquellen aus Straßenvektoren und georeferenzierten Bildern gewidmet, in denen systematische Fehler im Voraus korrigiert wurden.

Aus verschiedenen Gründen zeigen räumliche Datensätze des gleichen Objekttyps über das gleiche geographische Gebiet einen Lageversatz bzw. nicht-systematische Verschiebungen oder sogar eine unterschiedliche Geometrie, zudem kann das Objekt eines Datensatzes keinen Partner in dem anderen Datensatz besitzen. Das Verschmelzungsmodell hat die wesentliche Aufgabe zur richtigen Identifizierung der inhärenten Korrespondenzen, die diverse Kardinalitäten bspw. *1-1*, *1-0*, *1-n* und *n-m* haben. Ein neuartiges *Congruent Hybrid Model* (CHM) wurde entwickelt, um die Verschiebungen, die bei der Überlagerung zweier Datensätze bis zu 200 Meter betragen können, zu beheben. Theoretisch lässt sich eine optimale Übereinstimmung erzielen wenn der Matchingaufwand zwischen den aus Bildern extrahierten Mittellinien und dem vorhandenen Straßennetz minimiert wird. Das CHM ist ein linienbasiertes Verschmelzungsmodell. Sein Matchingaufwand wird jedoch anhand entdeckter Straßenkreuzungen optimiert. Nach dem Matching werden die Straßensegmente zu ihren jeweiligen Partnern im georeferenzierten Bild transformiert.

Das CHM verfügt über drei Haupteigenschaften: 1) ein Linienextraktionsansatz, der aus dem Elastic-Circular-Mask (ECM)-Algorithmus zur Extrahierung der Straßenelemente und dem Genetic Algorithmus (GA)-basierten Ansatz zur Verkettung der auf einander folgenden Straßenelemente besteht, 2) ein neuartiger Sparse Matching Algorithmus (SMA) zur Lösung des Problems mit mehrdeutigen Korrespondenzen, und 3) ein Vergleich von zwei häufig verwendeten Transformationsfunktionen - der schrittweise Rubber-Sheeting (RUBS) Ansatz sowie der Thin-Plate-Splines (TPS) Ansatz.

In den umfangreichen Tests, die verschiedene Aspekte der Straße in verschiedenen Szenarien aufweisen, zeigt der CHM eine hervorragende Fähigkeit, dem Rauschen entgegenzuwirken. Die Redundanz wird um 90% durch den ECM reduziert und die Restredundanz wird um rund 40% durch den GA-Ansatz entfernt. Der SMA ist in der Lage, zwei Datensätze mit einer Sparse-Rate von bis zu 14 zueinander in Beziehung zu setzen. Bei der Linienextraktion wird die Trefferquote vom ECM zum GA um rund 34% verbessert. Die Vollständigkeit variiert zwischen 97% ~ 193% in den Versuchen aufgrund der Überextraktion linienhafter Elemente. Der SMA erreicht in den Experimenten eine mindeste Effektivität von 72,6% und eine höchste Fehlerrate von 19,5%, zudem wird in den Experimenten eine Unsichtbarkeit-Rate der Straßen von 1,1% ~ 50,1% in Betracht gezogen.

Die Qualität der Verschmelzung stützt sich zudem sehr stark auf die Transformationsfunktion. Aus diesem Grunde vergleicht der Autor zwei populäre Transformationsfunktionen, die leicht unterschiedliche Leistungen in den Experimenten erzielen. Beide Funktionen eignen sich die Objekte ihrer Korrespondenzen nach in geeigneter Weise zu transformieren. Dafür werden Topologie und Geometrie des bisherigen Straßennetzes in den Transformationsergebnissen erhalten. Ein Unterschied zwischen den

beiden Funktionen liegt in der Basisfunktion selbst: RUBS transformiert das Objekt in der Tesselation der Triangulation mithilfe seiner Nachbarn, während TPS den globale Einfluss aller übereinstimmenden Punkte in der Testregion berücksichtigt. Ein weiterer Unterschied ist durch die Fähigkeit mit dem Rauschen umzugehen verursacht. Die transformierten Ergebnisse vom RUBS sind bei der genaueren Matchingqualität genauer, während der TPS eine bessere Leistung zeigt bei einer ungenauen Übereinstimmung der Punkte.

Die Experimente mit synthetisierten Daten sowie mit den echten räumlichen Daten haben die Zeiteffizienz des SMA, das eine stabile Leistungsfähigkeit besitzt, verifiziert, während die erforderliche Zeit von ECM und GA relativ empfindlich auf die Datenmenge reagieren. Im Vergleich zu den klassischen punktbasierten Verschmelzungsmodellen kann das CHM bessere Ergebnisse erzeugen, insbesondere im Hinblick auf die Geometrie. Des Weiteren werden die Inkonsistenzen in beiden Datenmengen erkannt und behoben, wenn die genaue Korrespondenz aus dem SMA abgeleitet wird.

Der Snake-Modell basierte Ansatz ist eine Nachbearbeitungsansatz des vorgestellten CHM, um die übrigbleibenden leichten Verschiebungen zu korrigieren. Das CHM kann zudem direkt für Geodatenvisualisierungsanwendungen eingesetzt werden, wenn das Untersuchungsgebiet bestimmte spezifische Eigenschaften aufweist, die vom Modell ausgegeben werden.

Contents

List of Figures

List of Table

Chapter 1

Introduction

1.1 Motivation

The rapid development of efficient and inexpensive geospatial data acquisition techniques has led to an ever-growing number of geo-databases and vectorized geo-referenced raster images. Various aspects of modern GIS research and development such as spatial database updating, spatial data visualization and disaster evaluation, require an integration of geo-databases and images. The homogenous integration of a road database (street map) and Very High Resolution (VHR) image counts as a typical case characterized by the typical problem of the misalignment between the homologous features in the road database and the image, which may be induced up to 200 meters displacement (Chen et al., 2006). The feature misalignments between two source datasets are caused by various factors such as different resolutions or scales, inaccurate camera models, projection errors, different data acquisition times and different data vendors (Xiong and Sperling, 2004). Moreover, the misalignments are not systematic and therefore can not be corrected by a uniform transformation function. In this thesis we regard the vector road database as reference spatial data and the image as target spatial data. By features we mean the tangible geographic objects that have been digitized in vector, however, the road segments mean the factual road objects in the road database. We choose the term “conflation” rather than “integration” in this thesis following the previous contributions (Chen et al., 2006; Song et al., 2009; Wu et al., 2007), which aim to improve the feature accuracy through alignment. In Chapter 2 we discuss more definitions used in this thesis.

When overlaying two inhomogeneous road datasets describing the same region, the homologous road objects may reveal four different forms of offsets as illustrated in Fig.1-1 where the blue lines are manually generated ground truth of roads from the image and the red lines are roads in the database: (1) perfect coincidence; (2) intersection; (3) disjunction; (4) feature inconsistency. Fig.1-1 is selected from our test data, which will be also discussed in the experiment in Chapter 5. Fig.1-2 and Fig. 1-3 are selected from the two favourite online map providers (accessed on April 15, 2012), and the obvious misalignments are highlighted by the red lines, which have been slightly displaced purposefully from the original road in order to show the wrong overlapping problems. It should be pointed out that the misalignments appeared not only on large-scale maps, for instance the extreme offsets shown in Fig. 1-3, they are also common when the image resolution changes during zooming.

The conflicting spatial relationship, e.g. the overlap of a road vector with buildings in the image, is unacceptable for the visualization, it also dramatically influences the performance of many feature extraction algorithms. For instance, the snake-based detection algorithm assumes the prior information i.e. the reference data is sufficiently close to the target object; otherwise it is difficult to converge. It is also obvious for disaster evaluation: the results are totally wrong if there are some misalignments in the original datasets.

Fig. 1- 1 Four types of misalignments

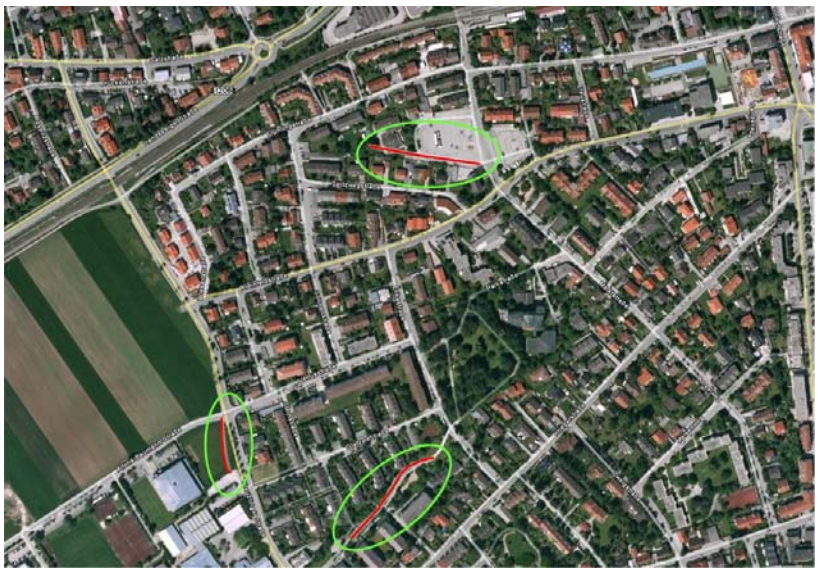


Fig. 1- 2 Misalignments in Google Maps® (48° 7'36.41"N 11°21'34.36"E)

Fig. 1- 3 Misalignments in NAVTEQ Map24® (48°10'2.04"N, 11°31'26.24"E)

Although the problem is straightforward when overlaying both datasets, detecting and correcting the wrong correspondences between an image and a road vector involves a manifold challenge:

1) Non-systematic misalignments: It is impossible to transform all features by one single global function. Instead, individual features in the datasets need individual treatments.

2) Uncertainty: Road features in an image are often disturbed by noise such as shadow, tree etc., which make them difficult to be recognized. Often, only fragments of a road can be detected from the image.

3) Inconsistency: The road features in the vector database are manually or semi-automatically created, therefore they have symbolized shape, especially at road intersections. However, the road features detected from the image only have rough shapes. This may increase the inconsistency between data sources.

These existing difficulties have been extensively evaluated in (Chen et al., 2006; Song et al., 2009; Wu et al., 2007) in the context of data conflation which usually includes three steps: 1) feature detection from image; 2) matching features from both data sources; 3) transformation of the matched features. Among them, the recognition of reliable features from the image plays a decisive role. Prominent road junctions are important clues and treated in the transformation function as Ground Control Points (GCP) in order to transform the reference features (roads in vector database) to the target features (the extracted features), where we assume the image has a higher actuality.

In spite of the existing misalignments, a reasonably correct overlay ratio is achievable in urban areas where many additional clues can support the matching work, for instance, many GCPs can help rectify the misalignments, therefore, reduce the difficulty of the conflation problem. In our study, we address the matching work for the suburban and rural area where both the clues from the context and the noise level are limited. We choose the Very High Resolution (VHR) aerial orthoimage with 0.4m resolution in which the individual features are well discernable. Feature inconsistency, i.e. feature in one database cannot find its complete correspondence in the other, is a main concern of the geo-data maintenance (Mena, 2003) and damage evaluation. This thesis is focused on solutions to a number of representative cases.

1.2 The Challenging Issue and the Goal

The conflation of road network and the image is an interdisciplinary problem, which can benefit from the achievements of different research areas. The essential task of conflation problem is to compute the reliable matching pairs from both spatial data sources. Two critical issues are involved:

1) Unreliable feature extraction from VHR image. Road features as well as the junctions are area objects on ground level which are likely sheltered by high-level features e.g. cars, trees. The most reliable features that are detectable in linear forms can be roads, building outlines or other open or closed linear features. The up-to-date road extraction algorithm (Franz, 2009) cannot provide reliable extracted road network because of the mentioned complexity. This has substantially impeded the research on conflation. Using the road intersection (Chen et al., 2006; Song et al., 2009) or the straightness of roads (Wu et al., 2007) for matching works only for certain areas.

2) Insufficient feature matching algorithm. The feature matching, for instance, point-to-point matching, requires the similar amount of feature in both data sources, otherwise, the matching results are unreliable or erroneous. Moreover, the definition of similarity measure is also a delicate problem for matching algorithm. The road-to-image conflation approach concerned in this thesis is a *1*-to-*n* matching problem where the reference road segments are relatively sparse as compared with the target features in the image and the correspondence could reveal *1-to-1, 1-to-n* or *m*-to-*n* relationship.

The road-to-image conflation approach faces not only the above-mentioned challenges, but also the increasing amount of new data. With the rapid progresses of spatial data acquisition technology, the updating period of data sources becomes much shorter than before. Manual conflation can no longer satisfy the demand of spatial data services. Without an automatic conflation method, many errors shown in Fig. 1-1~1-3 could not be efficiently corrected. If the misalignment between road network and image is larger than a certain threshold, some alignment algorithms cannot be applied.

The thesis aims to deal with the misalignments among the homologous objects between a reference road database and an image by developing an integrated system and verify its feasibility and efficiency with test datasets. The main problems which were considered in this thesis include: 1) a curvilinear feature extraction approach which assures the extraction rate and certainty; 2) a sparse matching approach to cope with the multiple correspondences; and 3) suitable transformation functions to snatch the matched objects from each other.

1.4 Organization of the Thesis

Following the introductory chapter, the theoretical background and related works are briefly summarized. The components of CHM and the experiments are addressed in respective chapters. The overall structure of the thesis is illustrated in Fig. 1-4.

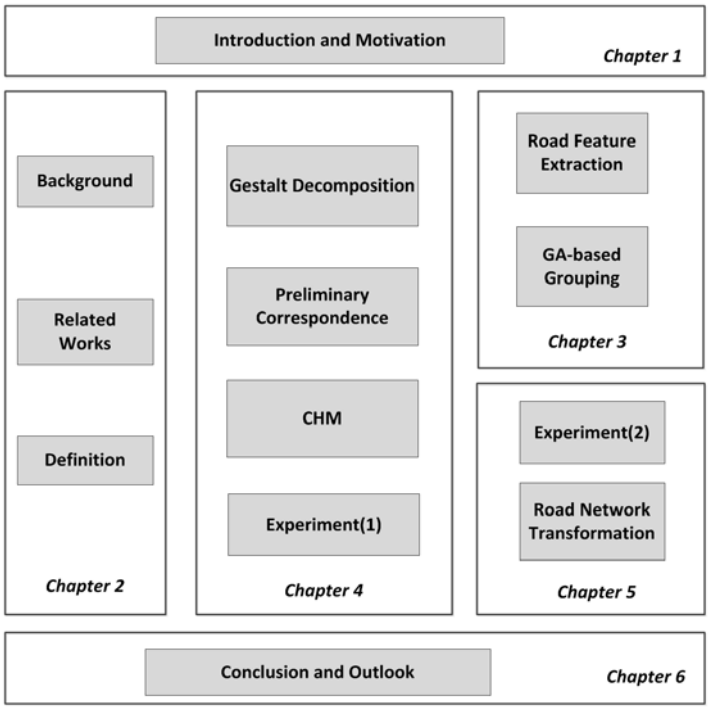


Fig. 1- 4 Organization of the thesis

In Chapter 2, the interdisciplinary characters of the conflation problem are comprehensively discussed. The definitions of the same concept from different discipline are combined for the image-to-vector conflation problem. The related matching algorithms are presented as elements of our conflation framework, but we focus on the matching approaches with available prior information. Moreover, the conflation approaches are classified on the basis of the matching strategies.

Chapter 3 is dedicated to the generation of the spatial data for the CHM model. The linear features are extracted from the image by means of the new detector - Elastic Circular Mask (ECM), and the primitives are further grouped using the Genetic Algorithm (GA)-based approach, which aims to reduce the noise in the extracted features.

The entire CHM model is presented and tested with synthetic datasets in Chapter 4. The noise from the inhomogeneous data sources is analysed. In CHM, the multiple correspondences between the two datasets are modelled as a mixed integer programming problem, but it can be simplified if the scope of the problem is limited, e.g. if there are less than 100 node groups in the experiment using synthetic datasets. The greedy idea is implemented in CHM to get the optimal matching pairs.

The experiments with the real spatial datasets are presented in Chapter 5 in order to verify the proposed model. Two favourite transformation functions are compared in terms of their suitability for road network.

The final chapter summarizes the advantages as well as the limitations of the proposed CHM model followed by a brief outlook about the emerging research issues for the future.

Chapter 2

Spatial Conflation: An Interdisciplinary Review

2.1 The Concept

2.1.1 Definition

The development of efficient and inexpensive geospatial data acquisition techniques has led to an ever growing number of geospatial data products and accelerated the comprehensive utilization of geospatial knowledge for our everyday life, for instance GoogleMap[1)], Nokia Map[2)], which provide us with efficient integrated spatial information e.g. street map, geo-referenced optical image or even street views from their portals. However, the multiple datasets bring about not only benefits, but also challenges, especially for the applications which require the combination of various spatial data sources. One conspicuous problem is the confliction between an existing spatial dataset of a region and a newly acquired dataset from the same region, because the one of the data sources usually has more information about the environment than the other, therefore, an integration approach for these more or less conflicting data sources is necessary to generate a better version than the original datasets. Data integration has been studied in several domains. In remote sensing technology, two geo-referenced satellite images can be registered, thus connected with each other; in medicine, datasets from Computer Tomography (CT) and Nuclear Magnetic Resonance (NMR) can be combined to obtain more complete information about a patient; in computer vision, a combined approach is commonly required to locate objects in an image; in cartography as well as in Geographic Information Systems (GIS), map updating also requires the integration of multiple data sources.

Despite the considerable amount of research works, the question about how well the applications can benefit from the new data as well as the archived data remains open. In order to address this question, we will review the prior knowledge-based spatial data integration approaches in Section 2.2. We will firstly introduce the concepts related with the integration from different research domains, which may bear unerring meanings in a specific research field but could sound confusing in other fields. On this basis, we then redefine the concept of spatial conflation to emphasize its characteristics in GIS community.

The term ‘object’ is perhaps most frequently used in GIS as well as image processing. It describes a tangible thing, e.g. a cup, a horse, a road etc. that you can touch or see. In previous chapter, we just mention the ‘cartographic object’, which indicates a concrete road, building and so on that needs to be shown on the map. The object bears the same meaning in image processing where an object corresponds to something with a specific size and form or shape. However, the detection of tangible objects from an image is a difficult task, up to now it’s still impossible to recognize the complete objects directly from noisy images. In most cases, only object parts such as segments of a road are directly detectable based on their geometric

1) http://maps.google.de/
2) http://maps.nokia.com/

or radiometric characteristics.

The term ‘feature’ indicates an important part of something that you notice because of its salience, relevance, or type. In the interdisciplinary research, feature is used to indicate object parts or groups, particularly when the objects are not so clear or they are hard to describe. In GIS, the feature layer is used to indicate the objects in the same thematic group such as the river, land parcel, road, etc., and it usually consists of geometric and semantic information. This thesis mainly deals with two different spatial datasets: the road networks and the road candidates extracted from a geo-referenced image. While road networks indicate exclusively the road objects, the extracted road candidates from an image may contain roads, road parts, other features or objects similar to or besides a road. To avoid confusion, we use in the thesis the term ‘edge’ or ‘primitive’ to describe the preliminary outcome from the image processing procedure, and the term ‘road feature’ to indicate the identity of ‘edge’ or ‘primitive’ that has been verified or assumed to belong to a road object. In this way, we discriminate the road from the road network, i.e. road object, road segment, and the road from an image, i.e. road feature, edge or road primitive, in the thesis. In Fig. 2-1, the terms, i.e. object and feature, are illustrated in two different spatial datasets.

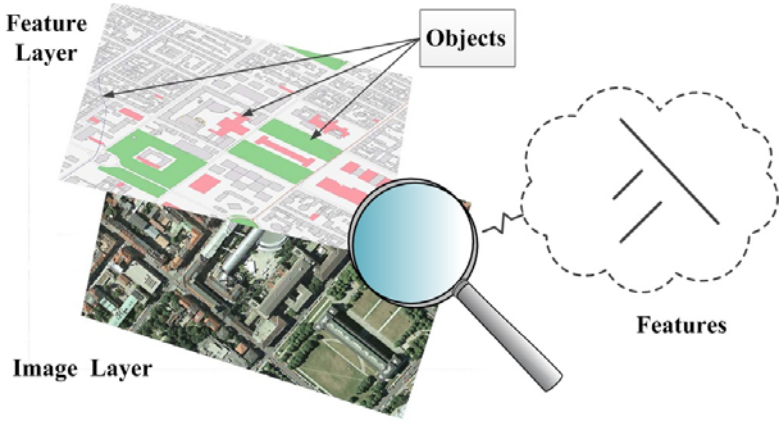


Fig. 2-1 Features vs. objects in spatial datasets

Both the term ‘segment’ and ‘component’ describe part of an object, but there is delicate difference between these two terms: the former emphasizes the part of a linear object, in other words, a linear object could be divided into many segments. The latter highlights a part of a whole object regardless of its shape or a member of a group. In our context, we prefer segment to component for the description of road network. If a road segment traverses between two road intersections, we also use the term ‘road arm’ to highlight the character of the road intersections.

Besides the definition of the object-related terms, it’s also important to understand integration frameworks from different domains, especially the difference between the image registration and the spatial conflation.

In remote sensing, medical imaging, computer vision, the concept of image registration is widely discussed for the fusion of two or more images from different times, different sensors or different perspectives. More than 1000 related papers were published between 1990s and 2000s according to the database of Institute of Scientific Information (ISI). The first comprehensive survey of registration was reported by Brown (1992), who defined the registration as a mapping function between two images with respect to the intensity. Zitova and Flusser (2003) reviewed the image registration works from interdisciplinary researchers,

and they reported the image registration as a process of overlaying multiple images which includes feature detection, matching and transformation. Finding and confirming the correspondence between the detected features from the sensed images is the critical step leading to accurate fusion results. Different from the data fusion which aims to generate the new data from multi data sources, image registration only computes the correspondences in images and transforms the matched features to each other.

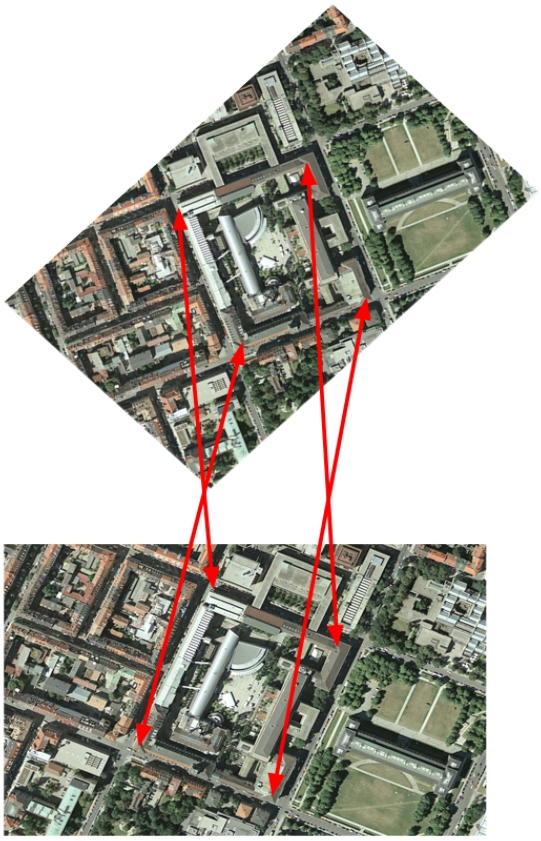


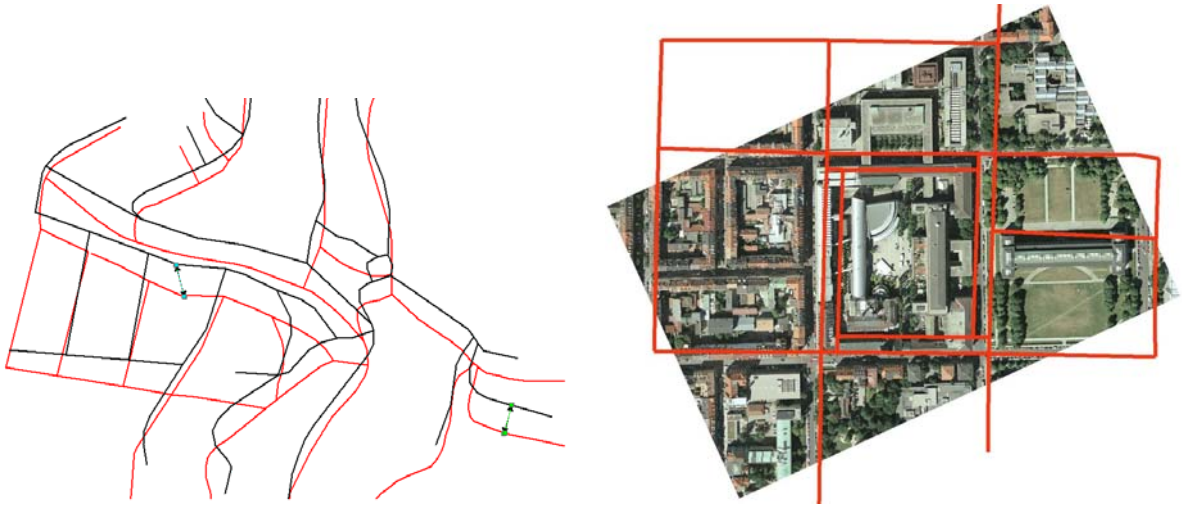
Fig. 2-2 Image registration by control points in order to correct the rigid deformation

There was no evident relation between the image registration and the conflation before 1990s when the satellite image was utilized to update or visualize the spatial data which was usually homogenous and from the same provider. Later, with the growing data sources from different providers, the requirements for the accurate correspondence or conflation between the image and existing spatial data have emerged. Two generations of conflation can be distinguished: 1) conflation of datasets with similar recording formats, for example, between two vector databases, or between two raster images; 2) conflation of datasets with different recording data formats such as between image and road network.

The term ‘conflation’ means to combine two or more text variants into one whole in rhetoric. Saalfeld (1985, 1988) coined the term ‘map conflation’ for geospatial data fusion or integration in the early 1980s to indicate the new generated map from two different map sources, although the integration of different spatial datasets was manually done before Saalfeld’s work. At that early stage of GIS development, huge geospatial datasets which usually originated from different agencies needed to be combined into one ‘better’ product. Lynch and Saalfeld (1985), Fagan and Soehngen (1987) and Lupien and Moreland (1987) also contributed to the first generation of conflation systems which specially dealt with the feature alignment problem in their research works. The up-to-date conflation process can be

characterized by: 1) the efficiency of combing or compiling a great number of digital maps; 2) the enriched features with the advantages from spatial data sources in similar or different recording data formats; 3) the rapid technological development and implementation of mathematical algorithms for recognition and matching of map objects.

The automatic or semiautomatic integration of spatial datasets from multiple sources has been an actual research subject since 1980s; moreover the new satellite images have been extensively used to update the topographic objects in the existing spatial databases since 1990s, which has extended the concept of conflation from similar to different recording formats of spatial data sources (Zhang, 2003). Mena (2003) reviewed around 250 related papers for updating the road database; Baltsavias (2004) reported the related works for cartographic object extraction using existing geospatial data as prior knowledge with the purpose of developing an operational system for updating the spatial database. The complexity of accurate detection of cartographic objects from the satellite image has stimulated extensive research works by numerous researchers from different research groups from photogrammetry, image processing and computer science. The main contribution of the second generation of conflation lies in the novelty to acquire homologous spatial objects from heterogeneous data sources, e.g. image, Digital Elevation Model (DEM) etc. Automatic or semiautomatic approaches have also been reported by Chen et al. (2006), Wu et al. (2007), Song et al. (2009), and Zhang et al. (2011). Recently, Ruiz et al. (2011) summarized the up-to-date research works with respect to the comprehensive classification of conflation processes which are mainly focused in the GIS domain, and they divided the conflation into two phases according to the works (Gillman, 1985; Gabay and Doytsher 1994): the identification of possible correspondences and the alignment of the matched features. Fig. 2-3 illustrates the two generations of the conflation approaches with two typical cases, where the road to be conflated is in red.



A) Vector (in red)-to-Vector (in black) conflation B) Vector (in red)-to-Image conflation

Fig. 2-3 Two typical cases of conflation

The fundamental concepts as well as the processes of the image registration are rather similar to those of the conflation: the correspondence and the transformation of the matched feature, if we neglect the data pre-processing. But it should be pointed out that the conflation also means to acquire more information from the partner spatial dataset to create a ‘better’ spatial data from the existing datasets regardless of their recording formats. From the point of

view of data process, we propose a general conflation framework with four phases: 1) data pre-processing; 2) correspondence computation i.e. by feature matching; 3) feature transformation; 4) quality evaluation. In practice, it is also reasonable to trigger only part of these processes to integrate the data sources, depending on the complexity of the original datasets. For the conflation of datasets in similar recording formats, the data pre-processing can be included (Zhang, 2009) or skipped over, but it becomes more complex and inevitable for the conflation of datasets with different recording formats where the high noise level may impede the detection of correspondences.

In this thesis, we build up multiple correspondences between image and road network by extracting the linear features from the image, and then model the spatial relationships as a mixed integer programming problem. In practice, the problem has been simplified, and a greedy algorithm is implemented to compute the estimated correspondences.

2.1.2 Strategies

The conflation involves a series of processes as discussed in previous section. Several factors may influence the organization of these processes into a feasible computational framework for specific spatial data sources. An appropriate strategy can accelerate the computational efficiency. In this section, we analyze the key factors involved in the processes in detail.

The pre-processing aims to assimilate the difference among the original datasets regardless of their recording formats. Its difficulty depends on the complexity of the input datasets. In case of similar recording data formats, e.g. the vector-to-vector or image-to-image registration, it needs less handling, whereas it requires more computing effort to normalize the datasets of different recording formats. The road features in the satellite image, for instance, need to be extracted in order to find their correspondences in the vector database. The low extraction ratio in the reported state-of-the-art road detection algorithms (Mena, 2003) is a bottleneck that makes it hard to provide adequate spatial data for the subsequent matching procedure. These algorithms suffer from the common problem to detect the road intersections so as to conflate the spatial datasets (Chen, 2005). The conflation with the purpose of integrating semantic information, e.g. post address (Zhang, 2006), or digital gazetteer (Goodchild, 2011), has to go through a similar process to normalize the data, but here the efficient rules instead of feature extraction are required. No matter what input datasets are chosen, the pre-processing has the key issue to set up effective criteria for the data harmonization so that the subsequent matching algorithm can be triggered.

Feature matching is the central process in the conflation framework, and it is also the decisive step to ensure the final quality of the conflated data. A lot of literatures on matching algorithms have been continuously reported in the last decade (Conte et al., 2004, Cui et al. 2009, Myronenko, 2010). The feature matching tackles two interdependent tasks: similarity computation and correspondence detection. The similarity measure can be cross-correlation, distance metric, turning function distance and so on between the homologous objects from both datasets and it decides the certainty of the matching results. The correspondence connects one data source (reference data) with the other (target data), and it is determined by the feature transformation functions. The general matching formulation between the reference feature *M* and its corresponding target feature *S* is*:*

$$\hat{f}(M,S) = \underset{f}{opt}\ simlarity(M, T(S,\Theta)) \qquad (2\text{-}1)$$

with $\hat{f}(\cdot)$ is the optimized correspondence value between the features; *T* is the transformation function; Θ is the parameter set in the transformation function.

As an important factor for matching, the similarity measure is the premise for the computation of the correspondence (Veltkamp, 2001), and we identify two different groups of measures: the feature-based measure and the intensity-based measure. The feature-based measure defines the similarity directly based on the geometric difference between the features, the Euclidian distance and its variants are the favourite measures because of the simplicity and intuition, furthermore, the distance-based measures can be generalized as (Chui and Rangarajan, 2003; Myronenko and Song, 2010):

$$E_{sim} = \sum_{k=1}^{K} \sum_{l=1}^{L} P_{kl} \left\| T(c_k^J) - c_l^I \right\|^2 \qquad (2\text{-}2)$$

herein E_{sim} is the degree of similarity between the feature c_k^J and its correspondent c_l^I; P_{kl} is the probability between the two features; T is the transformation function.

The intensity-based measure defines the similarity for the features with similar sizes, and the statistical comparison is the basis of this measure, which considers each individual element (pixel in case of image). The intensity-based measure has a higher computational efficiency as it does not require the feature detection process. The most popular intensity-based similarity measure is the Sum-of-Squared-Differences (SSD) (Hill et al., 2001):

$$E_{SSD} = \frac{1}{N} \sum_{n=1}^{N} (I_n - J_n^T)^2 \qquad (2\text{-}3)$$

where E_{SSD} is the degree of similarity between pixel I_n and J_n by the transformation function T; N is the number of pixels in the area, usually the same area from both datasets is chosen.

Some typical measures reported in the related literatures are summarized in Table 2-1 and Table 2-2.

Table 2-1 Feature-Based similarity measures for matching

	Feature-Based Measures
1	Discrete Metric
2	L_p Distance (Minkowski Distance)
3	Bottleneck Distance
4	Hausdorff Distance
5	Turning Function Distance
6	Frichet Distance
7	Nonlinear Elastic Distance
8	Reflection Distance
9	Earth Mover's Distance

Table 2-2 Intensity-Based similarity measures for matching

	Intensity-Based Measures
1	Cross-correlation
2	Phase-correlation
3	Mutual information
4	Area of Symmetric Differences
5	SSD

The transformation defines the method to deform or transfer the reference feature to its corresponding partner in the target using a rigid or non-rigid transformation function (Myronenko, 2010). Although the most common transformation function, e.g. rotation, translation, is perfect for rigid objects, and it is also simple to be implemented, it is the rare situation for the geospatial feature with rigid deformation. In the reality, most deformations are randomly distributed in the selected area; therefore, we have to consider both rigid and non-rigid transformations at the same time. It is a demanding task to discriminate and locate the rigid or non-rigid transformations for the matched features, and we compare two favourite transformation functions for the vector-to-image conflation problem in this thesis.

There are two typical strategies to transform the matched feature if we already know or assume it's a non-rigid deformation. The first one is the parametric approach which defines a series of transformation parameters from the matched features. The piecewise affine transformation is the most popular method in geoscience's research works because of its simplicity and it requires the subdivision of the space, e.g. by means of Delaunay triangulation or Voronoi diagram, and then the affine transformation is applied inside the regular grids. This approach is not smooth but continuous, so it's suitable for spatial features which usually obey the first law of geography (Tobler, 1970). It has been widely reported in conflation approaches (Ruiz et al., 2011). The first generation of conflation systems chose this method to transform the matched features (Saalfeld, 1985).

The Spline-based transformation function produces smoother transformation results than the piecewise affine approach and it is based on the linear combination of the basis functions:

$$f(x) = \sum_{k=1}^{K} w_k R(\|x - x_k\|) \qquad (2\text{-}4)$$

where w_k is the unknown parameter; the basis function R depends on the Euclidean distance from x to the given control point x_k, and K is the number of control points.

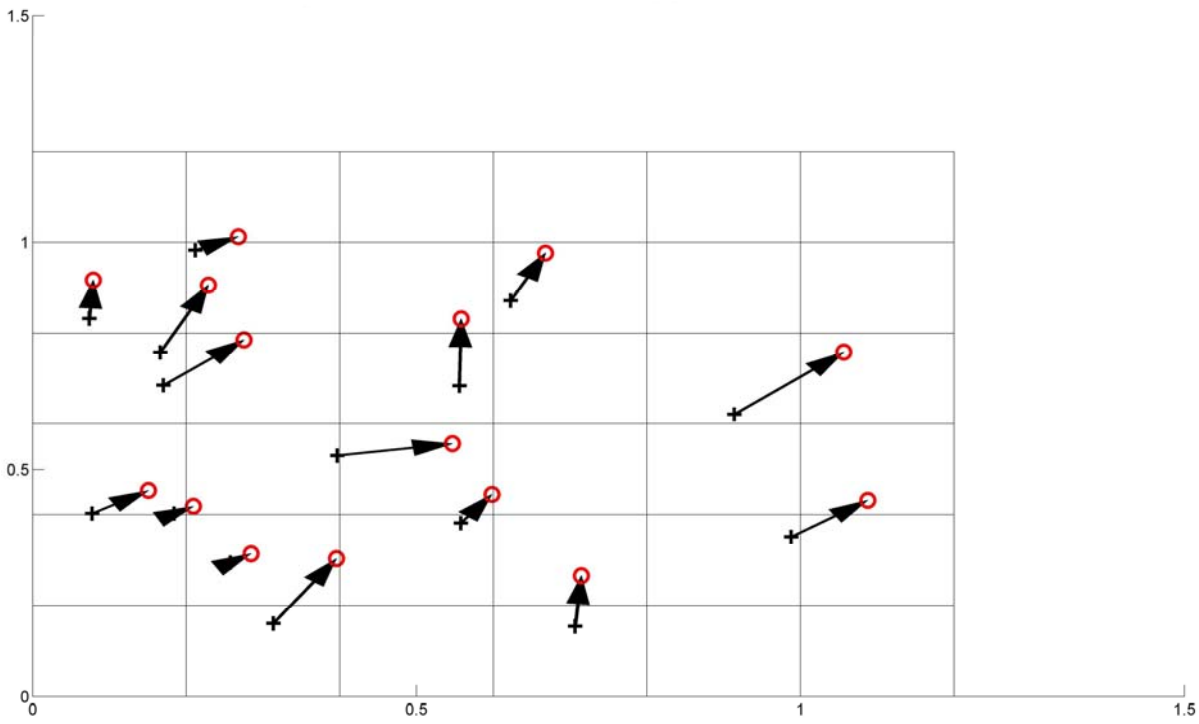
The Spline-based transformation function is a favourite method in medical image research, where the object needs smooth deformation; moreover, this method localizes the deformations with low computational complexity. Thin Plate Spline (TPS), for instance, is one of the Spline-based methods defined by (Bookstein, 1989).

$$f(x) = Ax + t + \sum_{k} w_k U(x_k, x) \qquad (2\text{-}5)$$

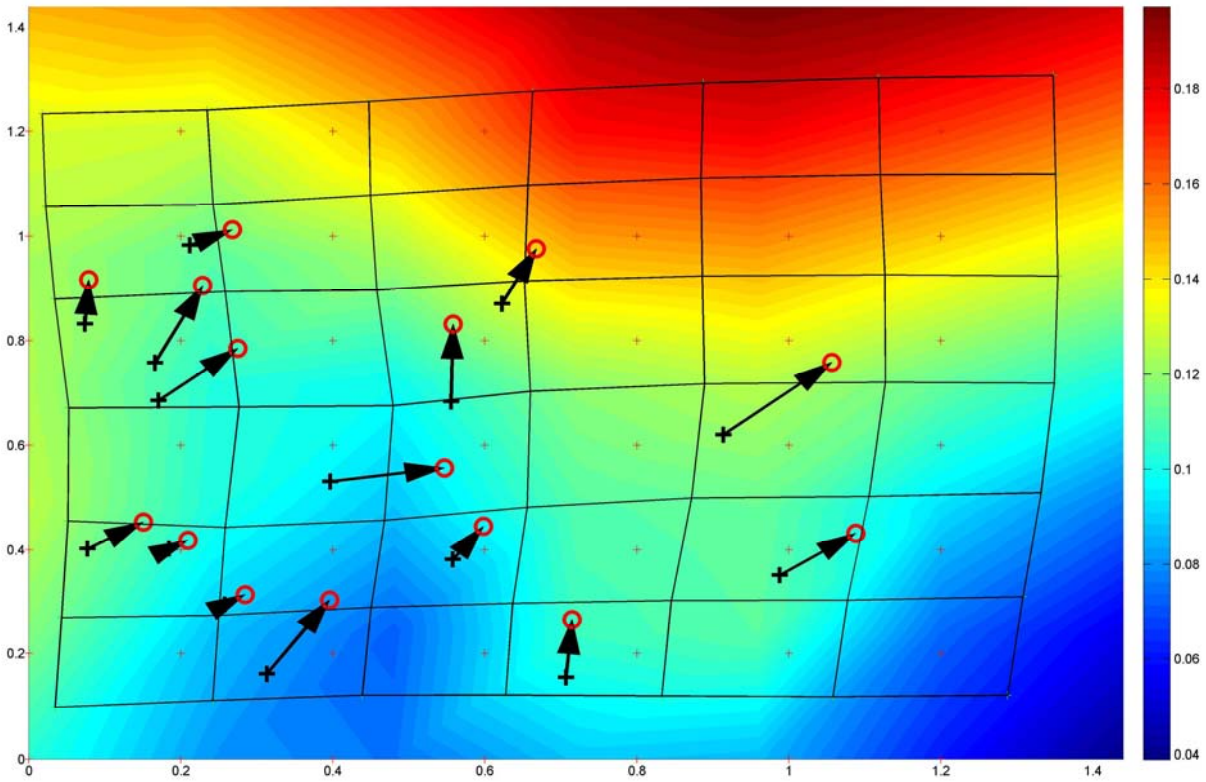
where A is the affine transformation matrix; t is a translation vector; w_k are the TPS coefficients and U is the TPS basis function. In 2D, $U = r^2 \log r^2$ where $r = \|x_k - x\|$ is the Euclidean distance from x to the given control point x_k.

The TPS method provides a global support for the matched features, which is preferred in many applications because it considers also the influence from features far away. However, this character makes the localization of the local transform difficult. Other Spline-based methods, e.g. cubic B-Splines, Multi-quadric Spline, can overcome the shortcoming of the TPS method to some extent, but they dramatically increase the computation complexity or they require prepared parameters.

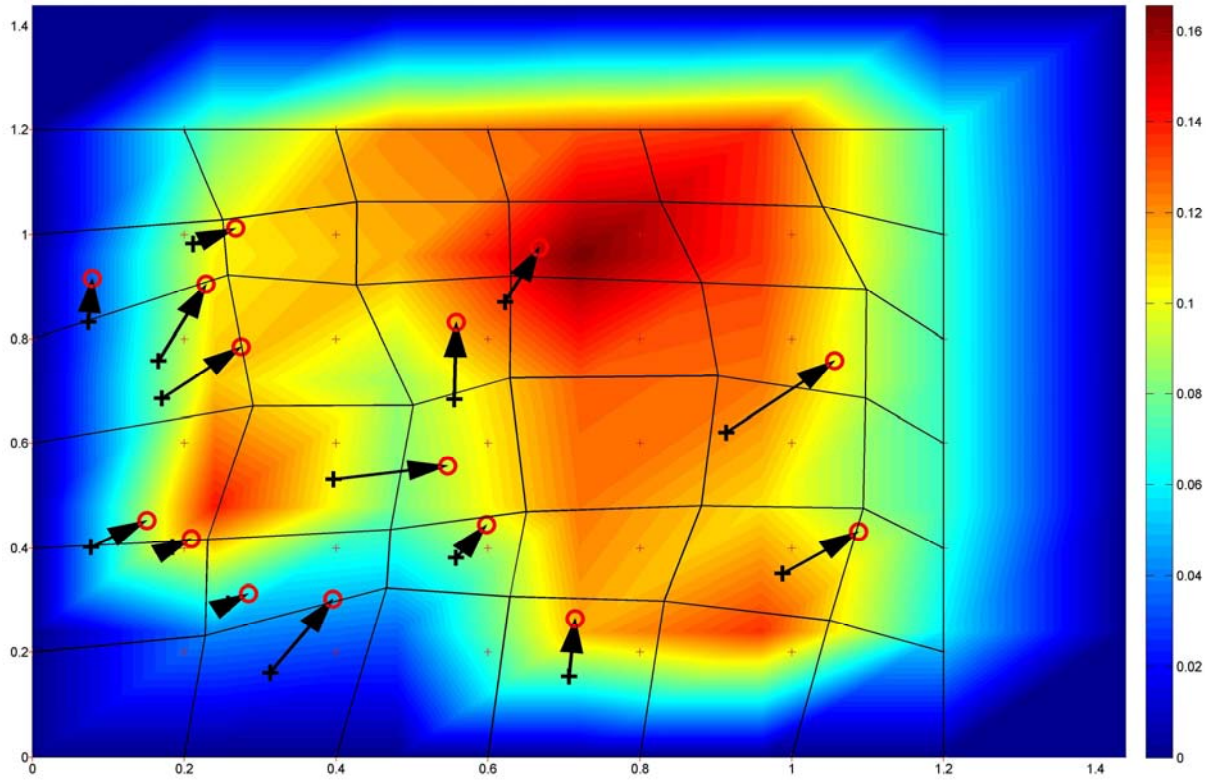
Both the smooth and non-smooth transformation functions have been used in the conflation applications for geospatial data (Wu et al., 2007; Song et al., 2009). Fig. 2-4 shows the main difference between TPS and piecewise affine transformation function such as the piecewise Rubber-Sheeting (RUBS) using the same control points. The reference data and target data are marked with '+' and 'o' respectively, and the colour bar in b) and c) is the deformation degree of the regular grid, where the degree has been interpolated with the consideration of its neighbors. The neighbor grids change smoothly in b), whereas the deformation in c) appears to be abrupt. The deformation in c) is mainly located around the point pairs, whereas the deformation in b) reflects more influence from neighbors. In the computation of c), we add the corners to the point pairs to get the deformation results near the border.



a) Point datasets in regular grids



b) The result from TPS



C) The result from RUBS

Fig. 2-4 Comparison of two different transformation functions

The second strategy assumes the transformation model as an unknown function to be optimized in terms of cost or target equation from the original datasets. The deformation can be defined as:

$$f(x) = x + T(x) \qquad (2\text{-}6)$$

where $T(x)$ is an unknown displacement for object x.

The target function for the conflation is defined as

$$E(T) = \underset{T}{opt}\left\{E_{sim}(I, J(x + T(x))) + \alpha E_{reg}(T(x))\right\} \qquad (2\text{-}7)$$

where $E(T)$ is the optimal result for the estimated displacement T; and E_{sim} is the similarity measure between the target dataset I and the reference dataset J for each object x and E_{reg} is the regularization term, and α is a regularization parameter.

Unfortunately, this function is an inverse problem (ill-posed problem) which has no unique solution from the matched features if there are outliers in the datasets. To get the constrained solution for the problem, for instance, it requires the regularization terms to suppress some undesirable results.

The optimization approach is advantageous due to its parameter-free computation procedure. It focuses on the optimal solution by the mathematical simulation such as soft assignment, relaxation etc. Furthermore, it doesn't directly require any correspondences in the matching results and can therefore match and transform the reference data to the target data at the same time. Its disadvantage is also obvious because it ignores the intrinsic characteristics of the data, although it can be used as a general solution.

Which of the above computing strategies is more suitable for the conflation depends on the nature of the original datasets as well as the requirements from the applications. The existing conflation approaches can be classified according to the characteristics of the spatial datasets as well as the conflation process.

2.1.3 Classification of conflation approaches

To understand the characteristics of various conflation approaches and to get an overview of their applications, a classification is necessary. Ruiz et al. (2011) comprehensively discussed the various conflation approaches according to four criteria: the matching strategy, the representation model, the categorization of matching results and the automation level.

We summarize only the approaches related with the vector-to-image conflation, but we also take into account the possible method from image-to-image registration without loss of generality. The reported conflation approaches can be differentiated from one another depending on three criteria: the geometric dimensions, the involved matching algorithms, the represent model. These criteria are interrelated in an integrated conflation framework.

The geometric dimensions of the feature are the most common criteria for matching as well as for registration approach. Points, lines and regions (polygons) are the favourite features because of their simplicity to acquire and handle. The point features consists of methods working with line intersections, road crossings, centroids of water regions, oil and gas pads and so on. The inherent advantage of the point feature is the simple choice of

similarity measure such as the variation of Euclidean distance. Point features are easy to acquire from vector data, but it is a challenging task to extract points or corners from an image. Line features can be the representations of general line segments, road network, object contours or elongated anatomic structures in medical imaging. The lines contain more geometric information than points, and the extra information from lines are the constraints for matching algorithm to get more accurate results. Line correspondence is usually expressed by pairs of line ends or middle points; however, line strokes have been verified as a more efficient means for road network matching (Zhang, 2009). The complexity of line-based matching method is not only caused by the difficulty to extract the lines from image, but also owing to the insufficient similarity measure to evaluate the similarity between two lines. Polygon feature is a closed-boundary region with an appropriate size, so it can be the projection of a lake, forests, a building or an urban area. The polygon has more geometric information for matching algorithm than point and line, but it requires a sufficient shape description which is not a trivial task.

The matching algorithm is the key component in a conflation application. Some novel conflation models are based on the new designs of the matching measures or procedures. Depending on the used mathematical method, a matching algorithm can be statistics-based, optimization-based or heuristic-based. The statistical-based matching method has many examples in several research fields incl. the GIS community. Walter and Fritsch (1999) proposed a buffer growing algorithm for vector-to-vector conflation base on the initial idea from the Iterative Closest Point (ICP). In principle, all of the graph-based matching algorithms (Conte et al., 2004) belong to this category. The optimization-based method regards the matching as an integrated mathematical problem, where the so-called target function needs to be set up to describe the similarity in the matching datasets, and then the optimal solution can be achieved via relaxation or linear combination method (Myronenko, 2010; Chui, 2001), moreover, the transformation function can be also integrated into the target function to simplify the procedure in conflation. In computer sciences, the heuristic-based matching method has been widely accepted and applied for object recognition. The shape context (Belongie et al., 2002), for example, is one of the well-known heuristic-based matching algorithm for object with deformable shape. The heuristic-based method usually needs some prior information about the object and a training procedure for the algorithm to learn the characters from the samples, therefore, the objects can be identified by the well-trained parameters. Many research works have proved that the heuristic-based method has a better performance in cluttered image than the classical shape-matching algorithm (David and DeMenthon, 2005; Ferrari et al., 2010).

The representation model in conflation approach is related to the fundamental aim. If the original data sources are in different representation models i.e. different formats, the initial step is to harmonize the heterogeneous data sources. In case of the vector-to-image conflation, the feature extraction from image and then matching the geometric feature is the common procedure. Walter and Fritsch (2000) presented a method to update the GIS database using satellite image, in which the vector data was converted to raster to compute the correspondence. The recording format can be in different formats such as vector data, raster, video sequence etc. Besides the recording format, the semantic-based model aims to use the semantic information to increase the accuracy of the matching results.

The applications of conflation approaches are mainly concentrated in geo-data updating, media technology, image processing or computer vision, and more recently also in disaster management.

2.2 Related Works

2.2.1 Road detection with prior information

Automatic detection and extraction of road objects from satellite image is one of the most challenging research tasks for photogrammetry and remote sensing in the last decade. Although progresses have been constantly made, the research questions involved in the detection task are far from being completely solved due to their intrinsic complexity. Among the large number of existing works, several reviews on the road detection algorithms made by (Fortier, et al., 2001; Mena and Malpica, 2004; Mayer, 2008; Rottensteiner, 2009) have been reported in different time periods. The approaches designed for the GIS data updating, are more interesting and also more closely related with our context.

Using satellite image to update the road database or using the road network to guide the road extraction can be traced back to 90's (Zhang, 2003; Mena, 2003). At that time both topics challenged the GIS community: updating the GIS data needs extensive manually work, whereas the low extraction ratio is the bottleneck of the image processing algorithm, so a combined approach of road network and image is a natural choice to increase the performance. Each dataset is the meaningful prior information for the other dataset. Stilla (1995) built up a map-aided analysis system to accurately locate the manmade objects in the aerial image using cartographic information. Zhang (2003) reported a map-based road detection method, which included two important applications: updating the GIS data (Klang, 1998; Fiset et al., 1998; Fortier et al., 2001), and validating or increasing the road detection ratio (Cleynenbreugel et al., 1990; Gunst, 1997; Bordes et al., 1997; Agouris et al., 2001). Mena (2003) reviewed the related research efforts from around 250 references for the updating of the GIS data and tried to classify the algorithms for road extraction and identified the challenges involved in the road detection using both vector and raster datasets. Baltsavias (2004) discussed the importance of prior knowledge embedded in multiple data sources for the geospatial objects extraction. He found out that the prior knowledge was quite hard to be represented and managed in spite of its usefulness claimed in many research works. Moreover, he pointed out the unsatisfactory performance of object extraction method. From our point of view, the achievements in the reported works are reflected in three aspects:

(1) The active contour model

The active contour model or snake-based approach (Kass et al., 1988) for the integration of GIS data and the satellite image provides an algorithm that takes the GIS data as the prior information and bends towards the feature of the image by minimizing the energy function in Eq. 2-9. If we define the parametric curve *C*

$$C(s) = (x(s), y(s)) \qquad (2\text{-}8)$$

where $s \in [0,1]$ is the arc length, and *x* and *y* are the coordinates of a closed curve *C*. The curve delineates the object on the image by setting and minimizing the energy function *E(C(s))*

$$E_{snake} = \int_0^1 \left(E_{\text{int}}(C(s)) + E_{image}(C(s)) + E_{con}(C(s)) \right) ds \qquad (2\text{-}9)$$

with 1) $E_{\text{int}}(C(s))$: the prior knowledge concerning the shape and movement of the object;

2) $E_{image}(C(s))$: the optimal description of object in the image;

3) $E_{con}(C(s))$: the additional external constrained force.

Klang (1998) detected the changes between the existing road database and the newly acquired satellite image using Ziplock Snakes, and the new roads which are not included in the existing road database are extracted by a simple tracking algorithm after detection of the seed points. In a similar way, Agouris et al. (2001) introduced the Differential Snakes for road change detection, and the local and global changes were identified. Bonnefon et al. (2002) presented a complete work to update geographic linear features in GIS, which could quickly find the optimal path by the dynamic programming, but there were unclear regions due to the noise, and then the snakes smoothed the unclear sections and controlled the curvature. Peteri and Ranchin (2003) proposed two different types of active contours for extracting the road network in dense urban areas. Based on the given topologically correct graph of the network, the road segments were reconstructed by introducing the parallelism constraint for the contour, whereas the intersections still used the original active contours. Youn and Bethel (2004) found that the road network had usually a block pattern, and they employed the active contour models to refine the initial approximations from the segmentation, and the road corners were rectified finally. The original snake-based concept was designed for enclosed object such as buildings, land parcels etc. Its usage was quite limited for open curves, e.g. road and river. Keeping this in mind, Butenuth (2008) proposed Network Snakes to extend the original snakes for delineating arbitrary curves. Rochery et al. (2006) also circumvented the limitation of the original concept by introducing the 'high-order terms' in the energy function for an arbitrary topology, and Peng et al. (2010) further extended the model to extract the road network from VHR satellite image, and the additional nonlinear and linear non-local terms were introduced for general road networks, e.g. straight narrow branch, sharp curves, dead end roads.

An obvious drawback of the snake-based approach is its sensitivity to noise. It may easily get stuck in the local minima during the computation (Amini, 1988). A more critical problem is that there is only a regularized solution, e.g. simulated annealing algorithm, for the target function which requires enormous amount of computation time. A further problem related to conflation task is the requirement of the snake-based approach for the prior information which should be very close to the features in the image. This proves rather critical for the road detection from the satellite image which could contain the misalignment up to 200 meters (Chen et al., 2006). Therefore, narrowing the misalignment between the two datasets is necessary (Song et al., 2009).

(2) Detection of road feature

Extraction of the entire road network from an image is a nearly impossible mission. It is constrained not only by the object model as the abstraction of the real world, but also by the cluttered scene, where the roads as the low-level objects could be easily shaded by the trees or shadows as shown in Fig. 2-5.

Road intersection, for instance, is a salient part of road and also the helpful information for detecting the whole road network (Negri et al., 2006; Hu et al., 2007), and the intersection-based algorithms reported in (Deschenes and Ziou, 2000; Barsi and Heipke, 2003; Ravanbakhsh et al., 2007) explored the characters of road intersections for the purpose of locating them in the image or in the scanned topographic map (Chiang et al., 2008). The intersection-based algorithms suffer from the disadvantage of their computing complexity to localize the intersections within large areas. Due to the recognition problem the existing

algorithms are not able to produce efficient solutions. We regard this situation as a useful hint for the conflation and are aware of the flaws in the extracted road data and the necessity of including some prior information for accurate road detection.

Fig. 2-5 A typical image with roads as low-level objects at a intersection

(3) Quality assurance of spatial data

A common step of conflation and updating of GIS data from image is to evaluate the quality of the extracted spatial data in comparison with the prior knowledge, which has been reported as a non-trivial problem (Gerke and Heipke, 2008). It's an important issue for the detection of correspondences and changes. Lemmens and Verheij (1988) reiterated the importance of the GIS knowledge for the image segmentation problems. Later, Gunst and Hartog (1994) and Gerke et al. (2004) verified the GIS data by modelling the prior knowledge as a spatial context. The graph-based verification approach (Gerke and Heipke, 2008) increased the detection performance of the road extraction algorithm, but it depended on to some extent the topology of the extracted features which may be unstable for complex scenes in image. Moreover, the combination of evidences is a tricky task. Based on the prior spatial knowledge, the evaluation of the extracted features should consider the broken features as well as their topology which may reveal complicated spatial relationships between the original object and the extracted feature. In case of road objects, there may be various cardinalities of correspondences, e.g. 1-to-0, 1-to-1 or 1-to-n from the extracted features.

From the literatures, we have also learnt a number of lessons. First, we are aware of the fact that it is unrealistic to extract the entire road network from the satellite image, even if the images are accurately geo-referenced. The related works with or without prior knowledge (Fortier et al., 2001; Rottensteiner, 2009; Mena, 2003; Baltsavias, 2004) show that the

performance of road detection is still quite limited, the detected features are noisy and not sufficiently reliable in comparison with the objects in the GIS database in spite of the great progresses reported by Poullis and You (2011). The fundamental progress of road detection is accompanied with by the algorithmic development for the detection of primitives such as edge or contour (Papari and Petkov, 2011). Second, the combination of multiple data sources is not always helpful for the road detection as we expected. Hinz and Baumgartner (2003) suggested the road extraction from multiple perspectives of the aerial image. The identification of the homologous object in different views is even more challenging than the recognition of the road, although the detection of 3D objects as a by-product proves interesting for many applications. Zhang (2004) proposed the approach to detect the road network from the image with the prior information of GIS data and DSM, and the combination of different quality measures for various data sources is a tedious endeavour before they can be used as prior knowledge. Similarly, the combination of SAR or laser image suffers the lack of representation criteria of the prior knowledge in addition to the difficulties of recognizing and localizing the homologous object in a different data source.

2.2.2 Feature-based matching algorithms

In Section 2.1.2, two general concepts for feature matching - similarity measure and correspondence are introduced. In this section we focus on the matching algorithms for the conflation of image and road network with the emphasis on feature-based approaches for points and lines. The semantic information embedded in road networks serves as the complementary information for the features in the image.

(1) Point-based matching

Point features have been most frequently treated in matching methods in a wide range of disciplines as they are relatively easy to be detected from the image. Nevertheless, many reasons make the point-based matching a complicated problem. Based on an analysis of progresses made in point-based matching with the specific focus on the deformation level as well as the 'dynamic' character of the data, Li et al. (2003) claimed the primary reasons which strongly influenced both accuracy and efficiency of matching are the data quality and the distortion level which can be subjected to data distortion, incompleteness and noise from different systems. Zitova and Flusser (2003) reviewed the point-based matching approaches for the image registration. Zhang (2009) regarded point-based matching approaches as a support to a network-matching algorithm. Among these works, ICP-algorithm (Besl and McKay, 1992) has been reported as the most popular method for point-to-point matching. It deals with an iterative procedure that computes the favourable correspondences with the initial matching by the assumption of knowing the prior information e.g. the similarity or transformation parameters.

Algorithm 2-1 ICP-Registration (Sharp et al., 2002)

Let *S* be the reference point set $\left\{s_1, s_2, \cdots, s_{P_s}\right\}$ ($P_s \in \mathbb{R}$), and *M* is the target data set $\left\{m_1, m_2, \cdots, m_{P_m}\right\}$ ($P_m \in \mathbb{R}$), and $D(s_i, M)$ is the shortest distance from M to the reference point s_i, and the algorithm starts:

1) Let T_0 be an initial parameter for the transformation;

2) Setup the correspondence function: $C = \cup_{i=1}^{P_s} \{T_{k-1}(s_i), D(T_{k-1}(s_i), M)\}$;

3) Compute the new transformation T_k that minimizes the mean square error between point pairs in C;

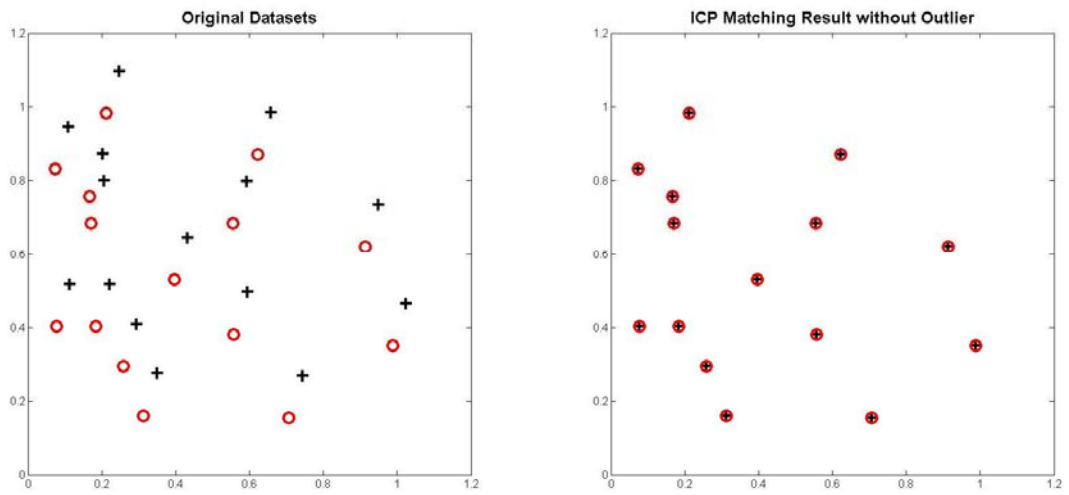
4) Repeat the Step 2~3 until met the termination criteria.

ICP utilizes the nearest-neighbor relationship to assign the correspondence at each step. It requires a good initial estimation in order to converge to the global minimum. ICP is simple and speedy, therefore, suitable for real-time applications. However, it is a difficult task to determine reasonable initial parameter. A fully automated ICP algorithm needs to search among multiple initial conditions which may increase the computing effort.

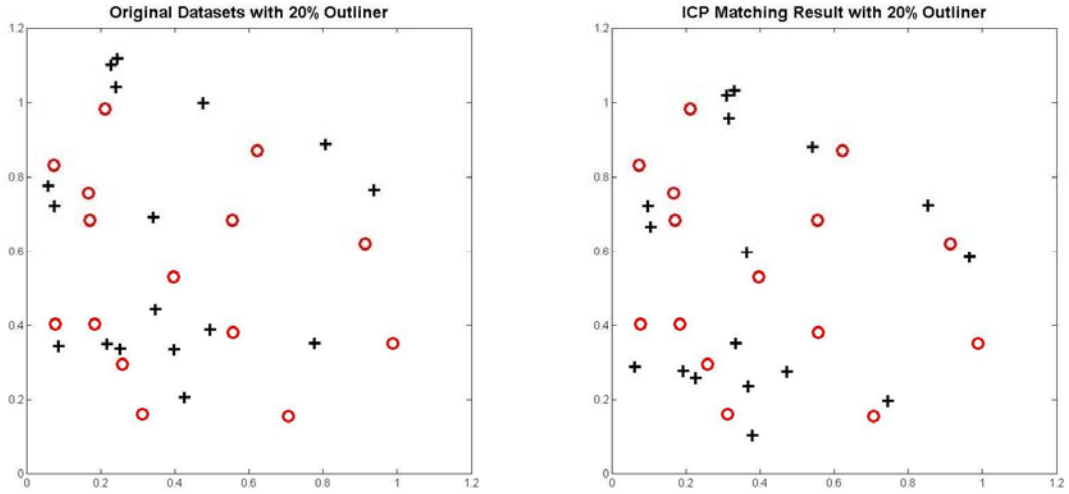
To find a desirable initial state of ICP, (Chui, 2001) and (Myronenko, 2010) formulated in their methods of TPS-PRM (Thin Plate Spline Robust Point Matching) and Coherent Point Drift (CPD) the point matching as an optimization problem which usually includes a target function. Unfortunately, the target function is an ill-posed problem because of large number of the unknowns in the function and it needs regularized term or constraint to get the optimal results. Moreover, the physical property of the spatial point is neglected. Apart from this drawback, these approaches are robust and general for different kinds of applications.

The point-based matching is a practical approach for point dataset, where the spatial relationship among the points is undefined or difficult to detect. The distance measure is a congenital deficiency of the point-based method. The nearest neighbor metric or its variation (Veltkamp, 2001) has been verified as a suitable measure for dense point dataset. The dense point dataset, however, is an unrealistic requirement for cartographic objects which are usually irregularly and sporadically distributed within the study area. Another challenge for the point-based matching method is caused by the outliners in the dataset and the noise may dramatically decrease the stability of the matching results if the noise ratio becomes larger in the target dataset.

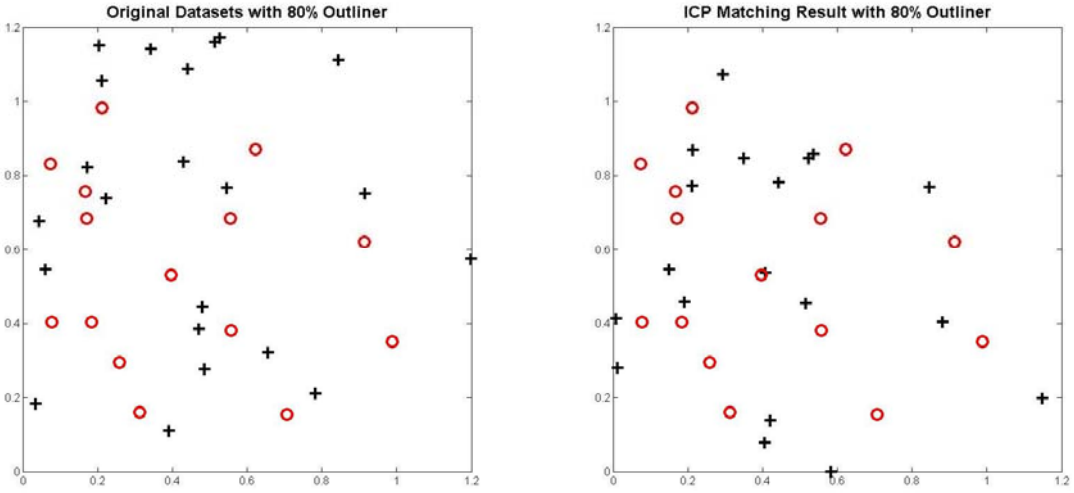
In Fig. 2-6 a), the reference dataset '+' and the target dataset 'o' have the perfect matching where there are no outliers and each dataset includes 15 points. But the matched results in Fig.2-6 b),c) become unstable when noisy points are added to the reference dataset with 20% (2 points) and 80% (12 points) respectively. Although many measures have been taken to improve the robustness against the outlier (Li et al., 2011), the fundamental idea remains the same: compute the similarity from the nearest pairs and minimize the overall matching errors. To reduce the influence from the outliers, more geometrical constraints among the points can be added, which has been typically considered as a line-based matching approach.



a)



b)



c)

Fig. 2-6 Performance of ICP algorithm for point-based matching with and without noise

(2) Line-Based Matching

Line-based matching approach takes more geometric information into account than the point-based matching method and is widely used for road network matching. The geometric, topologic and semantic characteristics of two road networks are useful clues for the establishment of an exact connection between them. The semantic information serves as auxiliary information (Zhang, 2009), while geometric and topologic information is more commonly explored. The line-based matching algorithms work on road networks with complete or incomplete topology.

The topology bears the most significant information about a road network. A road segment is also an elementary component of the graph. Therefore, line-based matching is mostly supported by the graph theory. Conte et al. (2004) reviewed the interdisciplinary graph matching approaches with regard to the applications and the algorithms. They deemed the graph matching as a general searching problem with the aim to refine the correspondence among the line segments.

Different from the graph matching method, the road network matching plays a significant role for the integration of geospatial data and considers more information about the roads and their catchments' areas. Among the related works, two algorithms are quite compelling. Walter (1997) proposed the buffer growing algorithm to identify correspondences inside a local region with a specific buffer threshold. Multiple correspondences are possible. However, the false correspondence can be minimized or verified as the buffer region moving to the neighbor segments. Zhang (2009) presented a line-based matching algorithm which helps construct delimited strokes from the road segments, and then introduce a buffer around the individual strokes other than the fragmental road segments. In further steps, the multiple correspondences are identified and optimized by means of a context-related topologic analysis.

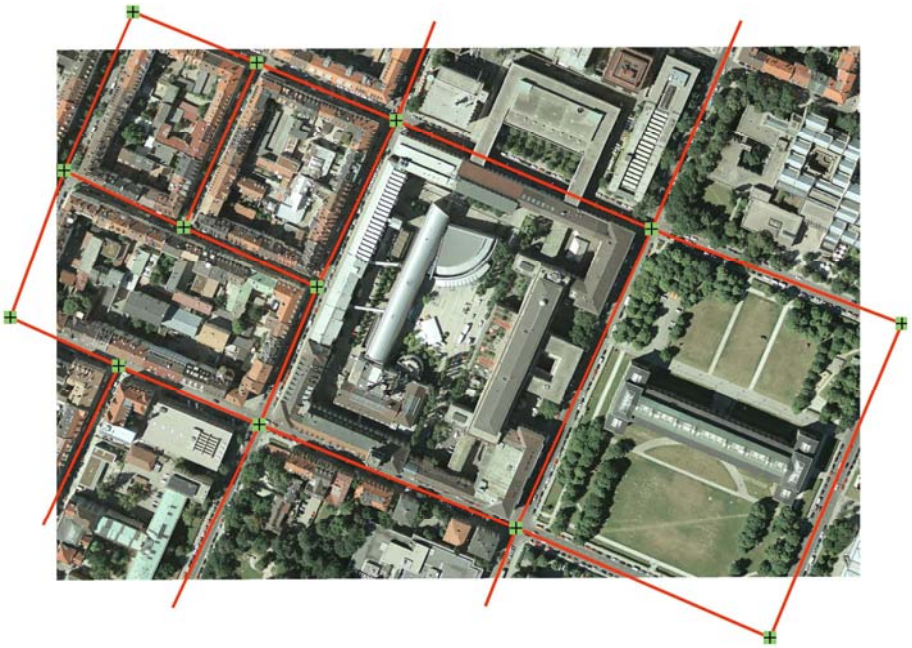


Fig. 2-7 The reliable topology of the road segments (red lines)

Fig. 2-7 shows the manually generated road segments and intersections from an image. The topology of the lines could also be incomplete. The extracted road features from the

image typically reveal some gaps between the line segments as shown in Fig.2-8. This may introduce uncertainty in the matching results, and may require to identify the correspondence between open lines (Cui et al., 2009). If the line segments in the target dataset are broken, the connection of the segments is not reliable. Whether it is possible to get the exact correspondence depends not only on the matching method, but also on the effective filter which directly or indirectly eliminates the noise in the dataset. It is essentially same as the traditional contour-based object recognition task which has to locate the prior shape such as cup or horse among the massive contour elements. A pre-processing is required. For instance, the perceptual grouping uses the convexity, cocircularity, connectness, parallelism or proximity to group the line candidates as a meaningful object. The local edge-based cliques (Ferrari et al., 2008) are also identified as reliable element for object detection. The contour-based object recognition algorithms usually deal with only one single object with well-defined boundary, and there are still no reports about their application for open line detection from satellite images.

Fig. 2-8 The unreliable topology of the extracted road lines (the green lines)

No matter whether the topology is reliable or not, the line is the fundamental element for the matching algorithms. Besides the matching measures introduced in Section 2.1.2, various approaches utilize different spatial relationships among the lines which are extensively discussed in (Walter, 1997; Parent and Spaccapietra, 2000; Safra et al., 2006; Mustiere and Devogele, 2008; Zhang, 2009; Zhang, 2011).

The spatial relationships of the lines are conceptually classified into four groups, i.e. one-to-null, one-to-one, one-to-many, and many-to-many (Zhang, 2009). The matching algorithms have the essential task to search the entire data space, detect the spatial relationships between homologous line features based on the established similarity measures. Usually, the determination of the spatial relationships starts from the line elements in the more reliable dataset. The correspondences for the one-to-one relationship may fall into one of 6 cases if we only consider the geometric information (Fig.2-9).

1) Zero correspondence

Due to different qualities of datasets to be matched, there are certain features without exact corresponding pairs in the partner dataset, for instance, the new road could exist only in one of the datasets. It is even worse for the conflation between road network and image where some roads are neglected by the detection algorithm.

2) Complete correspondence

This is the ideal but rare situation for the matching approach. It is the most reliable matching result if both lines in the networks reveal exactly the same geometry.

3) Overflow correspondence

The target line is longer than the reference line, thus contains both of the end points of the reference line. The overflow correspondence is the counterpart for the containment correspondence.

4) Extension correspondence

If one of the end points of the reference line can get the exact matching result in the target line and the other end point is also verified on the target line, it is the extension correspondence, which means the target line can be extended to have the same length as the source line.

5) Containment correspondence

The reference line is longer than the target line and contains both end points of the target line. It is the counterpart of the overflow correspondence.

6) Partial correspondence

It is common that there is no exact relationship as described above if it only considers the end points of the line, but at least one end point can be registered inside the reference line.

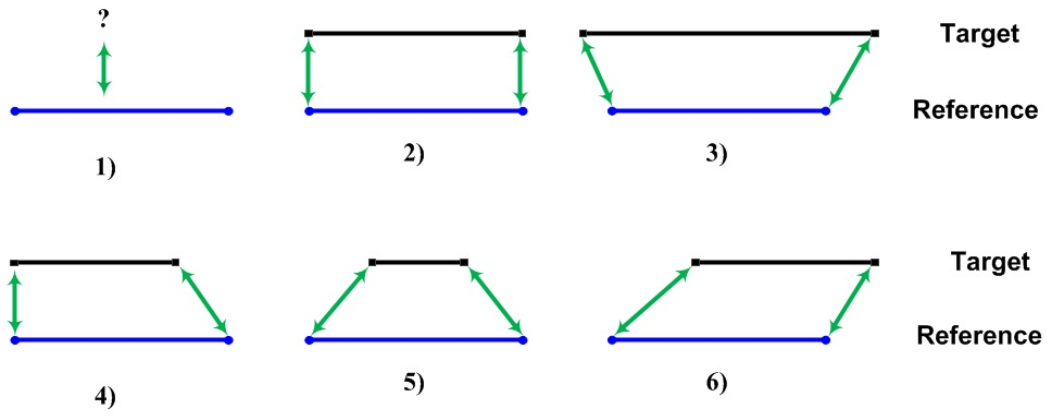


Fig. 2-9 Six variants of one-to-one spatial relationship between two lines

Only two lines with one-to-one relationship are considered in the above discussion. We can extend the discussion to complex situations that deal with spatial relationships among multiple lines. The correspondence 4)~6) usually involves more than one single line segment, if it takes the neighbors into account. Therefore, as shown in Fig.2-10, it evolves to a 1-to-*N* or *N-to-1* spatial relationship, where *N* is a positive integer. More generally, if we continue with the searching step, we may identify all the *N*-to-*M* relationships, where *M* is a positive integer that is not necessarily equal to *N*.

Using the matching results to set up unreliable connections with the target dataset is another aim of this thesis, and we will discuss more about the spatial relationships as well as the sparse line matching algorithm in Chapter 4.

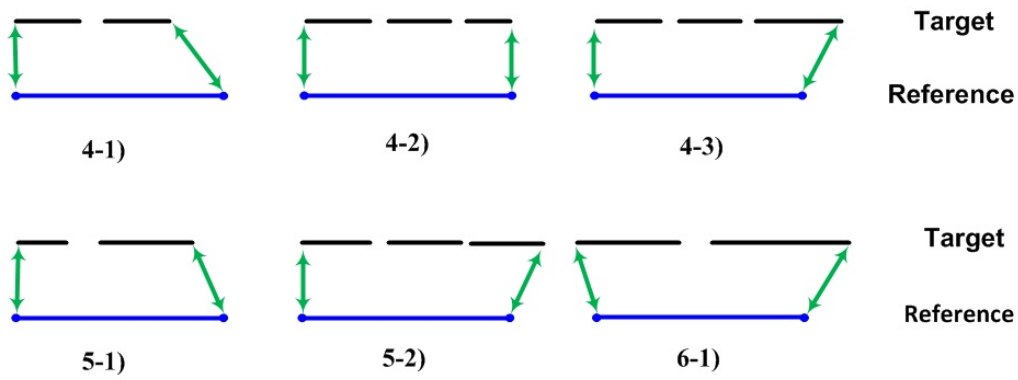


Fig. 2-10 Examples of one-to-many spatial relationship for lines

(3) Patch-Based Matching

The enclosed meshes or polygons by road segments also provide important clues for matching algorithm, and we term the approach as patch-based matching. It follows the principle of finding the appropriate shape descriptor or indicator and computing the similarity between two patches. Since road networks contain open lines, patch-based matching can only partly work on closed meshes in road networks. A review with general discussion about the patch-based matching method can be found in (Zitova and Flusser, 2003).

2.2.3 State of the art of conflation

In the two preceding sections, we reviewed some key issues of conflation process. This section is devoted to an overview of the most important progresses of conflation between image and road network. A complementary overview is given by (Ruiz et al., 2011).

Hild and Fritsch (1998) introduced a global alignment method between vector and images for geocoding according to which the vector features are rasterized to match the extracted polygons from the image. Chen et al. (2006) presented a conflation approach based on road intersection patterns recognized by using Bayesian classifier. They analysed the shape of the grayscale histogram and determined clusters on the shape using a size measure with the aim to detect road intersections that reveal certain statistical properties. However, the majority of image regions do not really display the statistical uniformity because of the remaining noise from small objects. To solve the problem, Ruiz et al. (2011) developed a non-parametric approach based on texture analysis for the identification and extraction of pixels that belong to road intersections.

Wu et al. (2007) reported an approach to localize the global alignment problem by dividing a large image into regular grids, and then computing the maximum correspondence between the road networks and the features in each image grid, and the matched features were transformed by the global TPS transformation function. The correspondence between the image and road network is identified by the peak responses from the corresponding gradient image, moreover, it assumed that there was a dominating orientation in each pre-tiled image. This assumption, however, proves too strict for the road pattern in the image, and is

only suitable for certain cities. The approach also computed the confidence factor for the correspondence, thus provided useful information for manually correction.

Song et al. (2009) presented a fully automatic conflation approach based on the road intersections as well as terminations which were extracted from imagery and compared with road networks by means of a relaxation labelling algorithm. The matched points were snapped using a Rubber-Sheeting transformation algorithm. Furthermore, the aligned features were input into the snake models to acquire the refined results. In this approach, the intersections were picked up from the imagery by setting suitable thresholds for Normalized Difference Vegetation Index (NDVI). The spatial context which assumed the road surface with a similar radiometric behaviour in the test area was considered. However, it was hard to decide the threshold for the image: A small threshold may increase the accuracy, but miss some important intersection; whereas a too loose threshold may increase the number of false intersections in the results.

These approaches share a common character of avoiding the detection of linear feature and using road intersections as the favourable salient features from the image. They have therefore a high computing complexity. Moreover, they ignore the shape of the road and presume the road segments are more or less straight lines, which may lead to inaccurate results after the transformation (Zhang et al., 2011). A combination of road intersection and salient road segments is therefore a more desirable approach, but it requires the extraction of linear features from the image.

2.3 Complexity of Spatial Conflation

2.3.1 Non-rigid deformation

The spatial datasets to be merged or conflated are usually collected from different times, different sensors and/or by different vendors, and there is inevitably some inconsistency between them. The same object e.g. the road segment in one dataset appears different in other datasets and has undergone some non-rigid deformation. If we consider all the features at the same time, we have to face an infinite number of unknown transformation parameters (Woods et al., 1998) because the solutions usually exist in an infinite space (Yoo, 2004). To make the non-rigid transformation manageable, we may look for discrete solutions instead (Zagorchev and Goshtasby, 2006). However, it is difficult to define the exact number for a sufficient estimation of the parameters, moreover, large transformation errors may be introduced if a feature to be matched is outside the convex hull of its matching partner.

Both continuous and discrete solutions can be reasonable for spatial conflation of road networks, but their relative efficiency and complexity need to be clarified in transformation applications. It should also be noted that the deformations in datasets or the misalignments between two road networks are non-systematic and therefore it is unrealistic to transform all features by one global function. Instead, each feature needs an individual treatment in an optimal way.

2.3.2 Incomplete matching measures

The matching measure decides the certainty of the correspondence. The common measures for matching approaches introduced in Section 2.1.2 are not sufficient for spatial features. For instance, the ICP chooses the closest point as the matching measure, but it can be misleading for well-structured spatial features. The distribution patterns of spatial objects

with natural or artificial structure are too diverse to be modelled in a single way. No versatile measures exist yet to effectively describe the similarity between the features and preserve their inherent characteristics. To choose or (re)define such measures remain the demanding task for matching approach.

Both the correspondence and the transformation parameters are unknowns of the matching system, but they are interconnected with each other (Chui, 2001). If we know the exact correspondence, the transformation is a simple least-square problem, and vice versa. However, the high dimensional parameter space for both correspondence and transformation has complicated the handling of the non-rigid feature matching approach. In practice, it is only possible to get approximations by means of additional constraints or regularized parameters.

2.3.3 Outliers (Robustness, Uncertainty)

One of the important preconditions of the matching approach is the equal or similar density of the datasets to be matched. However, it is not always the case for spatial data conflation, especially for the matching between road network and satellite image. From the discussion in Section 2.2.1, the road features in image are often disturbed by noise such as shadow, tree etc., which make them difficult to recognize and only fragments of road are discernable from the image. The performance of matching algorithm may drop further with the increasing noise or outliers in the dataset. This may lead to unstable and uncertain matching results. Under extreme conditions, for example, if the discrepancy between the two datasets is beyond a reasonable range, no reasonable solution can be achieved.

The matching or conflation quality relies obviously on the data quality. This thesis is focused on the strategy of increasing the stability of the matching approach, not the extraction ratio of the features from image.

Chapter 3

Curvilinear Road Detection and Denoising

3.1 Objectives

The decisive factors of the conflation approaches are the matching strategies and the quality of the spatial datasets. The feature-based strategy, for instance, relies on accurate geometric information. Both the point- and line-based matching approaches make use of the geometric or topological characteristics to compute the correspondences. In this chapter we discuss the generation of spatial data supported by our hybrid conflation approach. Two data sources - orthoimage and road networks are considered.

The point-based conflation approach which inherits the advantages of point matching algorithm has proved suitable for the road-to-image conflation in numerous practical applications (Liu and Wei, 2004; Jian and Vemuri, 2005; Chen et al., 2006; Song et al., 2009). The point clusters can represent different shapes around road intersections in 2D or even in 3D environments. Moreover, the point-based approach avoids the extraction process of line features which may miss many reliable road segments, but the detection of point features from the image is a challenging task which runs the risk of introducing much noise to the matching algorithm. The regularization treatment is therefore necessary in order to get reasonable results (Myronenko, 2010). The point-based approach has another advantage of having simple correspondences. A point from one dataset has either no or only one counterpart in the other dataset.

The line-based matching method is a natural extension of the point-based approach. However, the quality of the extracted lines influences dramatically the performance of the matching algorithm. As we have discussed in Chapter 2, different solutions have been presented in several research works (Barrow et al., 1977; Borgefors, 1988; Walter, 1996; Conte et al., 2004; Wu et al., 2007; Zhang, 2009). The line-based approach has the advantage of utilizing the accurate topological information which enables an incremental matching strategy, and it is suitable for road features spreading over the entire area and for the matching algorithm with homogenous data sources, for instance, road network to road network matching approach.

However, for heterogeneous data sources, it is difficult to acquire the reliable road topology from the image, and current feature detection approaches can only provide edges of road and need information from the partner dataset to increase the matching reliability. One of the challenging problems is related with the integrity and uncertainty of the datasets: the road features in the vector database have correct topology but ‘wrong’ locations in the image; whereas the extracted features from image have correct position but carry little geometric information and much noise. Therefore, the vector database and its image have to mutually benefit from each other in order to correct the misalignments and reconstruct the geometry of the extracted features.

Keeping the importance of the spatial datasets in mind, we generate two different spatial datasets from the image and land parcels respectively to conflate with the road from a navigation database which has a high topological accuracy. The extracted data from land parcels has better topology than the extracted features from the satellite image, and the proposed conflation model will be tested using both datasets so as to verify the feasibility of the model.

3.2 Road Extraction from Geo-Referenced Image

3.2.1 Extraction of primitives from image

In spite of the considerable progresses in recent research works and the widespread availability of Very High Resolution optical image, the road detection from the image remains a challenging task (Rottensteiner, 2009). The various existing approaches share the common techniques that make use of spectral, geometric, topologic, and contextual properties of roads (Fortier et al., 2001). The linear operator, e.g. Steger operator (Steger, 1998), Canny operator (Canny, 1986) is usually the starting step for the extraction of linear primitives followed by a grouping procedure based on geometric criteria such as width, length and curvature (Wiedemann and Hinz, 1999) or spectral criteria such as the NDVI (Lacroix and Acheroy, 1998). But it should be pointed out that the line operators have limited performance when they only work with low-level vision models where the roads with two parallel edges in the noisy image can be hardly detected. Therefore, clues from high-level vision models, such as shape model etc. are required (Papari and Petkov, 2011). Instead of pursuing a perfect road detector, we propose an Elastic Circular Mask (ECM) detector to generate the road candidates as shown in Fig.3-1.

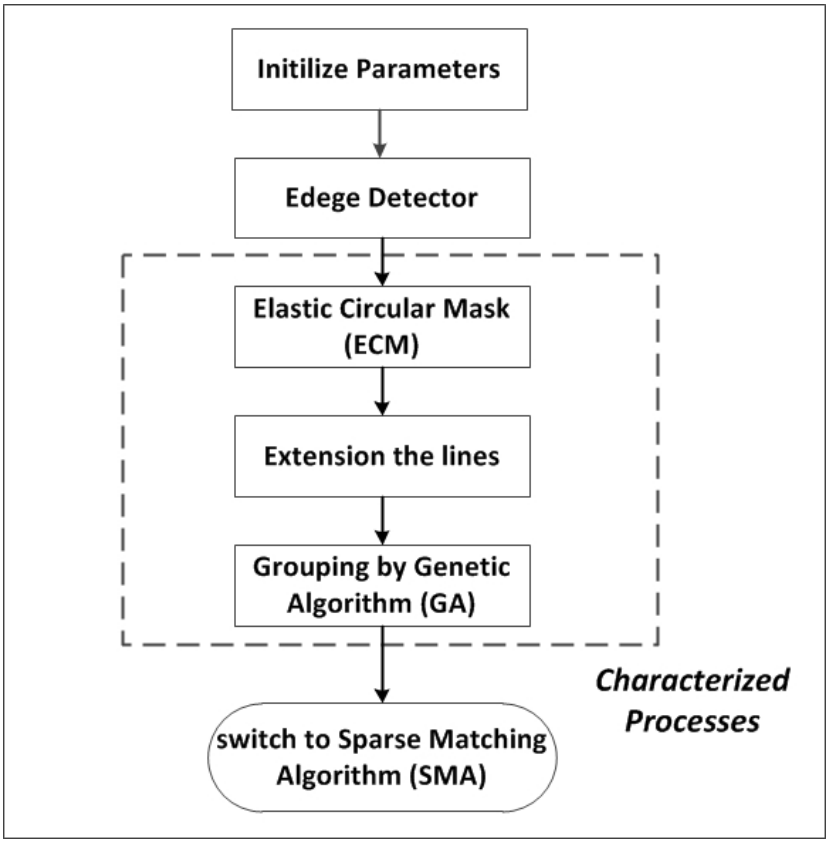


Fig. 3-1 The workflow of road detection from image

Assume the road network in GIS database is a graph model $R=\{V,E\}$, where $V \in \{v_1, v_2 \cdots v_i\}$ and $E \in \{e_1, e_2 \cdots e_j\}$ represent the road crossings and road segments respectively. The conflation needs the reliable features from the target image $I_{m \times n}$: either point features $P \in \{p_1, p_2 \cdots p_m\}$ such as intersections or curvilinear features $S \in \{s_1, s_2 \cdots s_n\}$ such as road segments.

To avoid the confusion, the end points of a road segment in the road network are referred to as nodes, whereas the end points of the extracted line elements from the image are defined as vertices. The intermediate points of a road segment or extracted line element are termed as shape points.

Roads in VHR image are reflected by their geometric and radiometric characteristics which have stimulated the development of various detection algorithms. However, the complete road objects can rarely be detected due to the existing noise. In most cases, only separate road fragments can be extracted. Moreover, linear structures from other object types such as building outlines have similar geometric and radiometric characteristics and can therefore be wrongly detected as road candidates. Fig 3-2 shows an example of the edges detected by Canny operator: the road has rarely the parallel boundary and the objects with parallel boundary are not always roads. Moreover, the zebra crossing and vehicles on the road are also detected as linear features, and the edges are frequently broken. Additional grouping algorithms are needed to filter out the noisy candidates. In short, the road segments that represent only a small minority among the edges are difficult to be distinguished from other features.

Fig. 3-2 Edges in red in the optical image

The geometric character of the edges can be employed as a useful indicator for objects with regular shapes. Features in VHR image with straight or curve edges on one or both sides, for instance, are possible road candidates. The road surface presents a homogenous radiometric intensity. These characteristics can be used in our ECM detector.

Edges in an image are located where there is a large radiometric difference to both sides and they contain a lot of tiny segments which may be missed during the grouping procedure. The width threshold for the road detection can be empirically preset as 1/3 of the lower bound of road width. Along the normal directions of the edges, a searching algorithm can most probably detect road features. For this purpose, a potential edge matrix is created for the filtered edges with intensity values used to find the correspondences as shown in Fig. 3-3. The intensity value of zero means non-edge pixel. The searching process starts from the normal direction of each potential edge, where the normal direction is computed from the pixel's neighbors on left and right, and it continues until the maximal road width i.e. width threshold is reached or it meets other edge pixels. The former case means that there is no valid road edge as the candidate, whereas the latter means a connection between two pixels can be established and the middle point of this connection is the potential counterpoint of the road. The searching process is repeated until all edges from the matrix have been traversed, thus, all the possible road counterpoints have been identified.

The searching procedure requires the maximal and minimal road width as control parameters. Fig. 3-3 demonstrates the searching procedure as well as its propagation to other pixel candidates of an extracted edge. The pixels of edge candidate (Fig. 3-3 left) have been labelled with its ID (Fig.3-3 right). The longest edge in the region is the searching start, and except the ends of the edge, each pixel is the point to probe new connection to the neighboring edge using the normal direction. We selected randomly three probing points with edge ID '1' to show the relationship to other edges.

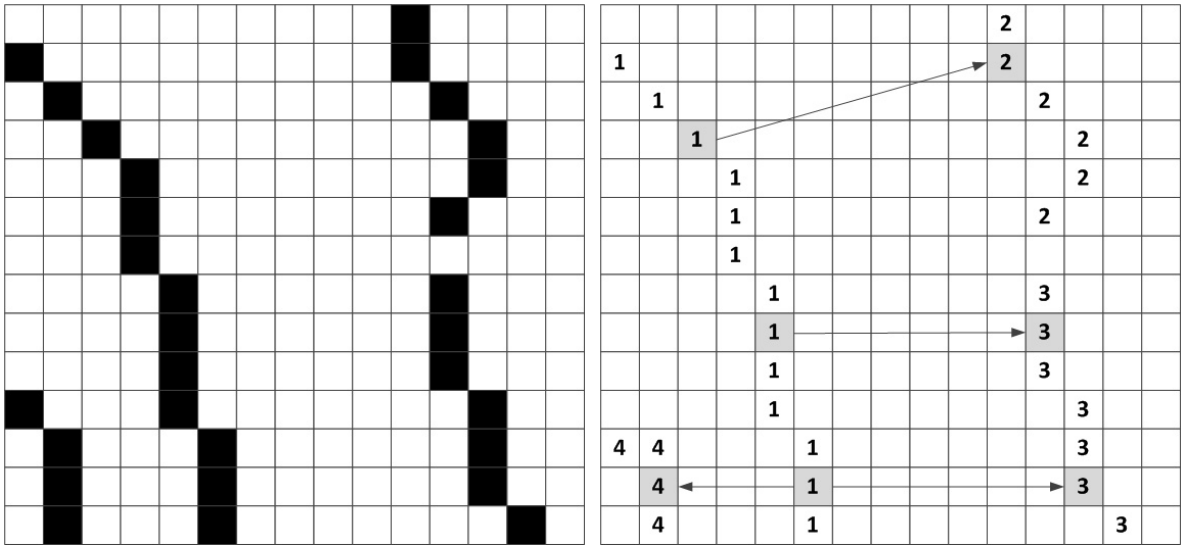


Fig. 3-3 Edge pixels (left) and potential edge matrix with normal directions (right)

The individual centre points between two opposing edge pixels infer the possible road area. Around each centre point, a circular region can be constructed whose diameter indicates the road width. However, the noise in image may cause some irregular changes of the diameter along different edge pixels. It is one of the reasons for us to adopt an ECM detector with a radius between a minimal ($W_{\min}$) and a maximal ($W_{\max}$) value. Fig. 3-4 illustrates the idea of ECM which can also handle gaps on edges.

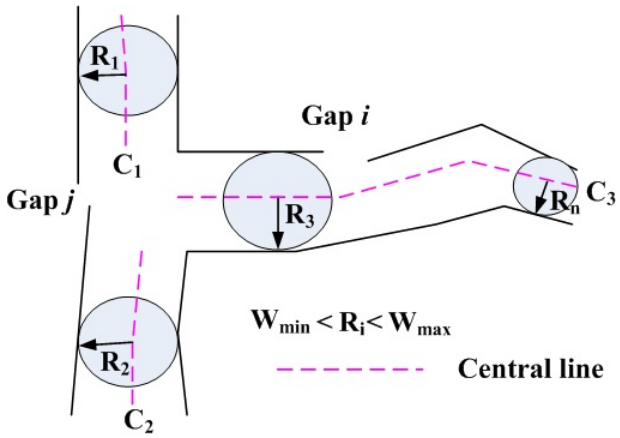


Fig. 3-4 Elastic Circular Mask (ECM) detector

Being guided by the geometric properties, we may determine road candidates among the extracted low-level edges. By connecting the neighboring circular masks, longer linear regions can be generated. We term this as a grouping process during which the edge ID and the curvature are used as two control parameters. The road is divided into different parts if the curvature is larger than a preset threshold; otherwise, the circular masks with identical edge ID belong to the same road segment. However, the centrelines through the connected centres of masks are still noisy and contain some non-road features in the image. On the other hand, the ECM detector may also miss some road candidates because it only chooses the normal direction of the edges as the possible searching direction. Some areas such as those covered by shadows are not included in any circular mask. If the road segments are extended on both sides, the coverage ratio of the road surface can be enlarged and some noise on the road eliminated. Although the spatial extension is not always stable and could introduce wrong candidates, it may lead to the generation of the candidates of possible intersection areas. As shown in Fig. 3-5, the initial detected road lines and the extended parts are visualized with a colour bar which indicates the ratio of average intensity in each centre segment and also for both sides, where the intensity baseline is 255. The grouping procedure, which will be discussed in Section 3.2.2, is necessary.

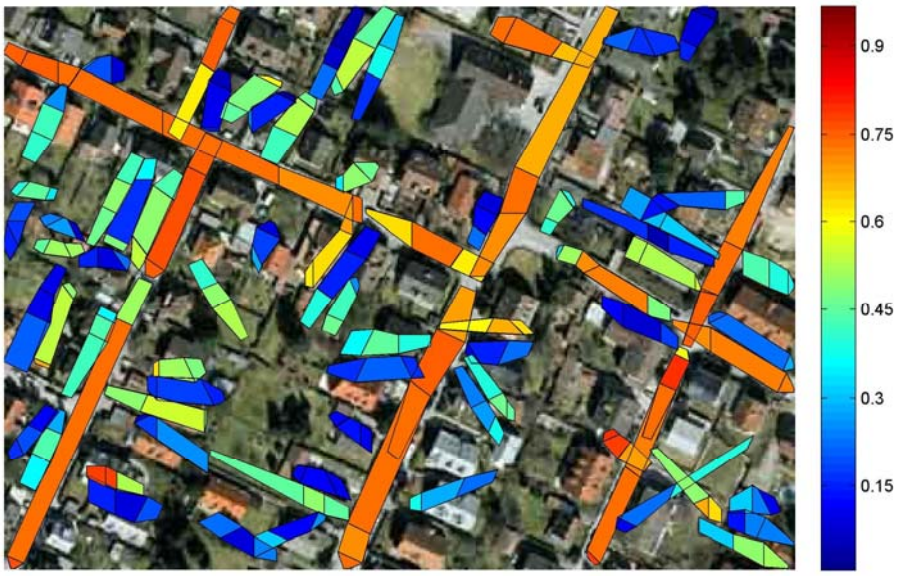


Fig. 3-5 Extended road areas from the road candidates

In spite of the noisy segments in current results, it is already one important step further from the edge primitives towards the road objects. We summarize the entire process as follows.

Algorithm 3-1 *Extraction of Linear Region*

Input parameters:

1) Edges detected by the edge operator;
2) Minimal and maximal road width $[L_{\min}, L_{\max}]$;
3) Extension threshold $[E_{\max}]$;

Output:

1) Road segments with extended parts S;
2) Connection matrix for road segments I;

Steps:

1) Search along the normal direction of all edge pixels in the potential edge matrix and get the possible circular masks;
2) Grouping the circular masks to create the centrelines of road segments;
3) Extend the segments to their maximal length by the threshold $E_{\max}$;
4) Record the intersecting segments and discriminate between the intersection and common connection.

In the above process, the curvature in the road segments can be ignored. But once the segments are connected, the curvature should be considered. Otherwise errors may be introduced in the feature matching procedure. The connection of the centrelines of roads from ECM detector to reach the maximal length needs the estimation of the curvature which can be described in Rueda et al. (2008) as:

$$K(p) = \frac{8S_p}{4S_p^2 + C_h^{\ 2}} \tag{3-1}$$

Where K is the curvature; p is point position; C_h represents the so-called *c-scale* element (Rueda et al., 2008); S is the distance from the shape point p to the *c-scale* element.

The curvature definition was used also in the road decomposition algorithm in Chapter 4.

After the curvature estimation, the detection algorithm chooses approximately the two end points to reinitialize the direction of the extended line if the maximal curvature is less than certain threshold. The curvilinear regions found by ECM detector usually are only parts of a road because of the existing noise. Comparing to the vector road database, the centre lines detected from satellite image are still very fragmental and many times in number as road objects in the vector database; moreover, the performance of the matching algorithm tends to deteriorate on datasets with outliers. The number of centrelines can be reduced if more information such as the radiometric characteristics and the perceptual factors that determine the course of roads are utilized to chain the centrelines into longer segments. The post-processing procedure aims to reduce the number of road centrelines from image to nearly the same number to road objects in the vector dataset which has more accurate topology, thus facilitate the subsequent matching efficiency. We choose the Genetic Algorithm (GA) which has many applications in image processing to integrate various characteristics of road features (Jeon et al., 2002; Jia et al., 2008). The GA has two functions: 1) grouping the

fragmental road centrelines; and 2) remove noisy segments by means of the radiometric property. By chaining and extending road centrelines, the spatial relationships among the extracted centrelines can be constructed and serve as essential clues for the matching algorithm. In the reality, it is difficult to ensure the same number of road segments in both datasets, so we propose later the so-called Sparse Matching Algorithm (SMA) to estimate the correspondence for the imbalanced datasets using the results from the GA grouping method.

3.2.2 Chaining of fragmental centrelines

Among the extracted linear centrelines, some non-road elements, e.g. buildings, trees, are involved and need to be eliminated, and the remaining road candidates will then be connected as much as possible into road networks. GA is a heuristic searching approach to generate optimal solution for noisy dataset. An introduction about GA can be found in Goldberg (1989). Jeon et al. (2002) reported the extraction of roads from Synthetic Aperture Radar (SAR) image using GA based on a region-growing strategy. Two modifications are conducted in our context to simplify the difficulty: 1) the chromosome initialization of the starting road features; 2) an improved fitness function with more strict perceptual parameters.

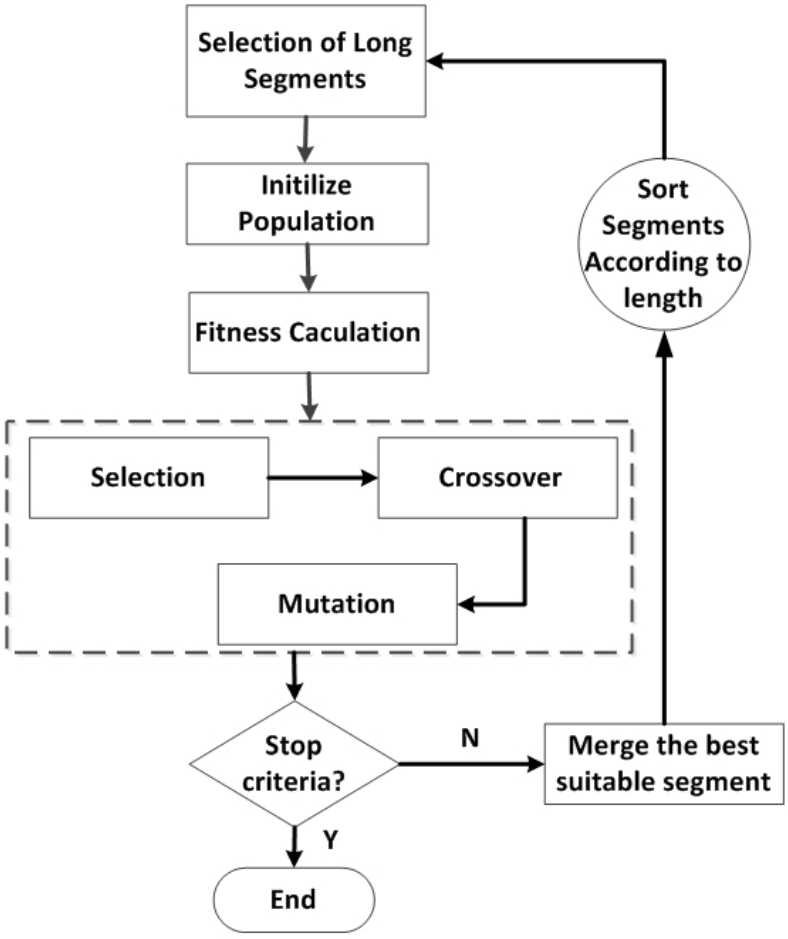


Fig. 3-6 The dataflow for filtering and chaining the line elements

The workflow of GA in Fig. 3-6 contains three main steps: the Chromosome initialization, fitness function and evolutionary rules. In the evolution process, the computation repeats until the preset conditions are fulfilled, so the result of GA is the optimal outcome from the searching space in terms of fitness, although it does not traverse all the combinational

possibility of the searching space, because the enumeration of all the combination from one candidate to the rest in the group is impossible in a large scale dataset. GA only considers the limited number of all possible combinations in the evolutionary process.

(1) Definition of chromosome

In GA we set the length of the chromosome the same as the number of centreline candidates and each bit of the chromosome represents the individual centrelines with a binary value 0 or 1. Since the features from the vector database provide the prior knowledge, their corresponding linear centrelines from the image are usually located inside the buffer (we choose 200m as the width for buffer), furthermore, the longest lines are selected as seeds of the chromosome. Other lines may or may not be connected to these seeds. The probability *1* indicates the capability of connection to the seeds, whereas *0* means the impossible connection.

(2) Fitness function

Both geometric and radiometric characteristics are considered in the proposed fitness function. Here the proximity function *P* indicates the closeness between two centreline candidates in Eq. 3-2. The $l_{\min}$ is the minimum distance of the two neighbors, *D* is the scale-independent density of the base feature, and *R* is the minimum distance between two neighbors.

$$P = \frac{l_{\min}^2}{2D\pi R^2} \tag{3-2}$$

The second parameter is the length of the candidate. The normalized length *L* is expressed as the probability function:

$$L_i = \frac{l_i}{\max_j\{l_j\}} \tag{3-3}$$

The third parameter is the colinearity *C* of the candidates, which has been recognized by Jia et al., (2008) as the key factor for the grouping procedure, *T* is the threshold for acceptable angle θ of the difference:

$$C = \begin{cases} 1-\dfrac{\theta}{T} \; with \; \theta \in [0,T] \\ 0 \quad with \quad \theta > T \end{cases} \tag{3-4}$$

The fourth parameter is the spatial relationship *O* between the candidates. It means the overlap ratio if they project to each other, O_{ij} is the length of the overlap and $A(l_i, l_j)$ is the sum of the length of both segments:

$$O = \begin{cases} 1-\dfrac{o_{ij}}{A(l_i,l_j)} & o_i \neq 0 \\ 0 \qquad else \end{cases} \tag{3-5}$$

Fig. 3-7 shows the closeness, colinearity and overlap for a specific road object P_1P_n, the

proximity between L_3 and L_5, and the noisy features L_2 and L_4.

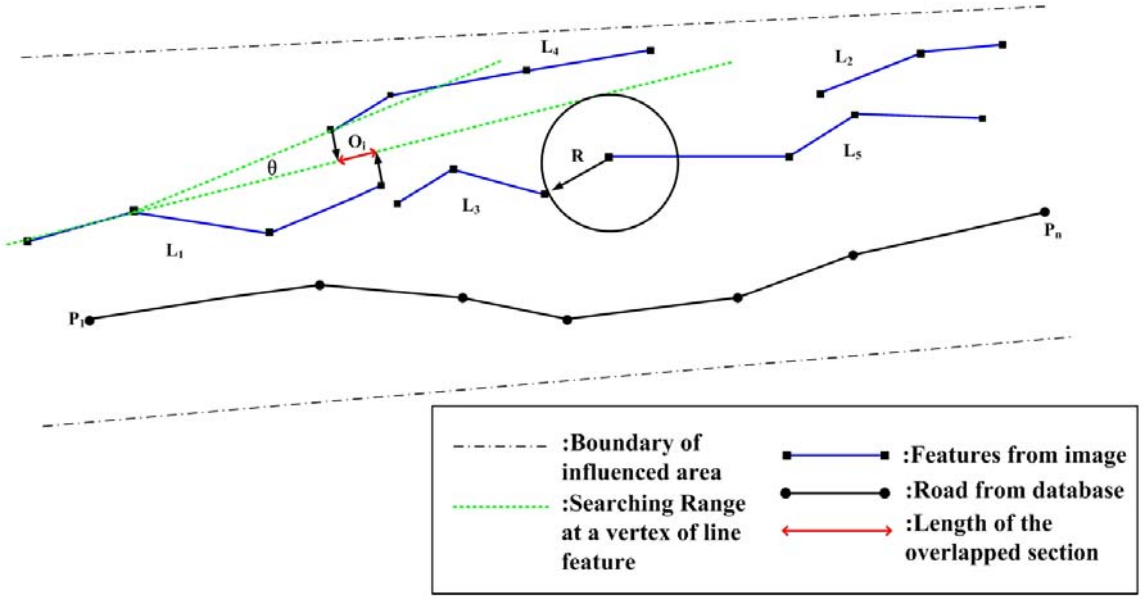


Fig. 3-7 Geometric parameters of the fitness function

The last factor is the intensity S of the candidates. It is the average radiometric value I of linear region or the road surface in the image with $M \times W$ pixels:

$$S = \sum_{i=1}^{N} \frac{h_i}{N} \quad \text{where} \quad h_i = \frac{\sum_{m=1}^{M}\sum_{w=1}^{W} I_{mw}}{MW} \tag{3-6}$$

The fitness function is defined as a linear combination of these five factors. If *N* is the number of chromosomes with *1*s, the fitness will be

$$F = \sum_{i=1}^{N} f_i / N \tag{3-7}$$

$$\text{with} \quad f_i = \max_j(\alpha P + \beta l + \kappa C + \gamma O + \eta S)$$

where $\alpha, \beta, \kappa, \gamma, \eta$ are the weights for different probability functions and their values are predefined according to the empirical values of the road characteristics. We choose $\alpha = 0.3, \beta = 0.5, \kappa = 1.0, \gamma = 0.5, \eta = 1.0$ respectively in the experiment to make sure the orientation has the first opportunity.

(3) Evolutionary process

The evolutional process of GA is realized by selection, crossover and mutation. In each generation, it simply selects half of the population with higher fitness values, and the selected individuals can generate the children by separation and connection of their chromosomes, so the new individuals and the selected individuals are the ingredients of new generation, where the new individual keeps the genetic diversity by mutation with predefined probability in the test.

In the experiment, we set three termination conditions for the evolution: 1) when it reaches the maximal number of generations i.e. *100* iterations; 2) when the maximal fitness of the population remains constant in last *15* generation; 3) when the individual with maximal fitness has already the same length of the prior road segment. In this way, the GA produces the optimized solution with values equal to *1* in the features.

3.2.3 Efficiency evaluation

In order to evaluate the efficiency of the extraction and grouping results, we test the proposed approach with a small dataset here, and a comprehensive verification will be presented and discussed in Chapter 5. In this simple experiment, there are only several road segments and intersections in the image (Fig. 3-8, a), and the noise is also kept limited.

a) Detected road centrelines

b) Grouped road centrelines

Fig. 3-8 Linear road segments before (a) and after (b) the grouping process

The road centrelines are extracted using *Algorithm 3-1* with the initial parameters, and there are 48 extracted candidate segments in total. In the grouping process using GA, 100 iterations as the maximal generations are chosen as the termination criteria, other criteria are the same as described in Section 3.2.2. Moreover, the length threshold of 30 meter is set to filter out the short line. There are only 21 road segments left (43.8%) after the grouping process. Fig. 3-8 shows the results before and after the grouping, and the colour indicates the ratio with respect to the longest lines in the area.

Premature is a common problem of the GA algorithm, but it performed quite well in this case, and it became stable after 50 iterations, which indicates that the fitness functions are suitable for the chaining neighbor lines.

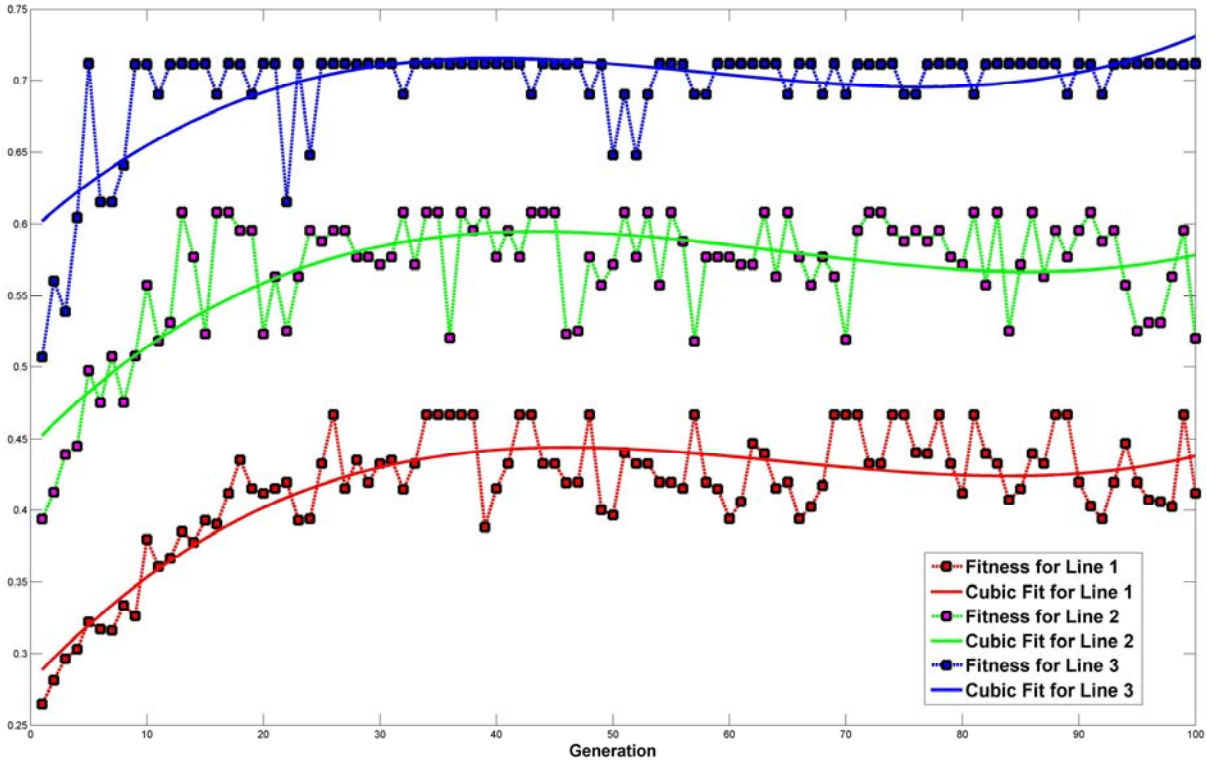


Fig. 3-9 Efficiency of GA for chaining line segments

Fig. 3-9 illustrates the fitness values of each generation for the top 3 lines with longest lengths, and it is clear that the value is almost stable at the end of iteration, which indicates the searching result is the optimal in the traversed searching space, and the further searching is not necessary.

In the experiment, 56.2% of the extracted candidate centrelines are filtered out, but the remaining centrelines are still quite fragmental for a common matching algorithm. Therefore, a special matching approach is needed to deal with the outliers.

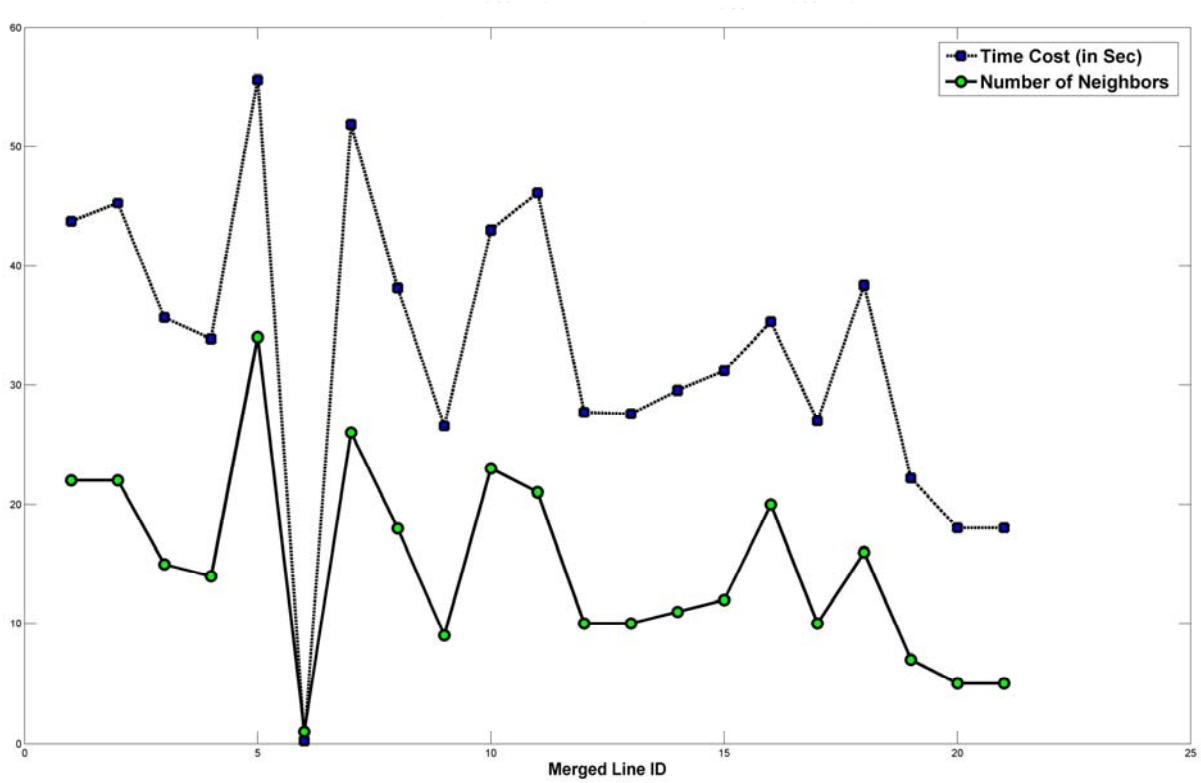


Fig. 3-10 Time cost and the number of neighboring lines in grouping

In the experiment, the time cost depends on the number of the related lines located on both sides of the current segment. If there are no related lines, the current segment remains and it requires a minimal time (0.2 second). If there are related lines, more time (34.7 second in average) is consumed in the searching process, and it is proportional to the number of neighbors.

3.3 Road Extraction from Land Parcels

3.3.1 Overview

The extraction of road features from image cannot yet satisfy the requirements for the matching algorithm because of the too low extraction ratio by the up-to-date detection algorithm (Rottensteiner, 2009; Poullis and You, 2010), although certain improvements can be reached, for instance, with help of the GA grouping. Other available spatial datasets can therefore be adopted as optional choices to derive road segments. We generate the road network from a database of land parcels, which is an important spatial data for the derivation and enrichment of road data, because the cadastral dataset usually has higher currency (Zhang et al., 2010). The quality of the extracted road from land parcels proves better than that of road segments from image, but lower than the road network in a navigation database. Fig. 3-11 demonstrates three land parcels (polygons) and three road centrelines (dot line). The space between the parcels is the road area. The centrelines can be computed using visual perception criteria. In this section, we use the modified straight skeleton operator to automatically derivate the road centreline. The skeletons are linear representations reduced from generally elongated patterns.

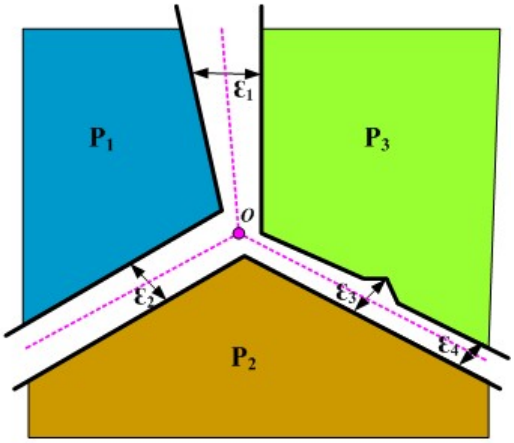


Fig. 3-11 Land parcels and road centrelines

A ‘good’ skeleton should satisfy the following criteria (Klette, 2003; Hu et al., 2007):

1) Being cantered inside the original shapes;
2) Conforming to human perceptions of the original shapes;
3) Being efficiently computable;
4) Being robust against noise and geometric transformation;
5) Having low complexity, but with high automation.

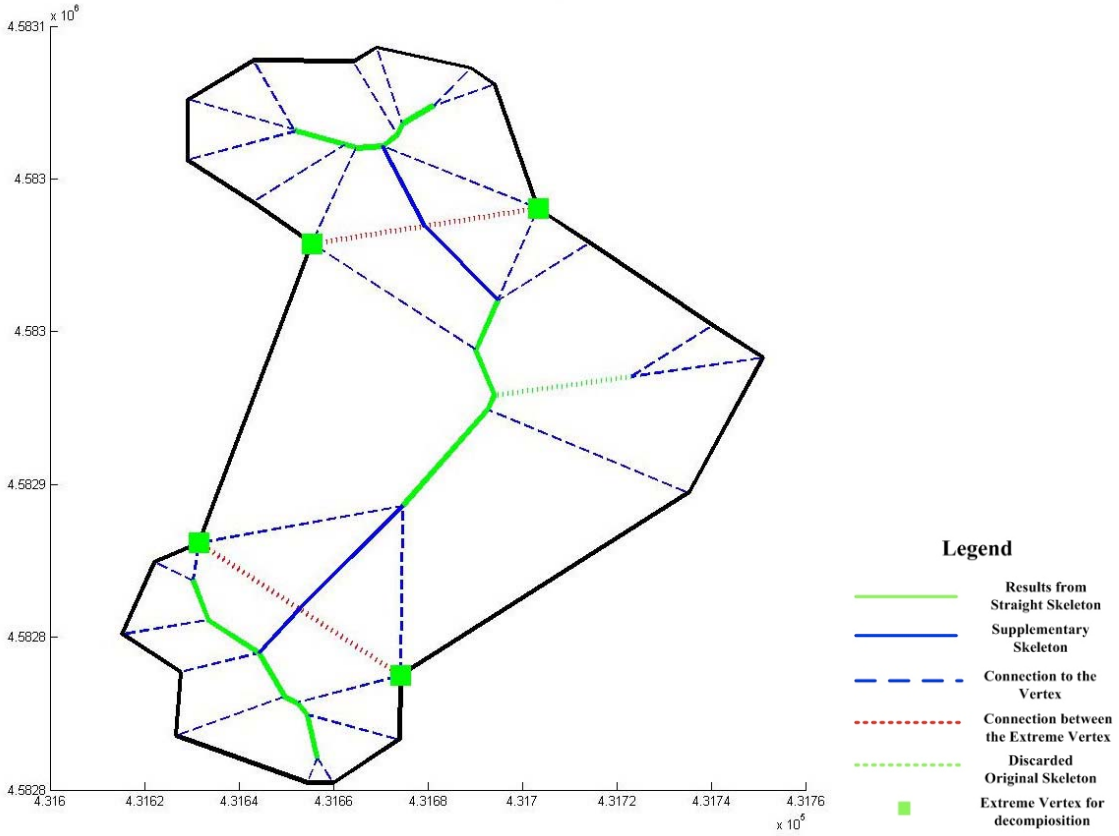


Fig. 3-12 Shape decomposition and skeletonization of an irregular shape

The existing skeleton algorithms can hardly satisfy all these criteria. Most of them suffer from some incomprehensible artifacts which need extra information or other constraints to get pruned, and the noisy boundary is an important cause of the artifacts. Any subtle noise or

fluctuation of boundary may result in redundant skeleton branches that may seriously disturb the topologization of the skeleton (Adluru, and Latecki, 2009). Shape decomposition can divide the noise into different regions, so the pruning step can get more prior information from the shared edge. These measures may help to create results that suit human perception better. Fig. 3-12 shows an irregular shape and its skeleton.

We compared three different skeleton operators i.e. Medial Axes (MA), Straight Skeleton (SS) and Triangulation Based Skeleton (TBS) (Zhang et al., 2010). They are fairly stable in small scale computation, but the result from Straight Skeleton algorithm is best suitable for the road network, where the smooth curves are approximated by short straight lines.

3.3.2 Shape decomposition

Shape decomposition is an efficient solution for large-scale computation, especially in GIS with very large number of features. Regardless of the difference between the man-made features (such as road, building) and natural features (such as river, forest), the topology is one of their common characteristics. The skeleton can be utilized to simplify the shape of the feature or to generate the centreline of the road (Haunert and Sester, 2008). Similar to the results from Cordella et al. (2000) and Siddiqi & Kimia (1995), (Cohen and Singh, 2007) demonstrated in a recent study on human perception that negative minima of contour curvature affect the observers' identification of contour segments in highest degree. Additionally, segment identification is determined by the contour length, the turning angle at part boundaries and the width at the part's base. For Straight Skeleton approach, the reflex and sharp angles on the boundary can influence the performance significantly, and therefore they should be regarded as noise and reduced during the shape decomposition.

Let $V=\{v_1, v_2, \cdots v_n\}$ be a set of vertices of a polygon P, and $E=\{e_1, e_2, \cdots e_n\}$ a set of edges with e_i being the length of the edge from vertex v_i to v_{i+1}, but e_n is from vertex v_n to v_1. Usually, the noisy edge can be described as the length ratio between two neighboring edges through a specific vertex, $R=\{r_1, r_2, \cdots, r_n\}$, $r_i=\frac{e_{i-1}}{e_i}$, and only the short edge with noise ratio around *1* consider as the noise vertex candidate. The protrusive candidate vertex list $L=\{l_1, l_2, l_3, \cdots, l_n\}$ for the shape decomposition is computed from the negative minima of contour curvature and the reflexive vertex.

Two categories of vertices are selected as candidates for decomposition: 1) Vertices with the negative minima of contour curvature; 2) Reflex vertices. The local curvature extreme is derived from the differences in average orientation between two neighboring vertices (Siddiqi & Kimia, 1995), whereas the smooth edge can be evaluated by the second-order shocks (Kimia et al., 1995). Siddiqi's method is modified in our case because we use straight line shape instead of smooth curve. Fig.3-13 shows the curvature and local extreme of each vertex in Fig.3-12.

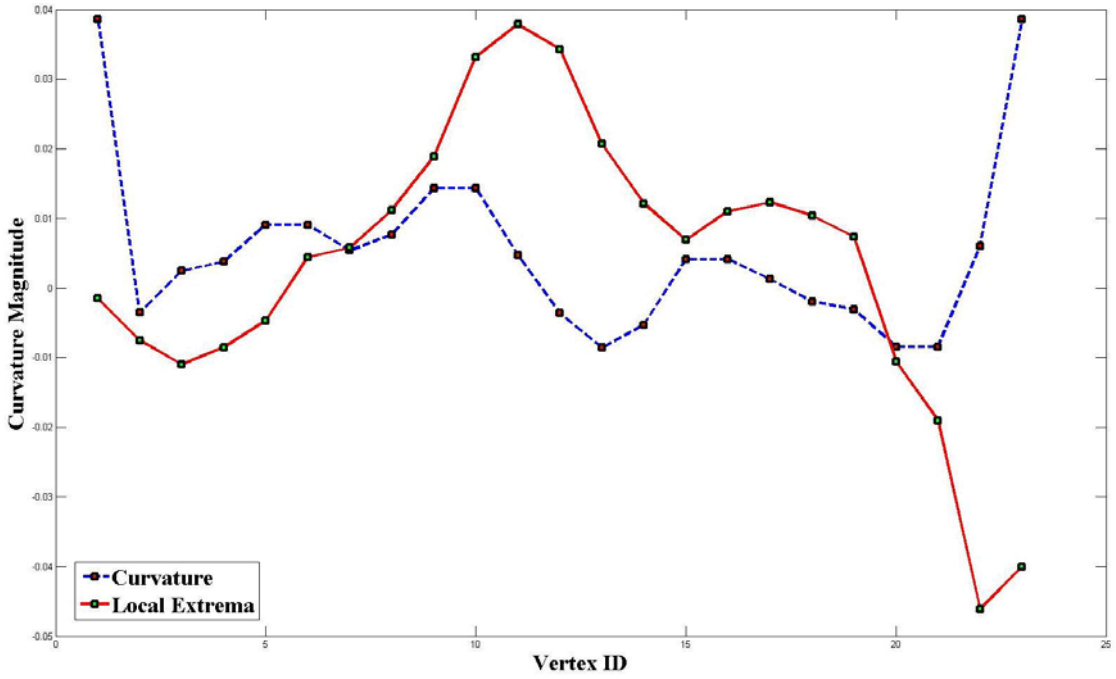


Fig. 3-13 Curvature and local extreme of vertices on the polygon in Fig. 3-12

The small noise on the edge can have great influence to the skeleton results, and it should also be excluded in the candidate vertex list. The noisy edge has two characters: 1) shorter than the pre-set threshold, and 2) the neighboring edge is also short. So we define the noise as:

Definition 3-1 *Edge* e_i *on polygon* P *is noisy if its length is shorter than a specific threshold and the ratio* r *is near to 1.*

In this thesis, the reflex vertex v with the permitted ratio $r = 10\%$ has been used as the threshold, and Fig. 3-14 shows the relationship among the curvature, ratio and the edge length.

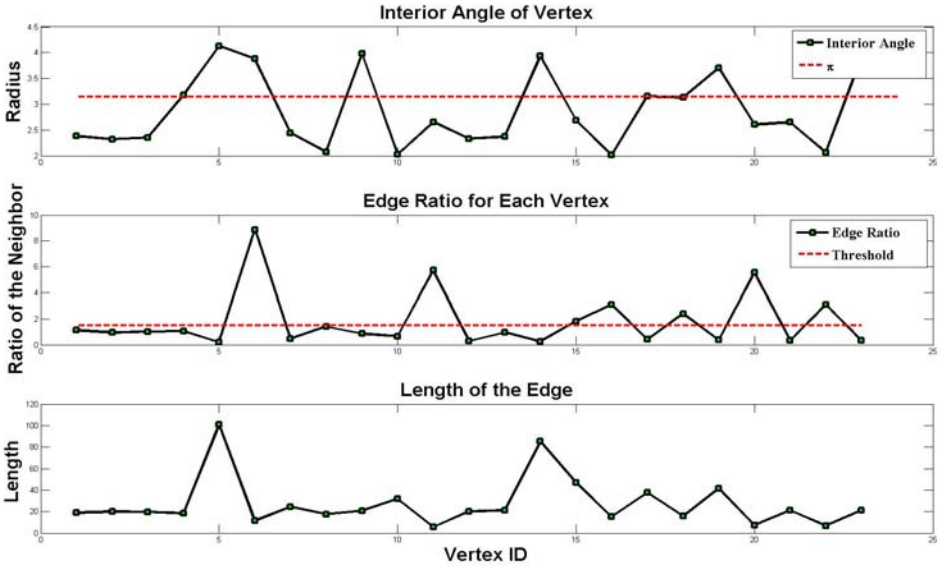


Fig. 3-14 Interior angles, ratios and lengths of the edges

Intersections or intersection regions (Zou and Yan, 2001; Tang and You, 2003; Haunert and Monika, 2008) play a significant role in the pruning procedure, and we choose the typical intersections as the prior knowledge, which includes UV pattern, cross pattern, T pattern, K pattern, X pattern, V pattern, Y pattern and two randomly selected intersections from the test area. Some of the skeleton results are demonstrated in Fig. 3-15. The selected intersections are included in our test with geographic coordinate system.

In the results, the Medial Axes produces smooth change for reflex angle, and it has the same result as Straight Skeleton if the shape is convex. The Straight Skeleton provides the best approximation to Medial Axes in case (2-1) and (2-2). The TBS creates strange skeleton because the 0-edge triangle inside the shape has great influence on the main skeleton (1-3, 2-3) for highly symmetric shape. For the shape with no prior knowledge, the triangulation-based skeleton (3-3) can provide more details in the skeleton, depending on the sampling points on the edges.

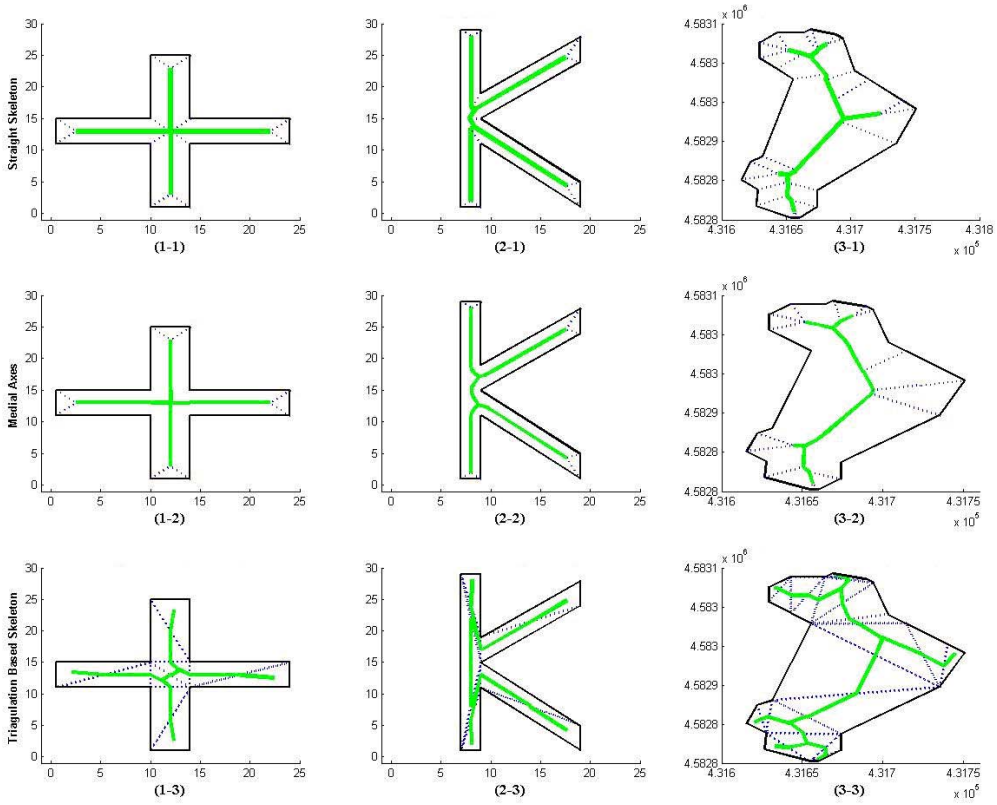


Fig. 3-15 Skeleton results created by three different operators for typical intersections

The stability of these operators for small test dataset is fairly good, but their computing efficiency varies largely from one other. In general, the Medial Axes has low efficiency for the construction of smooth arcs. The Straight Skeleton and the TBS have similar performance, but the computing time of Straight Skeleton is related with the number of reflex vertices in the original shape. The relative computing performances of the operators for the typical intersections are summarized in Table 3-1.

Table 3-1 Computing performance (in millisecond) for typical intersections

ID	Intersection pattern	Number of vertices	Number of reflex angles	SS (ms)	TBS (ms)	MA (ms)
1	V	6	4	1.3	0.7	95.1
2	UV	7	1	1.9	0.9	101.0
3	T	8	2	2.3	0.9	1188.6
4	Y	9	3	1.8	0.9	164.7
5	K	11	4	2.7	1.0	280.9
6	Cross	12	4	5.4	8.2	1028.3
7	X	12	4	2.4	1.1	499.6
8	Random1	23	6	4.7	1.9	1027.4
9	Random2	47	21	12.1	3.4	4941.9

Algorithm 3-2 *Shape Decomposition Algorithm (SDA)***:**

The entire vertex set V of the polygon and the protrusive candidate vertex set L are the input. The vertices in L are traversed one by one. For each vertex being treated, its distances to other vertices in L are computed and sorted in ascending order. Two rules should be kept to compute the new edge which divides the shape: 1) the new edge should not be one of the edges in original shape, and 2) the new edge should have no intersection with any edges. The edge is constructed if it fulfils the requirements and the vertices delimiting the edge are then marked as used. This procedure goes on with the next unmarked vertex in L until no further vertex remains.

Fig. 3-16 shows the results from the above algorithm: the green square markers are the protrusive vertices, whereas the dash lines in brown are the new constructed edges. The constructed edge is a connection between two vertices which are not neighbors in the original shape. Each sub-region after decomposition can be then skeletonized without or with less artifacts.

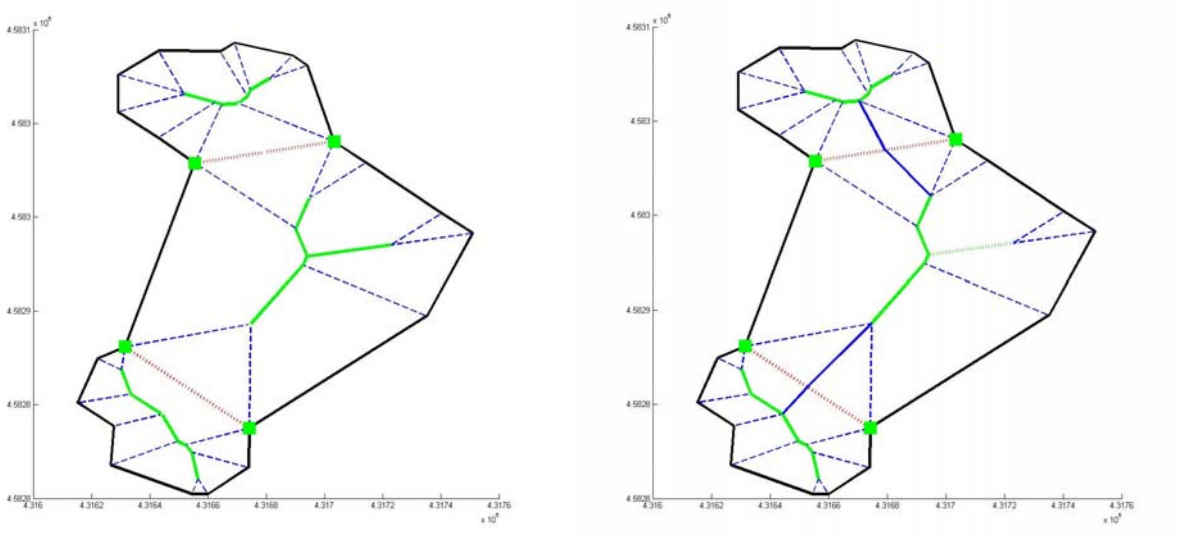


Fig. 3-16 Shape decomposition (left) and skeleton (right)

The decomposition divides the shape into several sub areas, and then we define the new shape by the generated new edge. Three types of the new shape are defined by their connections:

1) Terminal region which contains only one new constructed edges, others are the edges from original shape;
2) Connection region which has two new constructed edges, and it can be deemed as passage among the neighbors;
3) Intersection region: which has more than two new constructed edges, and it is the difficult part for skeleton algorithm, because it can be a very complex intersection in the road network. Usually it is located around the road intersections, and there are many examples in Fig. 3-17.

3.3.3 Extraction of road segments using straight skeleton algorithm

The straight skeleton is a mathematical structure that is defined by tracing the corners shrinking at each vertex of the shape. Edge event and split event need to be treated properly during the shrinking procedure. In the section, we firstly modify and implement the straight skeleton algorithm (Felkel and Obdrzalek, 1998), and then prune the skeleton for each decomposed polygon.

Similar to other skeletonization methods, the straight skeleton algorithm may produce some artifacts such as spurious skeleton branches. Haunert and Sester (2008) presented their method to prune the skeleton by the defined road axes, which is effective when the shape contains long straight edges. We modify their method for the decomposed regions. In spite of the different definition of the road for cadastral administration and the navigation applications, we define the road centreline as the skeleton of the parcel in order to get the centreline network instead of the isolated skeleton.

Definition 3-2 The *road centrelines from the parcels should satisfy the requirements:*

1) Road centreline is the approximated line, which is located between two or more parcels;
2) Each parcel is enclosed by the generated centrelines, except the border of the test area;
3) The road centreline of a present parcel is connected with those from the nearest neighboring parcels in all directions. The connect centrelines reflect the intersection region after decomposition;
4) The centreline of a parcel keeps the direction when it is connected to the centreline of the neighboring parcels.

The long straight edges characterize the road network, but skeletonization will become complex in urban area, especially around intersections. After the decomposition, spatial relationships are set up between the neighbor polygons via the shared edges. The shared edges can be also used as the prior information to reconstruct the connection of the decomposed polygon. The topology reconstruction algorithm based on Definition 3-2 can be summarized as follows:

Algorithm 3-3 *Topology Reconstruction Algorithm (TRA) based on shared edges for skeleton.*

The inputs of this algorithm are edge set E, decomposed shapes P and the set of shared edge $E^{'}$. The Straight Skeleton K of the shape p_i is firstly computed, then the standard

deviations between ratios of edges in E and E' are calculated in order to estimate the main direction of the shape. In each polygon in P, three constrained rules are used to control the reconstruction process: 1) mean slope difference between E and computed skeleton K, 2) the standard deviation of the skeleton, 3) the shapes which do not fulfil the above rules. The candidate segments connecting to the shared edge are selected from the skeleton by computing the shortest distance between the vertices of the skeleton and shared edge E'. This selection process is controlled by the above rules. However, there are some exceptions where the shortest connection intersects with the skeleton or shape. In this case, the original skeleton is kept. The skeleton segments on both sides of the shared edge are smoothed by minimizing the difference of their orientation, and the output is the reconstructed skeleton K'.

This algorithm assumes that the skeleton is reasonable if the mean slope difference between polygon edges and skeleton is less than the threshold (30 degree in the experiment), but the input skeleton contains some artifacts and may influence the mean ratio. Therefore, the mean ratio is chosen only if the standard deviation σ of the polygon is small (2σ in the experiment). It firstly uses the shared edge to approximate the skeleton, but it keeps the original skeleton if it fails to use the shared edge.

In facts, the reconstruction results from the straight skeleton are rather reasonable except the artifacts, and we may detach the artifacts in order to keep the skeleton. All the lines in skeleton which have connection to the shared lines are kept; otherwise the lines will be deleted.

Algorithm 3-4 *Skeleton Pruning Algorithm (SPA): The pruning algorithm for detaching the artifacts in the computed skeleton.*

Inputs of the algorithm are the computed skeleton K of the decomposed polygon and the shared edge E'. The valence for each vertex in K is calculated after the topological reconstruction from algorithm *TRA*, and then added to the vertex set V. The vertex is marked as shared vertex if it is on the shared edge E'. After finding the edge set of skeleton K and relating it with V, the vertex from V is selected if it has more than three connections. Each edge of the selected vertex is evaluated and kept if it connects to the shared edge or has inside loops, otherwise it is marked as artefact and pruned.

The vertex with only one connection is the possible start of the artifact or connection to neighboring polygon. It is always true if the shape can be perfectly decomposed, but the dead lock roads are exceptions. Fig. 3-16 shows the detached branches in dash line. Terminal region and connection region can be simply recognized by their location, but the intersection region is more complex. As noted in (Zou and Yan, 2001; Haunert and Sester, 2008), the intersection region is the most complex part due to the flexible shape of the intersection itself. Complex intersections are also very common in real road networks. The non-processed skeleton created by TRA and SPA algorithms can be kept, although they contain some unreasonable results.

3.3.4 Reconstruction of road network

Here we demonstrate the reconstruction results of road network for the selected test area which is part of the city Barcelona and only the parcels are provided. There are 743 parcels covering an area of 12 km^2. Fig 3-17 shows the test region, and it includes part of the new and old town which reveal different patterns. The grey areas are the parcels, whereas the

decomposition results, i.e. the areas among the land parcels, are classified by the number of the constructed edges. In total there are 1434 decomposed polygons.

Using SDA, the entire road area is decomposed into 91 (6%) terminal regions (Conn 1), 774 (54%) connection regions (Conn 2), 556 (39%) intersection regions (Conn 3 or more), and 13 (1%) isolated regions (Conn 0), which has no connection to the network, and usually the isolated region is at the border areas. The connection region (Conn 2) and the intersection regions (Conn 3 or more) dominate in the test area.

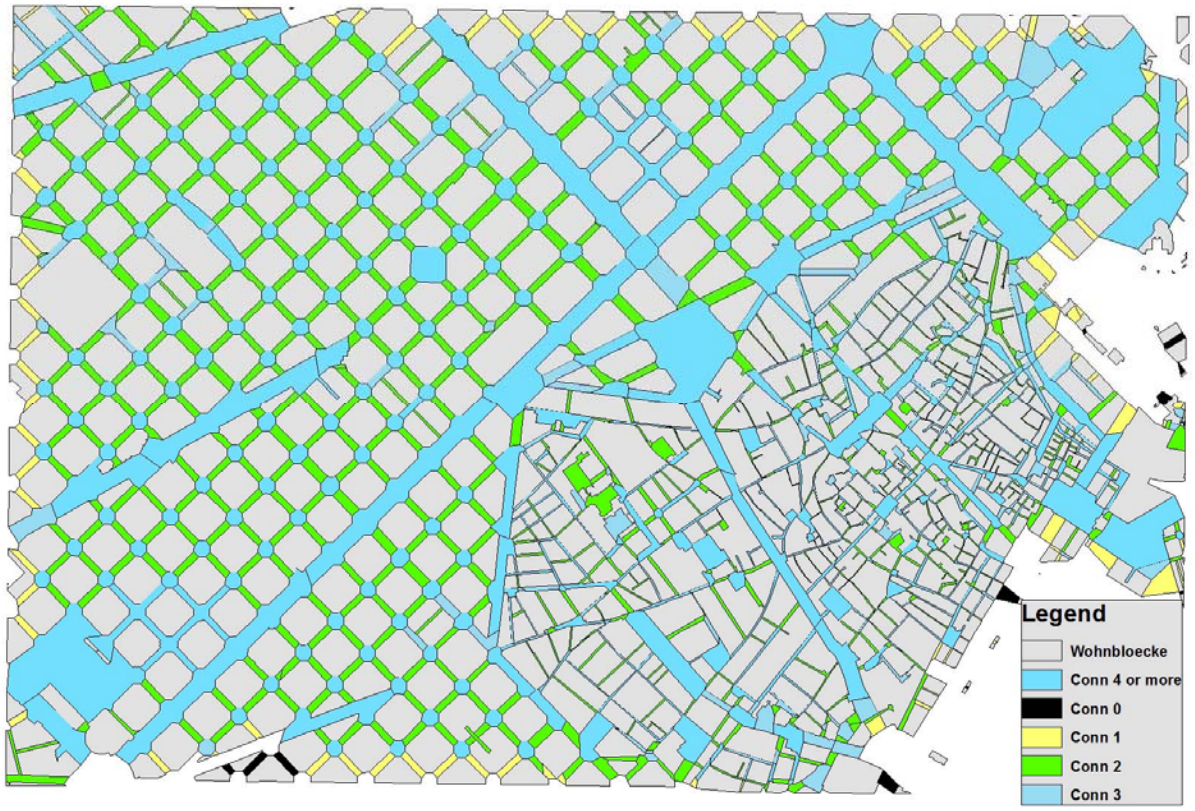


Fig. 3-17 Results of shape decomposition

Among these regions, the intersections are the most difficult part to be skeletonized. 24% complex intersection regions cannot be classified as the typical intersections (ID 1~7) defined in Table 3-1. Therefore, only 76% intersections from the test area are treated.

The decomposition procedure takes the typical intersection patterns described in Table 3-1 as prior knowledge, so the decomposed results skip the protrusive vertices in such intersections. During the decomposition procedure complex 2D shapes are converted into smaller areas, which lead to a reduced complexity of the skeleton computation, and some additional information for the pruning procedure is obtained. With help of the prior knowledge of the road network, the shared edges can be used as constraints to construct the topology of the entire network. The results are smoother and more favourable for human perceptions.

SPA and TRA are used to generate the pruned and reconstructed topology for the decomposed regions, and then the skeleton of the neighboring regions is connected via the shared edges. The minimal distance between the shared edge and the skeleton is the new approximated skeletons, and an example of this skeleton type is shown in Fig. 3.16(right) with the solid blue line.

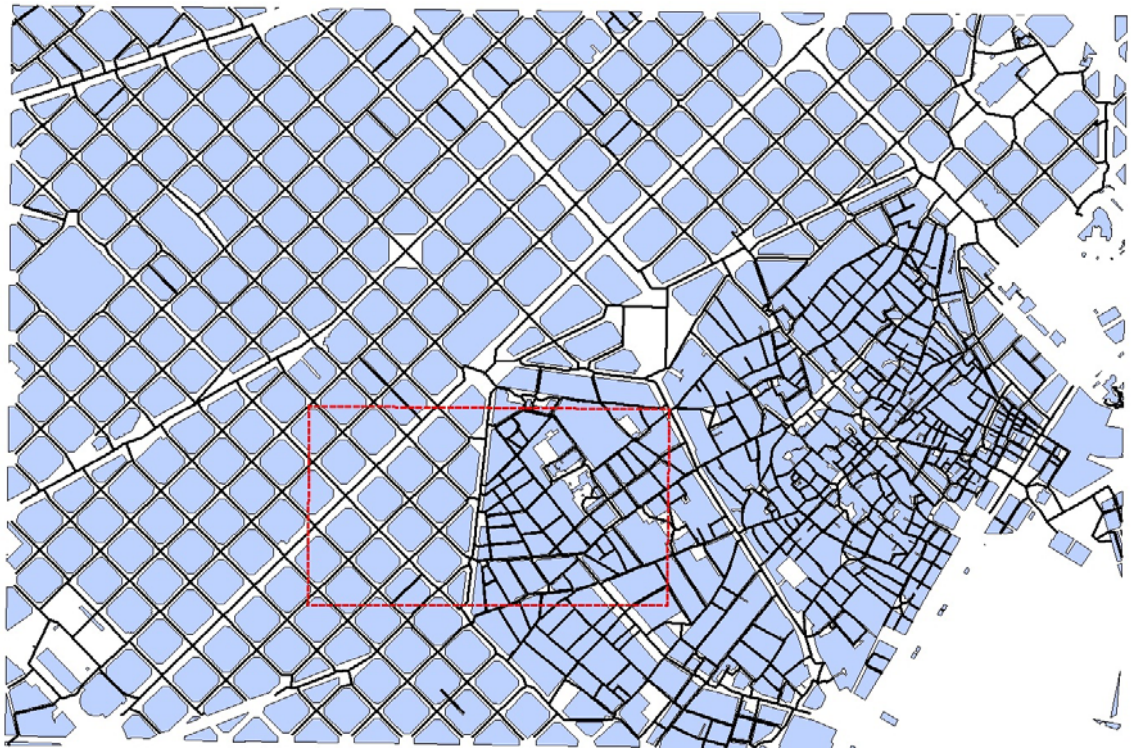


Fig. 3-18 Generated centrelines in the test area

Fig. 3-18 shows the final result of the test area. The centrelines are well detected in the new city area with regular shapes, but the intersection regions still have some problems. In an enlarged part of the test area as shown in Fig. 3-19 (the rectangle with dot line in Fig. 3-18), the details between the new and the old town are compared.

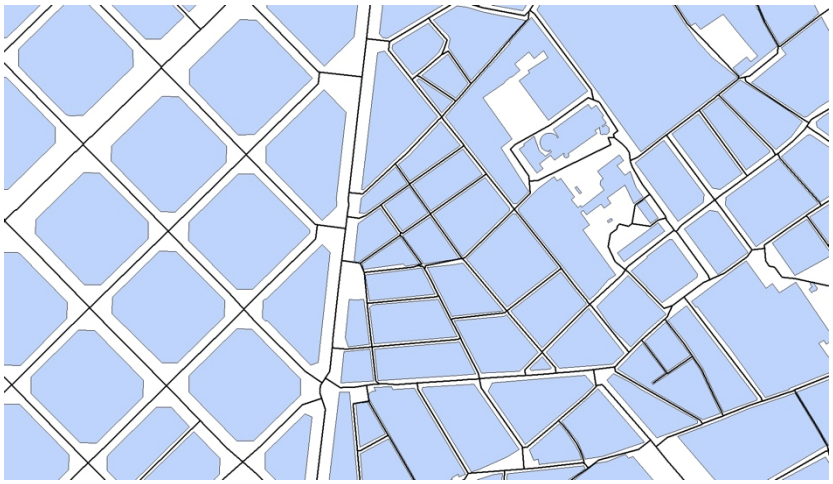


Fig. 3-19 Comparison of centrelines between the new and the old town

There are also some complex regions with irregular shapes. The TRA chooses the results directly from the Straight Skeleton algorithm as the final skeleton, but the dead-end segments are rejected.

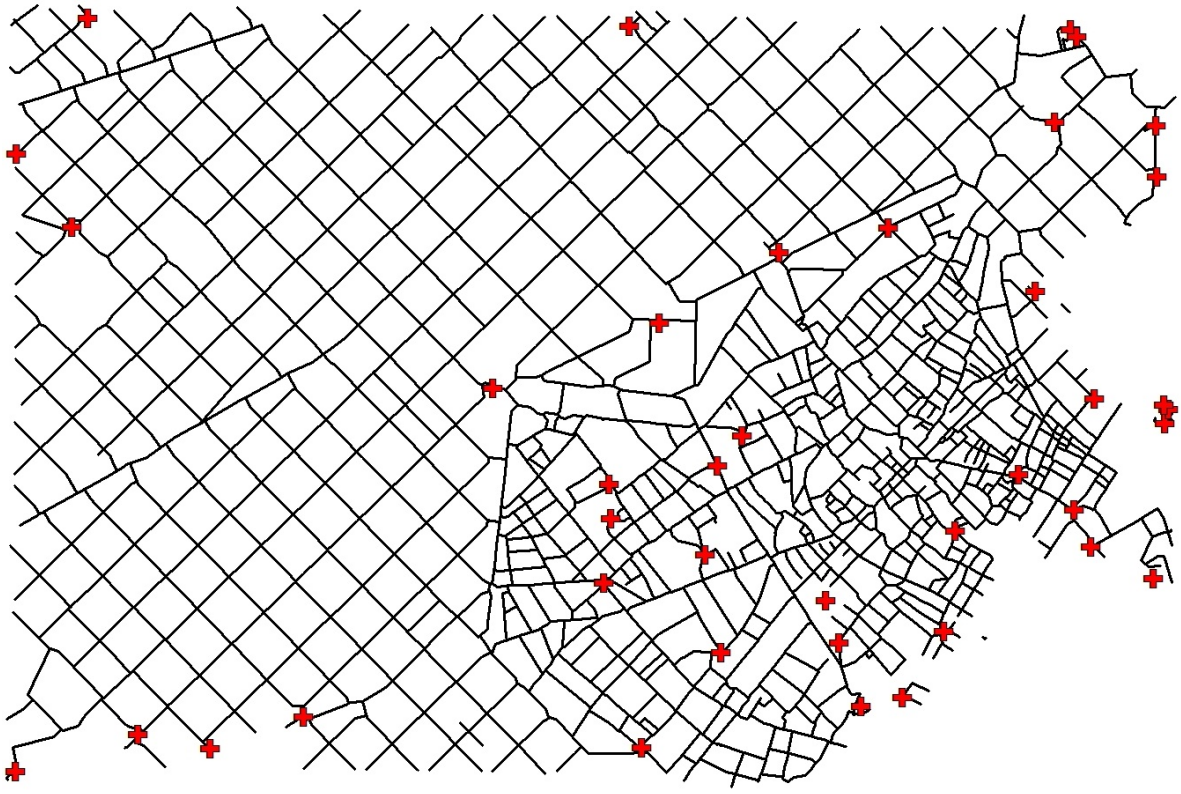


Fig. 3-20 Distribution of unsatisfactory results marked in red

Fig. 3-20 shows the distribution of the problematic regions marked in red. The solid lines represent the generated road network. The old town proves more problematic than the new town, and region in the border area is also more sensitive because of their fewer constraints during the reconstruction process.

Table 3-2 Correctness of each types of region

	Correct	**Incorrect**
Terminal region	84 (92%)	7 (8%)
Connection region	768 (99%)	6 (1%)
Intersection region	529 (95%)	27 (5%)
Total:	1389 (97%)	40 (3%)

Although the topology of the road centreline is reasonable, some unwanted connections still exist within the intersection region. Table 3-2 shows the relative correctness of the skeleton results in the decomposed regions.

3.3.5 Efficiency evaluation

The computing efforts include the decomposition, skeletonization and skeleton pruning. The skeletonization runs in $O(nm+\log n)$, the same as reported in (Felkel and Obdrzalek, 1998), where n denotes the total number of polygon vertices and m is the number of reflex ones, and the most time consuming part is to deal with the reflex vertices. In the worst case, the decomposition runs in $O(n^2)$, but definitely the number of the vertices with negative minima of contour and reflex angle is definitely less than n. The skeleton pruning also runs in

$O(n^2)$, so the proposed algorithm does not decrease the complexity, but it can minimize the number of the reflex vertices m through the decomposition following the perception laws, and the computation time for shapes with fewer vertices is more efficient.

3.4 Summary

We generate road network from additional data sources, for instance, from orthoimge and from spatial land parcels, besides the existing vector road network, which has been collected for car navigation purpose.

Two experiments revealed the different difficulty levels in the extraction algorithm. From the orthoimge, around 90% road segments have been extracted, but there is still much noise (40%) in the results, which include not only the road objects but also other spatial objects with similar geometric and radiometric character. Comparing to the existing road network, the high noise ratio is a challenge for the matching algorithms to get the spatial correspondence, and this is also a driving force for the development of new matching algorithms.

From land parcels, more than 90% correction ratio of the road network has been reached in the test region, which means the high spatial similarity between the two datasets. In previous research efforts (Zhang, 2009), it was exactly proposed for the matching approach in case of two road networks with similar data quality, so the conflation can be supported by the so-called delimited-strokes-based matching algorithm.

To focus on the matching problem with a high noise ratio, we solve the problem in two steps: reducing the multiple correspondences by a gestalt decomposition of the road network, and a new matching approach which optimizes the matching cost for the road network with regard to geometric and radiometric characters in two datasets.

Chapter 4

A Congruent Hybrid Model for Spatial Data Conflation

4.1 Overview

As already discussed before, conflation indicates to fuse two or more homogenous or heterogeneous data sources so as to generate a new single congruent product which inherits advantages from the existing spatial datasets for advanced applications. If both spatial datasets are composed of the same object type, such as road networks from different suppliers, and at a similar abstraction level, the homologous features will reveal a similar distributional density over the space. However, the conflation of the heterogeneous datasets has to cope with the large difference between the spatial datasets and the correspondences are usually sparse. The existing matching algorithm, e.g. buffer growing, has a notable problem that it adopts a sequential strategy which usually fails to achieve a global optimum (Li and Goodchild, 2010; Cobb et al., 1998).

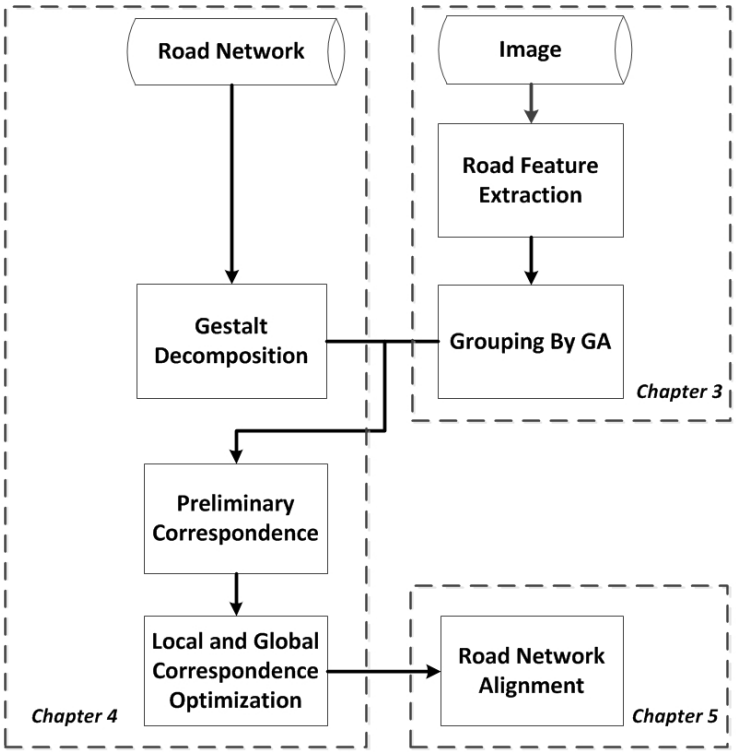


Fig. 4-1 Modules in the congruent hybrid model

The afore-mentioned challenges are addressed in this chapter with the objective to find correspondences between the geo-referenced image and the vector road network. Being aware of the complexity of the harmonization between the data sources, we name our approach a Congruent Hybrid Model (CHM) which includes three parts: feature extraction and

grouping; spatial data harmonization, and estimation of sparse correspondences. The first part has been discussed in Section 3.2, and we will focus on the two last steps in this chapter. The general procedure is shown in Fig. 4-1.

The spatial data harmonization is a pre-processing step with the aim to narrow the difference between the road network and the extracted lines from the image. After the harmonization, the conflation is reduced to a line-based matching problem which however, cannot be formulated as traditional graph matching. The geometry-based matching provides a possible solution. The existing line-based matching algorithms consider the nodes of the lines, but disregard road curvature (Safra et al., 2006), and our approach tries to overcome this shortcoming.

The estimation of sparse correspondence, i.e. Sparse Matching Algorithm (SMA), is a critical stage to get the optimized matching results. Due to the limited information, the SMA chooses a local-to-global strategy. To be more specific, the local correspondences, i.e. the decomposed cliques, are firstly estimated, which may include many inaccurate links, so a global optimization is necessary to clean the preliminary matched results using the global constraints, e.g. connection of neighborhood, the cost of possible deformation.

4.2 Spatial Relationship from the Matching Sources

4.2.1 Preliminary correspondences by gestalt (geometric) character

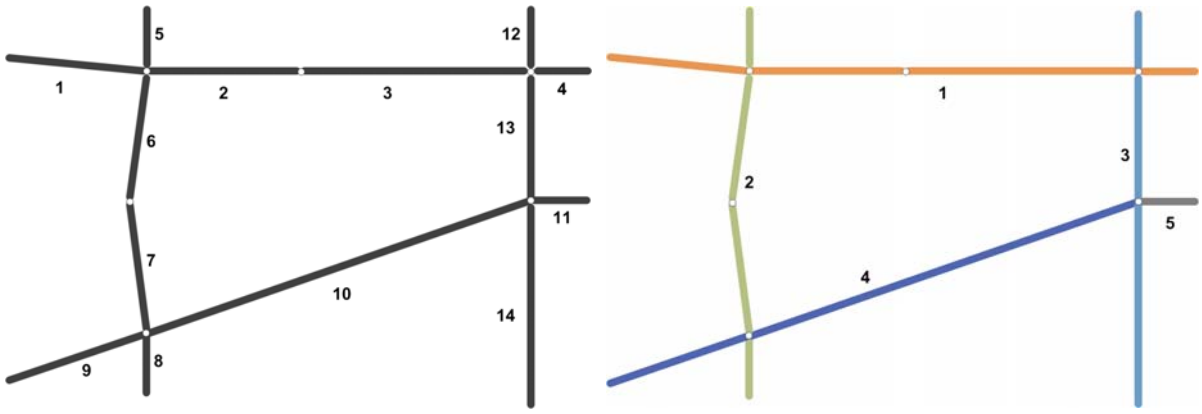
The datasets to be matched, i.e. the reference and target spatial datasets are usually represented in different granularity. For instance, the road network from the road database has integral topological information, and we also assume the topology is suitable for many practical applications, whereas the extracted road candidates from image have relatively low quality in terms of topology because they are quite fragmental and contain lots of gaps even after the grouping process.

From the geo-referenced image and road networks (Zhang et al., 2011), the displacements between the two datasets usually fall within a specific range of up to 200m. In spite of the relative small offsets, the exact localization of each other in the corresponding dataset is rather difficult because of the unsystematic distortion. Comparing the geometry of individual road segments from both datasets alone is not sufficient. First, the geometry in neither dataset is inaccurate; especially the extracted line elements are not always road features. Second, the abstraction level of both datasets is different: the extracted lines have a lower level of abstraction than the road networks, therefore, there is no straightforward way to compare the two datasets.

For a given reference model, i.e. the road networks, with $A \in \{A_1, A_2, \cdots, A_m\}$ connected segments, the correspondences to the target model, i.e. the grouped line elements $X \in \{X_1, X_2, \cdots, X_n\}$ from Section 3.2.1 with $m \ll n$, should be limited to the permutation of the features C_m^n , which requires an enormous computational effort. In order to avoid the computationally expensive approach which might lead to no reasonable solution for large spatial datasets, we calculate the preliminary correspondences for individual local areas based on the prior knowledge about the spatial datasets. The offset between the geo-referenced image and the road network is usually restricted within a reasonable threshold λ ,

although the correspondences are quite complex and can be multiple correspondences or no correspondence in each individual area if specific locations are considered.

An accurate correspondence depends on a reasonable spatial division based on geometric characters. An individual local region is in this sense a natural fragmentation of the road network. According to the gestalt principle, the straight linear features generally provide the most prominent information or proximity of a road network (Desolneux, et al., 2008), likewise, the skeleton of an irregular shape represents the most important character of the object. As we all know, the node-segment model is the dominating data structure of road networks in GIS applications, and the fragmentation of the spatial dataset by maximizing the length from the straight linear segments, so the straight as well as curved road segments can be recognized and regrouped from the road networks.



A) Typical road network based on node-segment model

B) Regrouped road network

Fig. 4-2 Decomposition of the road network following the gestalt principle

Different from the classical space decomposition approach, e.g. the Delaunay triangulation, the gestalt-based decomposition takes into account more regional information than the characteristics of the whole dataset. It is a decomposition procedure from the point of view of the entire dataset and a grouping procedure if only the individual road segments are considered. Similar to the GA grouping approach for the extracted line features and the decomposition approach for the land parcels, the decomposition of the entire road network leads to smaller groups with prominent topological characters which can be efficiently exploited in the matching algorithm.

***Definition 4-1*: Gestalt-based decomposition of a road network**

For a road network in the reference model with M segments and K nodes in planar Euclidean space $\mathbb{R}^2$, the decomposition of the road network is an iterative process of finding an optimal clique M_s of straight segments, which includes the maximal length of grouped segments below a specific curvature threshold ν and a by-product clique M_c of curve segments.

M_s and M_c completely divide the original space into individual local regions which are connected to each other, each with a reduced node matrix K' carrying the primary information for the matching algorithm. Each element in the clique indicates a prominent orientation of the

road, and the maximal length is more suitable to guide the accurate correspondence than the shorter segments. The correspondence-seeking problem between the road network and the grouped centrelines from the image is thus converted to the correspondence-seeking problem for each clique.

The decomposition is quite efficient for the road networks because the topology is already well organized. But the curvature of each segment needs to be individually considered so as to identify the segments with large curvature. The curvature of a specific road segments is expressed by the offset to the main direction. Rueda et al. (2008) suggested a local curvature scale to describe the straightness for individual segments, and we adapt this method for the road networks.

Let $P \in \{p_1, p_2, \cdots p_n\}$ be the shape points of a road segment, the curvature of this segment is estimated by the relations between the length of the longest connection among the shape points and the distance from other point to this connection.

The curvature at each shape point can be represented as:

$$K(p) = \frac{8S_p}{4S_p^2 + C_h^{\ 2}} \tag{4-1}$$

Where K is the curvature; p is point position; C_h represents the so-called *c-scale* element (Rueda et al., 2008); S is the distance from the shape point p to the *c-scale* element.

The curvature K of a specific segment is defined by the shape point which has the maximal curvature (Eq. 4-2). For instance, the segment between the end points p_s and p_e in Fig. 4-3 is the longest connection among all the shape points along the curve, so it is the *c-scale* element C_h in Eq. 4-1. The point p_{i+3} to the C_h is the distance S as shown in the figure.

$$K = \max(K(p_i))(i = 1, 2, \cdots n) \tag{4-2}$$

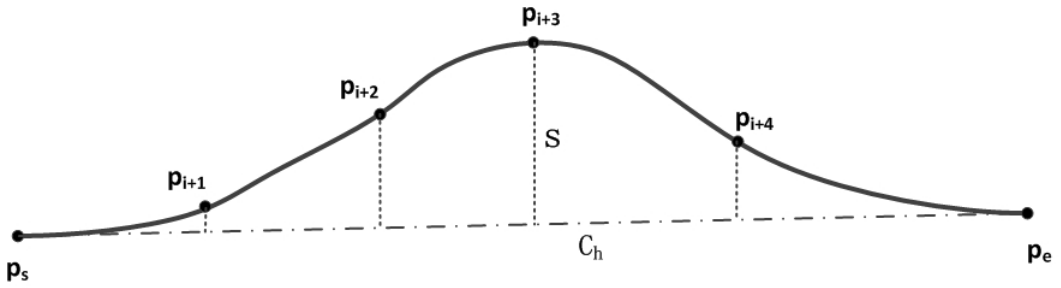


Fig. 4-3 Estimation of curvature of a road segment

In order to make clear the concept of curvature, we compare the curvatures of 4 different road segments from the whole road network to show the difference in more detail. According to the definition in Eq.4-2, the curvature depends on the maximal distance to the chord, and a longer chord has a smaller value for the same maximal distance. In Fig. 4-4, four segments have the similar curvature values, although they have diverse shapes.

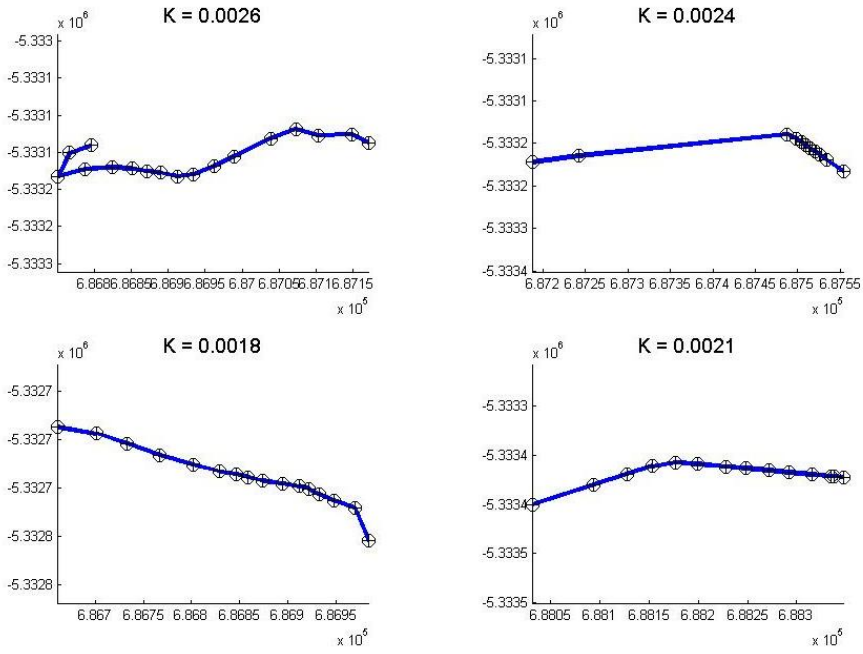


Fig. 4-4 Examples with similar curvature values from a real road network

The decomposition is implemented in an incremental manner, starting from the road segment with the maximal length. We summarize the detailed procedure in the Algorithm 4-1.

Algorithm 4-1: **Fast gestalt-based decomposition**

Input: 1) Road network $M \in \{V, E\}$ with nodes V and edges E;

2) Curvature threshold α and orientation threshold θ;

Output: 1) Gestalt clique M_s;

2) Reference matrix to the road segments $G \in \{M_s^i \to M\} \quad (i = 1, \cdots n)$

Steps:

Step 1) Sort road segments M by length in a descending order;

Step 2) Select the next unmarked segment in the sorted matrix M, if there is still a remaining unmarked segment, otherwise go to Step 6);

Step 3) Calculate the curvature of this segment: if it is larger than α, mark it as curve segment and go to Step 2); otherwise continue with Step 4);

Step 4) Extend the segment on both nodes, if the orientation difference between neighbor segment is less than threshold θ, then mark the segment in the sorted matrix M to indicate it has been used; continue the procedure of extension until the orientation difference is larger than θ, then save the passed segments in new clique M_s^i;

Step 5) Record the reference information from M_s^i to the road segment M, then go to Step 2);

Step 6) Return the result.

Comparing to the orientation threshold, this algorithm permits relative large tolerance for the curvature threshold, because the main purpose of the curvature threshold is to increase the matching quality in the next step if the segment is known as a curve.

Assume there are m segments in M, the computing complexity of Algorithm 4-1 is mainly dependent on the computation of sorting and extension, and the time complexity in the worst case is $O(m\log m + 2(m-1)^2)$.

The preliminary correspondence for both datasets can be derived from the prior knowledge, i.e. the offset is usually limited within a reasonable range, although non-correspondence is also possible for some segments. The generated road cliques, i.e. the decomposed groups of road segments, are the basic elements in the reference dataset, so the preliminary correspondences are the local matching results despite of the noise, if we consider the individual clique as an independent object.

The topologic information in the reference model has a relatively high reliability which can also guide the pruning of the noisy information from the preliminary correspondence, whereas the topologic information in the target model i.e. the extracted and grouped lines need also to be preserved and the data structure will be discussed in Section 4.2.3.

Analogue to the decomposition algorithm, the incremental strategy is also suitable for the computation of preliminary correspondence. However, a single line candidate in extracted data may be matched several times with the decomposed segments from the reference model. The preliminary correspondences can be identified by using Algorithm 4-2, and the accurate matching algorithm will be discussed in Section 4.3.

Algorithm 4-2: **Preliminary correspondence by predefined distance threshold**

Input: 1) Generated clique M_s;
2) Grouped lines N from image;
3) Distance tolerance λ;

Output: 1) Preliminary correspondences C^P;
2) Multiple-mapping matrix N_i^t;

Steps:

1) Select an element *i* from M_s;

2) Searching for the selected element its correspondences from the grouped lines N with the tolerance λ, then save the correspondence c_i;

3) Mark the lines in N with the number of correspondences n_i;

4) Stop the function if all the elements in M_s have been calculated; otherwise go to Step 1).

These preliminary correspondences are the candidates for the final accurate matching which follows the basic principle that the individual line in N has no more than one correspondence to the specific clique m_i from M_s, so the multiple-mapping matrix is a starting point for noise reduction. We visualize the preliminary correspondence of the road segments (the longest green line) with the lines (in black) of the clique in Fig. 4-5, where the height i.e. z-axis means the quantity of tolerance from the candidates. The value has been added if there is overlapping area between two candidates. The prior distance tolerance λ is proportional to the segment's length, and the value in each grid has been simply added if

there is more than one correspondence inside. The influence from the neighbor clique has been omitted for convenience.

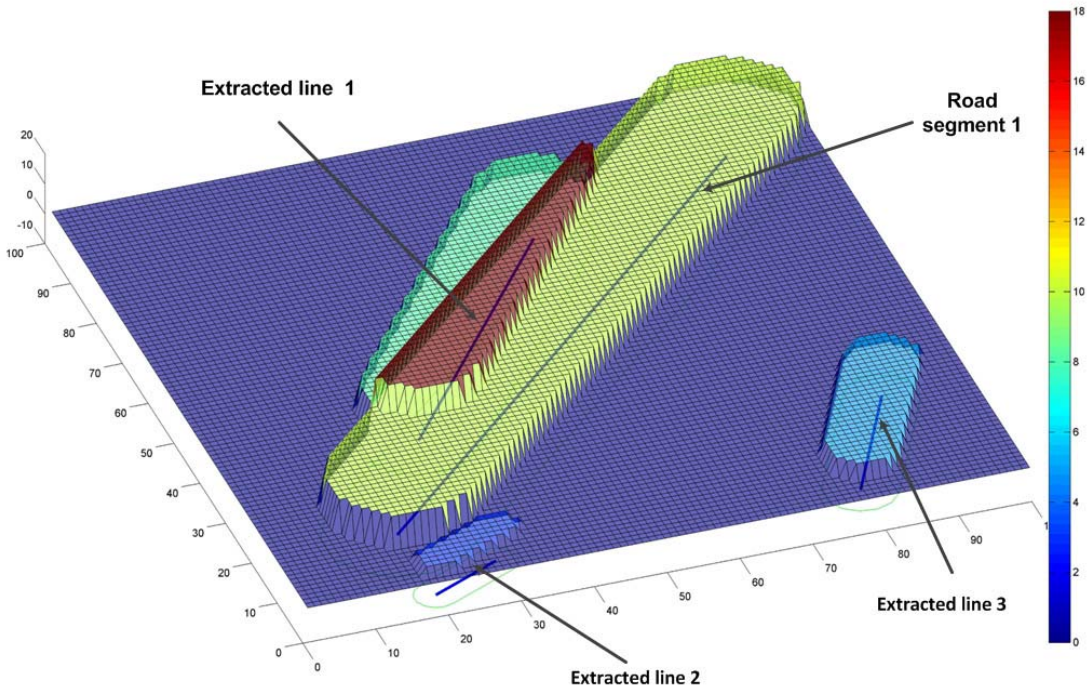


Fig. 4-5 The preliminary correspondences by dynamic λ

The computation of Algorithm 4-2 is quite efficient, and it has the time complexity in $O(mn)$ for m elements in M and n lines in N. Comparing to the GA grouping algorithm in Chapter 3, this time complexity can be ignored, although both computations can be done in pre-process for a huge spatial dataset.

The preliminary correspondence is mostly dependent on the distance measure, where the Euclidian metric is fundamental for Algorithm 4-2. However, other similarity measures also count for the final accurate matching.

4.2.2 Similarity measure

In many applications including image registration (Modersitzki, 2007), object recognition (Li et al., 2012) and cartographic objects matching (Zhang, 2009), the suitable similarity measure or dissimilarity measure is the essential factor. In matching approaches (Veltkamp, 2001), Euclidean distance is the most favourite similarity measure, and usually it has been chosen as the default similarity measure, especially for point matching approach because the distance is the only available information from points. However, the Euclidean distance is not sufficient if the datasets to be matched reveal unbalanced number of objects such as road segments, additional similarity measures are necessary to be integrated in the matching process so as to get the optimal correspondences.

The similarity measure for line-based matching algorithm has an advantage in using the geometric and topologic information from the road networks, although it also increases the complexity of algorithm. A number of methods have been developed based on the idea of graph-matching (Conte, 2004), which captures the relationships such as left of, right of, cycles,

collinear with the specific object etc, as well as topological connectedness, however, the requirement of full topology for both matching datasets is impractical for the extracted features from the geo-referenced image, which contains many broken lines even after the grouping process.

After the gestalt decomposition, the matching task has been divided into two related subtasks: 1) the exact correspondence in each decomposed area; 2) the optimal correspondence in the whole area. The former is to get the local matching results, whereas the latter is to pursue the optimized global correspondences. The similarity measure is decisive for the first subtask. In Fig. 4-6, it shows a spatial relationship example of the road segments (in dash line) and extracted line elements (in solid line). Five geometric measures are described here.

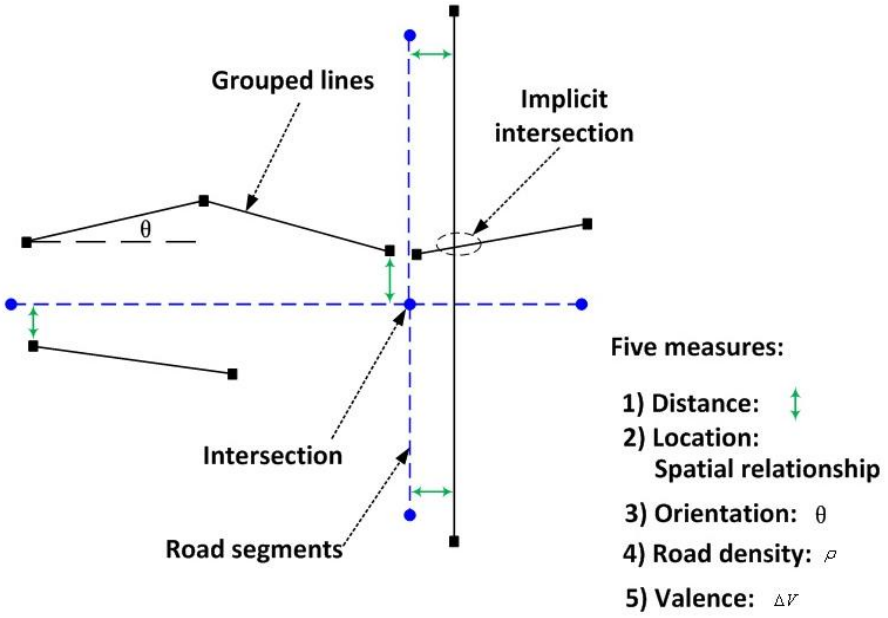


Fig. 4-6 Similarity measures for line-to-line matching

A single similarity measure such as distance, can hardly describe the complex spatial relationship between the elements in both spatial datasets. We have identified five different similarity measures which can complement one another for the line-to-line matching. These measures include two types: 1) the geometric parameters: distance, location, orientation, road density, and 2) the topologic parameter: valence in the local clique.

The distance measure, e.g. L_p distance, Bottleneck Distance, Hausdorff Distance, Frechet Distance, Turning Function Distance, are the common similarity measures in various matching approaches (Veltkamp, 2001). Frechet Distance is the most suitable measure to represent the similarity between two curves. Essentially, the Frechet Distance refers to the minimum distance among all monotonically increasing functions of the maximal distance $d(\cdot)$, and in mathematical notation it is

$$F(A,B) = \inf_{\alpha,\beta} \max_{t\in[0,1]} \left\{ d(A(\alpha(t)), B(\beta(t))) \right\} \tag{4-3}$$

where A and B are two curves; $d(\cdot)$ is the Euclidean distance between two points; α and β are arbitrary continuous non-decreasing functions with parameter $t \in [0,1]$.

Although the Frechet Distance is a powerful measure to represent the similarity between two curves, it is limited if the segments in the local clique and the grouped lines from the extracted line elements do not reveal the simple one-to-one correspondence, rather one-to-many or many-to-many correspondences. This principle of the conflation makes the distance measures not so effective. The offsets in the datasets to be matched are non-systemic distortions, which mean that the distance alone can not represent the similarity among the features. The correct matching may be the farthest instead of the nearest neighbor of the current object. In such cases, the relative location of the features may additionally contribute to the similarity.

The location is a related concept with the distance measure, but it focuses on the spatial position. Usually two or more data sources are considered. However, similar to the distance measure, location along is insufficient for the determination of similarity, because the nearest intersection is not always the corresponding road crossing. In Chapter 2, we already discussed the line-to-line spatial relationships, which considered the distance and location at the same time. For multiple spatial data sources, the distance and location measure can be integrated to derive the preliminary correspondences via the Minkowski sum, i.e. dilation in geometry. In fact, both location and distance measures have been used to compute the local cliques in Algorithm 4-1.

The length of the road segment is also an essential measure for road network matching approach (Zhang, 2009), however it usually requires that both datasets have the similar quality, in particularly, the similar topology, which is quite difficult for the extracted features. In order to identify the matching pairs by utilizing the length of the segments, we introduce an associated similarity measure: the road density, which indicates the standardized ratio between the length of total road segments and the area of the local clique.

$$\rho_i = \frac{\sum_{k=1}^{K} l_k}{S_i} \tag{4-4}$$

where ρ is the road density in the local clique; S is the area of the clique with $i = \{1, 2, \cdots N\}$; l is the length of the road segment in the clique which has K segments with $k = \{1, 2, \cdots K\}$.

In an ideal situation, ρ in corresponding cliques are quite close to each other, but they may be different if one or both cliques contain much noise or there are no matching pairs. The noise needs to be filtered away in an additional procedure.

The orientation is yet another useful measure that can describe the similarity between two lines, particularly between two straight lines. Collinearity is an explicit representation of the spatial relationship between lines. But in matching datasets, the orientation becomes unstable because the corresponding segments in different datasets can hardly keep the similar orientation, especially when the two datasets have large difference in a specific clique. Even so, the orientation can be used as a suitable similarity measure because the orientation of an individual segment is relatively stable as compared with its neighboring segments. Furthermore, the orientation is naturally related with the location, and the relative orientation of line elements is the focus of attention according to the gestalt principles. To be more specific, the relative orientation touches two related aspects: 1) the relative orientation inside its own dataset, i.e. among the neighborhood, and 2) the relative orientation to the partners in the

preliminary corresponding outcome. The calculation of the orientation for a specific clique will be discussed in Section4.3.1, where the orientation is the prerequisite of the matching cost.

The valence measure, i.e. the connections at a specific intersection, is the topological character of the road network, as shown in Fig. 4-7. If the topology in the reference dataset is relatively stable, it can serve as an important factor for the computation of the matching pairs. However, the topology in the extracted dataset is more unreliable than in the reference dataset. There are three instances of the valence measure when the road segments match the extracted line elements: 1) the same number of valence and road networks and extracted line elements have perfect matching; 2) the same number of valence, but the matching results need to be verified in further functions; 3) the unbalanced valence between the road networks and extracted line elements.

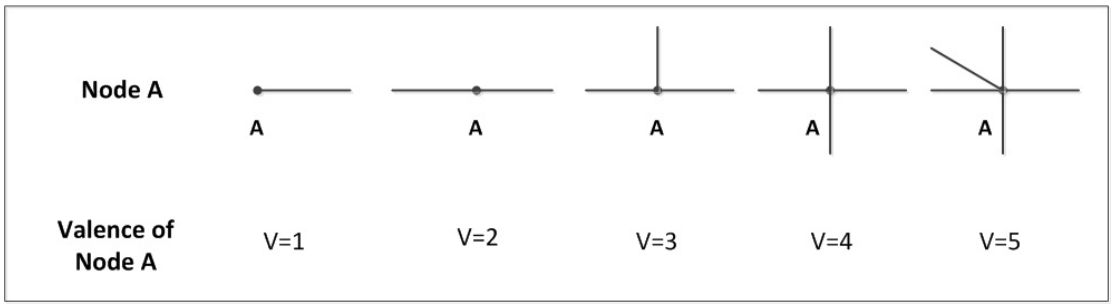


Fig. 4-7 Valence of the road intersections

The importance of the valence measure is based on the principle that the nodes or vertices with the same valences are potential matching pairs from both datasets. In fact, the final objective of the matching algorithm is to find the exact matching pairs of the nodes from the road segments and the extracted line element.

As already discussed before, the final matching result can not be derived from a single similarity measure because the quality in the matching datasets is rarely consistent. Therefore, we integrate the multiple similarity measures in a high-level algorithm. The subsequent section is dedicated to the data structures used in our matching approach.

4.2.3 Multi-level representation of road network

The implementation of road matching algorithm necessitates a suitable data structure of the road network which is no easy issue due to different qualities in the road network datasets. Moreover, the data structure is required to meet the requirements of many operations ranging from searching to the reconstruction of road network which are the critical processes for the accurate correspondence computation. Different from the route planning algorithms aiming to seek routes that satisfy certain optimization criteria, the extracted and grouped lines from image need flexible data structures to deal with the broken lines. The rearrangement of the ‘road network’ is a necessary procedure to fill the gaps. We design a multi-level data structure to cope with this challenge for various spatial data sources.

The node-segment model is the basic spatial data structure in modern GIS, and it mainly explores the geometric characters of the underlying spatial data. The road datasets in our case have the similar character. No semantic information is available from image, and we have to rely heavily on the geometric information. Our data structure contains three different

levels: 1) ground level; 2) grouping level; and 3) dynamic level. We show the definition of each level as well as the features in Table 4-1.

Table 4-1 Multi-level data structure for the representation of road network

Level	Definition	Feature
ground level	The basic node and segment of road network, or the extracted lines.	node, line
grouping level	The rearranged road segments following the gestalt principle or line elements by the GA grouping algorithm.	gestalt lines
dynamic level	The reconstructed or optimized lines or segments from the preliminary correspondence, and it may change during the estimation of the exact correspondence.	reconstructed lines

The above multi-level data structure is specially designed for the matching approach that works on both data sources with large quality differences. In principle, each level corresponds to a different operation for the matching approach.

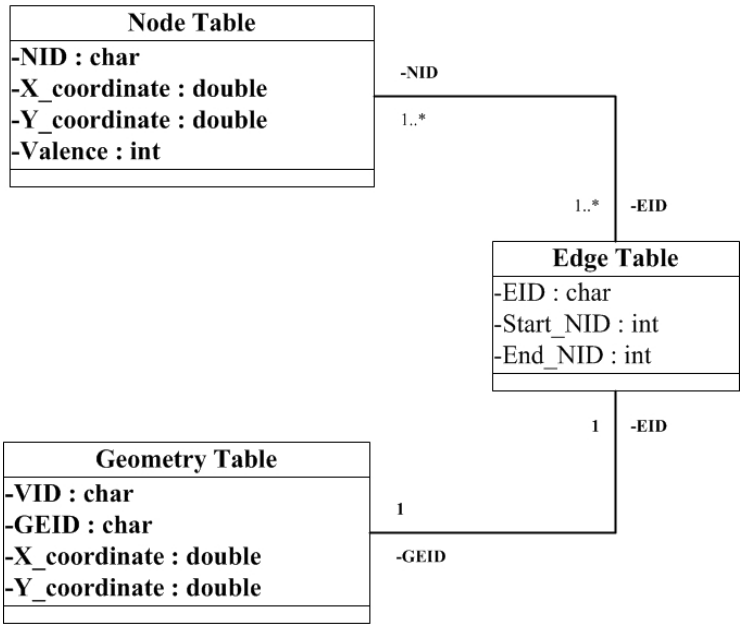


Fig. 4-8 Data structure for the ground level

The ground level deals with the basic elements of road network - node and segment, and the elementary topological information. To be more specific, it includes three main tables to record the information from both data sources, namely Node Table (*NT*), Edge Table (*ET*) and Geometry Table (*GT*). The ground level does not care about the data quality, and it simply keeps the entire information of the spatial data. The *NT* contains the entire nodes of the road network incl. the valence of each node which is useful for other operations in the matching process. The *ET* records only the starting and ending node for each segment, whereas *GT* contains the detailed information of shape points from each segment which is usually called

polyline. For the extracted lines, it stores the line elements from ECM detector instead of the extracted edges. The data structures in the ground level represent the characters of both spatial datasets. The information of edge has been used as the commutative media for the nodes and geometric information, so the topology has been preserved for the next procedure in the matching algorithm. The keyword in each table has been identified with '*ID' respectively. The data structure of ground level is shown in Fig. 4-8.

The grouping level deals with the road segments at an aggregated level. The edges have been chained to form new and longer segments with larger prominence than the individual fragment lines, or the road network has been decomposed into individual cliques which include several segments with the similar character. The prominent segment or clique has been used to fill the gaps in the partner data source, because we already have the prior knowledge about the quality in the datasets to be matched. As shown in Algorithm 4-1 as well as the grouping in Section 3.2.2, the new segment is derived from the existing segments and nodes following the gestalt principle. The new segments do not change the geometry of the road segments, so the data structure can be simply adopted with a minor modification from the *NT* and *ET* at the ground level, as shown in Fig. 4-9.

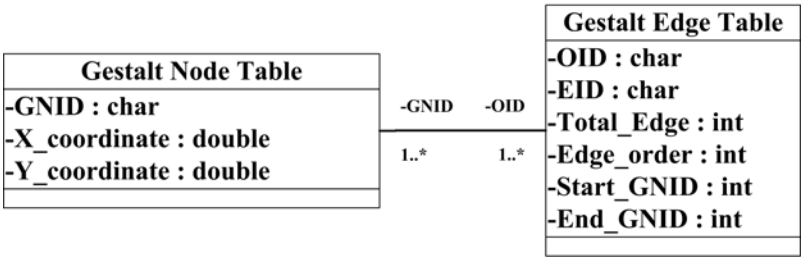


Fig. 4-9 Data structure for the grouping level

The dynamic level deals with the auxiliary features for the computation of the correspondences. The poorer the quality of the spatial dataset, the more auxiliary features it needs. During the matching procedure, either some new segments should be added or some existing segments should be ignored so as to get the optimized matching results. Since the gaps in the dataset of lower quality can not be directly recognized, an iterative procedure is necessary to test the usability of each segment. Moreover, the intersections in the extracted line element are not always explicit. Therefore, we need additional approximation of the possible location for the intersection which can be estimated from the centrelines of grouped segments and has a temporal nature.

Both the auxiliary segments and nodes are recorded at the dynamic level, but it is also based on the data structure of the ground level. The new nodes or segments will be created only if they are not included in the ground level. The data structure is shown in Fig. 4-10.

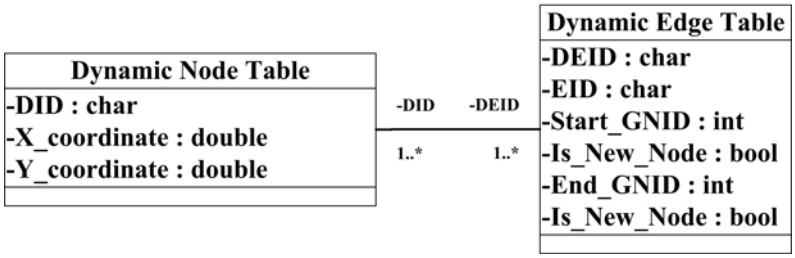


Fig. 4-10 Data structure for the dynamic level

In the pre-processing step, both spatial datasets to be matched are firstly arranged according to the data structure of the ground level, and then the datasets are rearranged into the compact datasets at the grouping level. In the matching procedure, new auxiliary segments need the dynamic level data structure, which permits many alternative operations during the optimization of the matching results.

4.3 Sparse Matching Algorithm

4.3.1 Refine the preliminary correspondence

For each local clique, we already get the preliminary correspondences from Algorithm 4-2, which only utilizes the distance and location as the constraints, but it implicates two critical problems: 1) the local correspondence in each clique contains many false corresponding segments; and 2) the single line could be included in two or more different cliques. The refinement of the preliminary correspondence is necessary in order to solve these problems, and the procedure is only restricted locally around each clique. Moreover, the prior topological information in the road network is useful to reconstruct the spatial relations for the extracted lines. The approximated intersections from the grouped lines are the essential candidates for the entire conflation approach.

As discussed in Section 4.2.2, the geometric and topologic similarity measure is reasonable for our conflation approach, especially for this refinement procedure. Regarding the multiple preliminary correspondences from both datasets, the essential task is to find the location of the matched node as well as the line orientation in the feature. As already known in Chapter 2, the fragmental lines and over extraction from image make the end nodes of the line unreliable for the matching procedure, and the neighboring features have to be used as the additional information. Fig. 4-11 shows the idea of the refinement procedure in a clique: the individual segments in each clique are classified based on the difference of the orientation, and then the line group with small orientation difference from the segments in the road network is chosen.

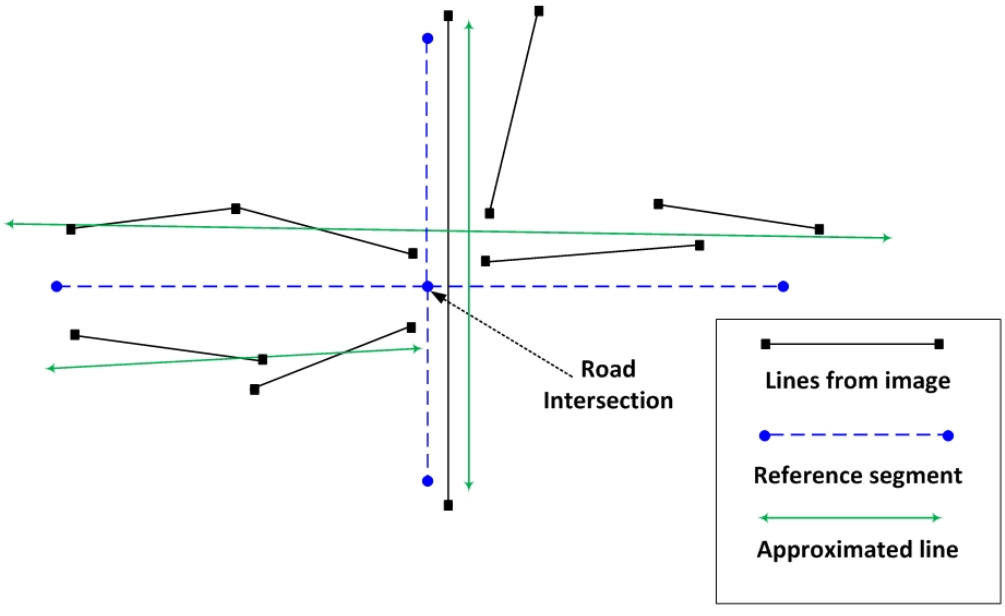


Fig. 4-11 The refinement of preliminary correspondences

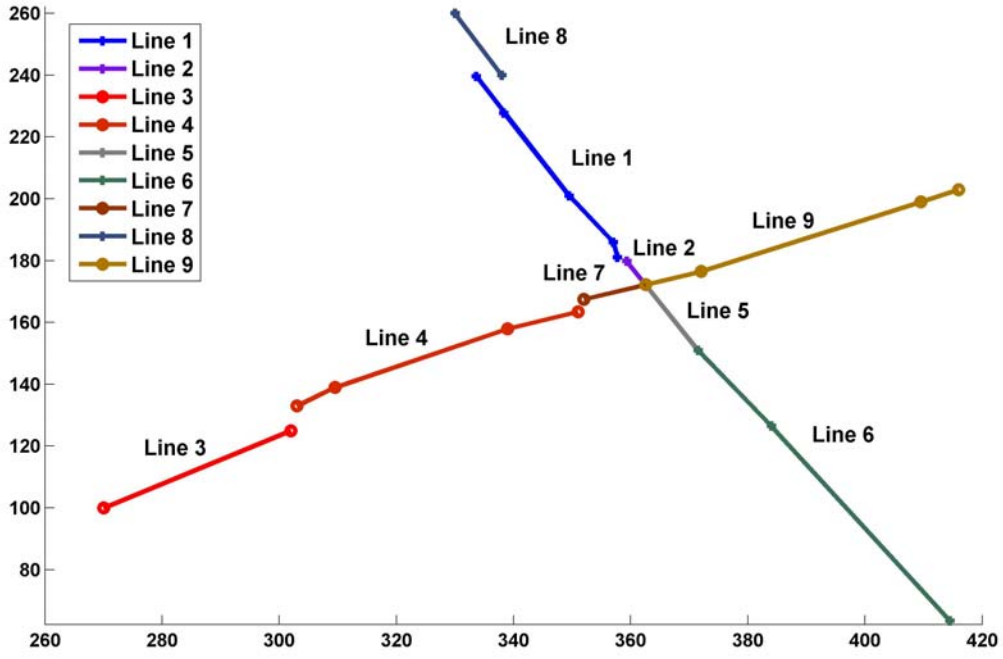
The extracted lines (solid lines) in Fig. 4-11 contain the road candidates of another clique, so they reveal diverse orientation characters, but the lines with the similar orientation to the road segments are the most prominent candidates, which can be recognized by the classification algorithm. The lines with large orientation difference can be filtered out, but there could be two or more lines which have the similar direction as the road segment.

Wen (2006) reported that the orientation of the line elements is more reliable than other measures such as length and location. The centrelines from our *ECM* detector have the same characteristics and they lie around the centre of the objects in principle. Assuming we have *n* extracted lines in the i^{th} clique $C_i = \{l_1, l_2, \cdots l_n\}$ and the orientation difference to the road segments is $\Delta O_i = \{\Delta o_{l_1}, \Delta o_{l_2}, \cdots \Delta o_{l_n}\}$, so the aim is to find the group of lines which have the smaller difference in ΔO_i . It is a typical classification problem, and *k*-means reported in (Jain, 2010) is a suitable approach for the problem.

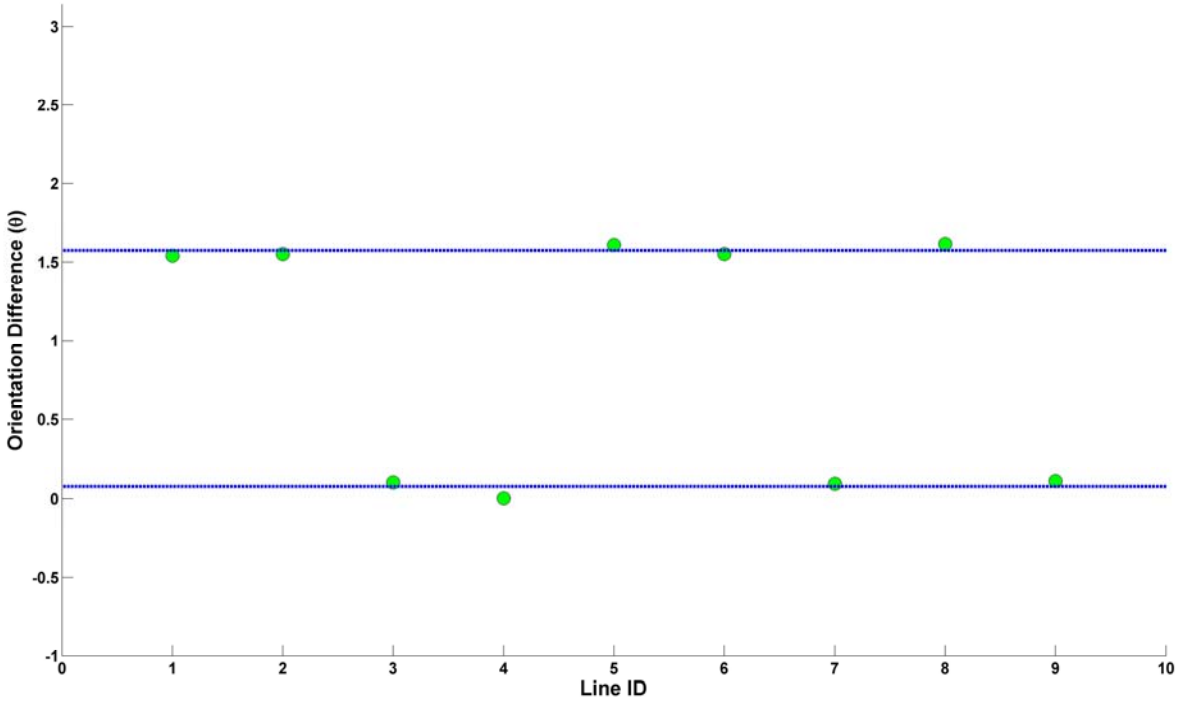
$$J = \sum_{k=1}^{K} \sum_{i=1}^{I} \left\| \Delta O_i^k - c_k \right\| \tag{4-5}$$

where *K* is the number of the lines in the group, $\|\bullet\|$ is the Euclidean distance; ΔO_i^k is the i^{th} variable in the k^{th} group; c_k is centre of the k^{th} group.

K-means method is used to find a suitable cluster such that the squared error between the empirical mean of a group and the data in the cluster is minimized. If we set *K=2* in the computation, we will get the group with the smallest orientation difference to the reference maximal road segments.



(a)



(b)

Fig. 4-12 Orientation classifications for one clique using K-means

Fig. 4-12 shows the classification results of an example with a randomly selected clique which includes 9 individual lines (the green point). From the K-means method, the group centres are [0.08, 1.57] respectively, and they are expressed as dash lines in the figure. The group with smaller difference is the expected groups. The detailed information about the grouping result and the distance to the respective group centre is summarized in Table 4-2, and the distance to the group centre has been minimized with the total sum of the distance being around 0.01.

Table 4-2 Grouping result using the k-means

Line ID	Orientation Difference (Degree)	Distance to group centre (Degree)	Group ID
1	1.54	0.0001	2
2	1.55	0.0001	2
3	0.10	0.0001	1
4	0.003	0.006	1
5	1.61	0.001	2
6	1.55	0.0001	2
7	0.09	0.0001	1
8	1.61	0.002	2
9	0.11	0.001	1

The preliminary correspondence is derived from the distance constraint and it contains the most important line elements near the longest segments in the road network. The orientation alone, however, may also generate inaccurate line candidates. In order to include as many as possible lines for the optimization process, especially for very noisy data, we also consider the total length of the lines. From the above example, if the group of lines with smaller orientation difference to the corresponding maximal segment clique is $G_i = \{l_1, l_2, \cdots l_k\}$, we modify Eq. 3-3 to get the ratio between the two line groups:

$$r_i = \frac{\sum_{k=1}^{K} l_k}{L_i} \tag{4-6}$$

where r_i is the length ratio for the i^{th} clique; l_k is the extracted lines with $k = \{1, 2, \cdots, K\}$; L_i is the total length of the road segment in the i^{th} clique.

The constraint in Eq. (4-6) as supplement is employed only if the orientation constraint does not generate any line candidates. The suitable threshold for r is strongly dependent on the complexity of the application, and we choose 40% in our test.

Using the orientation and the length constraints, each clique keeps the lines with similar orientation, but one single line can be counted by more than one clique, if the cliques are near to each other. To eliminate the ambiguity during the intersection estimation, where each line can be calculated only once, the confliction information is kept in a matrix which records the line ID as well as the clique ID, and the priority is given to the clique with long segments used to calculate the intersection.

In each clique, the refined line elements keep the similar orientation, but they could also belong to different approximated lines. The principle to group the lines with similar orientation is to compute Euclidean distance to the primary direction. The multiple groups may cause multiple correspondences of the road intersection, where an optimization process is necessary to compute the minimal total bending cost.

The orientation and length constraints as well as the distance constraint in Algorithm 4.2 filter out the noisy data in extracted lines from the image, and the results are the partner group of lines corresponding to the longest segment in each clique. However, the current correspondence is still ‘noisy’, which means that the multiple correspondences still exist, i.e. each of the longest segments may have zero, one or more than one correspondence in extracted line groups. But one thing is certain that the corresponding line groups reveal the similar orientation to the road segments, so the preliminary topology in the road network can be introduced to the extracted line elements at intersections which exist in the corresponding line groups either explicitly or implicitly. If we can convert the multiple correspondences from line groups into intersection groups, the optimization can utilize both the line groups and the intersection to estimate the similarity between the objects in the reference and the target data sources, which will be described in Section 4.3.3.

The valence v of the intersection in the road network is derived from the topological relationship which is the number of the connected road segments in each specific road crossing. Two intersection classes are identified according to the valence: 1) the terminating node of a dead end ($v = 1$) and 2) the normal intersection ($v > 1$). The former can simply

acquire its matched points from the corresponding cliques. The node at the other end of the road segment which usually connects with other road segments can be determined.

It is a simple intersection problem if there are only two line candidates, but the complexity to estimate the intersection increases dramatically with the growing valence. The situation gets even worse if the orientation error ε_i for each line group is considered. In such case, a single intersection of multiple lines does not exist. There exists an intersection area instead.

To simplify the problem, we choose the mean orientation of each clique to indicate the final direction of the line group. For a intersection with valence = *n*, $(n-1)!$ intersection candidates may theoretically exist, and the best one can be determined with help of a distance filter *T* and the minimized orientation angle difference.

According to the uncertainty model from (Shen and Palmer, 2000), the possible location of the end point of the extracted lines can be described by the exponential probability density function:

$$f(d) = \begin{cases} e^{-\alpha d} & d \geq 0 \\ e^{\beta d} & d < 0 \end{cases} \tag{4-7}$$

where α and β are parameters determining the uncertainty degree in extending and shortening the end point along the line by a distance *d*.

Besides the uncertainty model, the intersection is located within a limited area around the corresponding node in the road network, e.g. < 200m. These two requirements can serve as rough filters through which the candidates with large errors are excluded; the final candidate is selected according to the minimal variation of the orientation. For a candidate intersection *i*, its summed orientation difference C_i can be calculated by:

$$C_i = \sum_{k=1}^{K} \Delta\theta_k \tag{4-8}$$

where $\Delta\theta$ is the orientation difference from the line elements in the image and *K* is the number of lines with orientation differences. Fig. 4-13 illustrates a group of lines and the tolerance area T of the node from road network. The final candidate of the intersection is the one with the minimal value of C_i. In the example, the intersection is derived from line I_3 and I_4, which of course has no orientation difference, but line I_1 and I_2 have small difference if the intersection i is the candidate with the minimal orientation difference for the line group.

The generated intersection builds up the spatial relationship between the line groups for each clique from the road segments, but it does not change the nature of multiple correspondences in both data sources, which means that the intersection correspondence is also one-to-many. To be more specific, one node in the road network, there may exit many intersections among the extracted line elements.

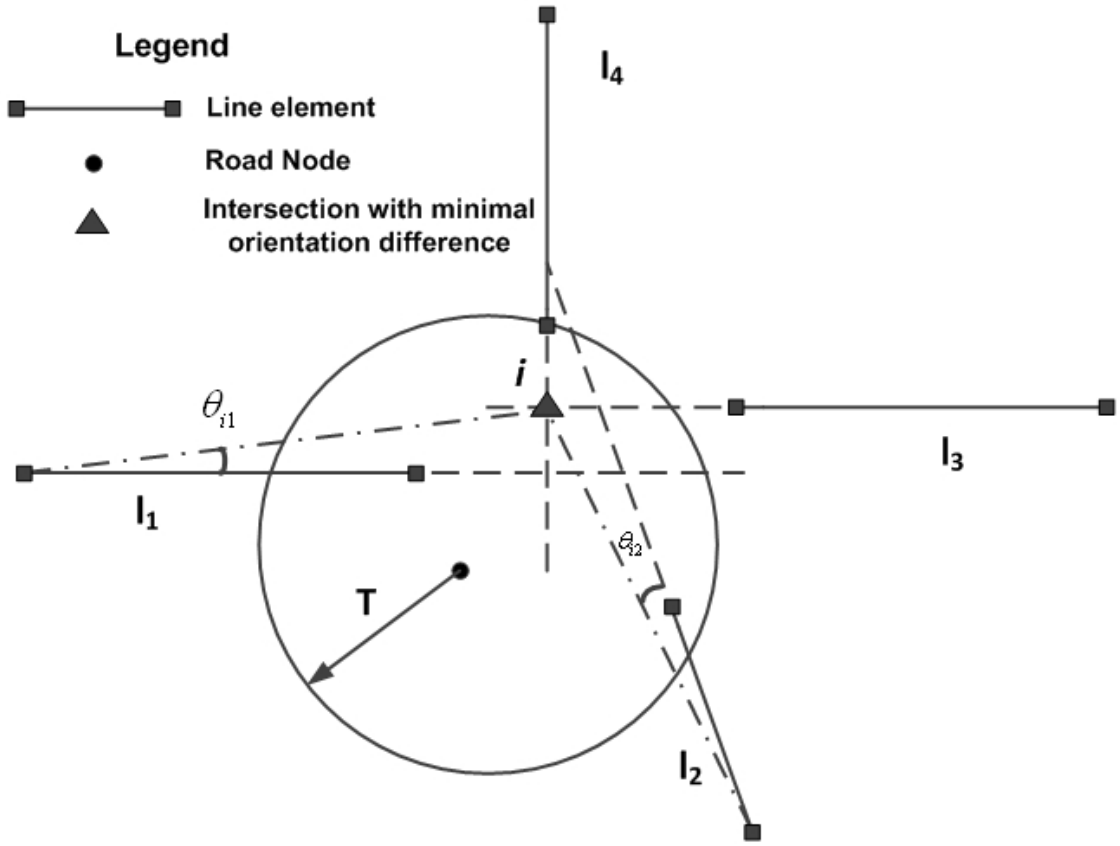


Fig. 4-13 An intersection candidate within the tolerance scope

4.3.2 Definition of the optimization problem

The multiple correspondences from both data sources have been converted to the correspondences between intersections which are the desirable features to build up the spatial relationships. Furthermore, they can be used as the control points in the transformation function to compute the final results. We use an optimization algorithm to eliminate the false correspondences.

For each clique $m_i = \{s_1, s_2, \cdots s_k\}$ $(i \in \{1, 2, \cdots I)$ from the road network, it has multiple partners in the extracted line element groups $G = \{g_1, g_2, \cdots, g_n\}$, where each line element group g_j contains one or more individual extracted line elements, i.e. $g_j = \{l_1, l_2, \cdots l_J\}$ $(j \in \{1, 2, \cdots J)$. The road crossing t_i in group *T* from the prior road network has also multiple partners in the estimated intersection group N_i from extracted line elements, noted as $t_i \leftrightarrow N_i$, where N_i is a group of intersection $N_i = \{n_1, n_2, \cdots n_J\}$, and we name it as super junction to avoid the confusion with the concept of intersection in this thesis. However, the prior knowledge grants only one-to-one matching from the road network to the extracted features, and in the worst case, for instance, if both datasets are inconsistent at the intersection, no matching exists. But we will focus on the former in our test, because the change ratio of the road is the relatively low comparing to the whole road network.

In Fig. 4-14, it shows a case of multiple correspondences at intersection I_4, which has two partners $\{n_4^1, n_4^2\}$ in the line elements. It's assumed that the neighboring nodes $\{I_1, I_2, I_5\}$ have already got their partners $\{n_1^1, n_2^1, n_5^1\}$, and the segments in $\{M_1, M_4, M_6\}$ have the corresponding partners $\{\{L_{11}, L_{12}\}; \{L_{41}, L_{42}\}; \{L_{61}, L_{62}\}\}$ respectively.

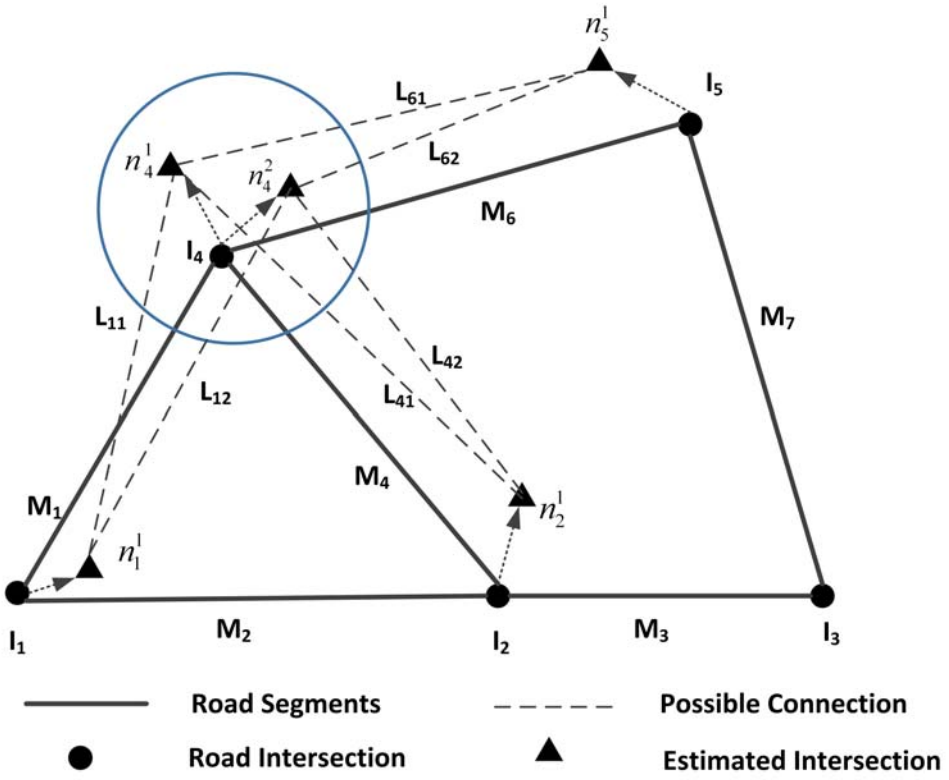


Fig. 4-14 Multiple correspondences for nodes

The multiple correspondences spread over the entire road network instead of around an individual node, therefore the solution should consider the local as well as the global influence. In this way, the basic matching problem has been transformed into the optimization problem with the aim to find the optimal intersection from super junctions N_i to reconstruct the road network from the features with minimal difference to road network. However, the selection of the suitable candidates from N_i for t_i is not a trivial task because it influences always the neighboring intersections, therefore the local optimal solution can not guarantee the global optimal, which requires compromise in a specific application.

Let's take an overview of the super junctions in a small network as shown in Fig. 4-15. The solid lines are the topological information from the road network, whereas the dash lines are the virtual connection i.e. the possible connection between two nodes. The cost for each virtual connection can be derived from the similarity estimation from the matched segments. The multiple virtual connections at the super junction N_i represent the multiple correspondences for the matching results.

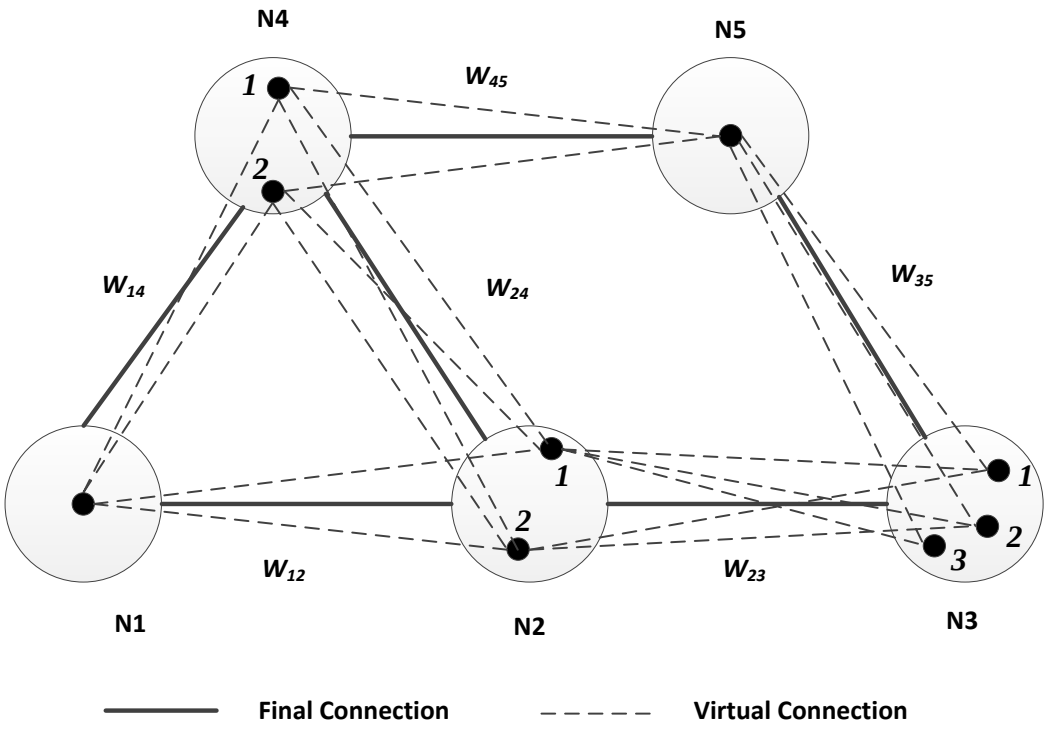


Fig. 4-15 Networked multiple correspondences

If the connection can be ignored, i.e. there is no topological constraint for the super junctions, we can generalize the problem as Generalized Travelling Salesman Problem (GTSP) by Laporte et al. (1987) with the aim to find the minimal cost for the salesman to visit each city (node) clusters.

Generalized Travelling Salesman Problem (GTSP): given a clustered graph $G=\langle N,E,C\rangle$ with K super junction sets N, edge set E and the edge cost set $C:E\to I^{+}$. Each super junction contains multiple points $N_i=\{n_1,n_2,\cdots n_{i_k}\}$ and has no common point with other super junctions:

$$\begin{cases} 1)\bigcup_{i=1}^{K} N_i = N \\ 2)\, N_i \bigcap N_j = \varnothing \end{cases} \quad \forall\, i,j=1,2,\cdots K, i\neq j \tag{4-9}$$

We seek a minimal cost subgraph $S=\langle V,Q\rangle$ with $V=\{v_1,v_2\cdots,v_K\}\subseteq N$ containing one point from each super junction, i.e. $v_i\in N_i$, and different requirements have to be fulfilled by the subset of the edges $Q\subseteq E$. The cost of S is the total edge cost $C(Q)=\sum_{i\in Q}c(i)$.

The result of GTSP is the minimal cost for the super junction groups, it corresponds a round trip for all super junctions without any redundancy edges. Huygens (2002) and Hu (2008) extended the GTSP to the biconnected edge of the network. Hu proved it as a NP-hard problem (Hu, 2008, pp. 94).

Generalized Minimum Edge-Biconnected Network Problem (GMEBCNP): For a clustered graph $G=\langle N,E,C\rangle$, we need an edge-biconnected network of minimal cost which connects

one point from each super junction. A feasible solution to the GMEBCNP is a subgraph $S=\langle V,Q\rangle$, which $V=\{v_1,v_2\cdots,v_K\}\subseteq N$ and Q connects all nodes in V via edge redundancy, i.e. for each pair of points $v_i,v_j\in V,i\neq j$, there must exist at least two edge-disjoint paths.

The biconnected graph has no articulation nodes and it only reflects the general character of the nodes, and we can add general topologic constraints for the super junctions, and the constraint is the prior information of the road networks.

Constrained Minimum Perfect Matching Problem (C-MP^2): Given a clustered graph $G=\langle N,E,C\rangle$ and a partner graph $S^{'}=\langle V^{'},Q^{'}\rangle$ which has a topological relationship with $G:V_i^{'}$ is the single corresponding node for $N_i^{'}$, i.e. $V_i^{'}\rightarrow N_i^{'}$ and $Q^{'}$ is the connection for $V^{'}$. We seek a minimal cost subgraph $S=\langle V,Q\rangle$ with $V=\{v_1,v_2\cdots,v_K\}\subseteq N$ containing one point from each super junction, i.e. $v_i\in N_i$, and $Q\subseteq T^{'}$. The total edge cost of S , i.e. $C(Q)=\sum_{i\in Q}c(i)$ is minimized.

Feremans et al. (2002) regarded the application of the C-MP^2 as a generalized minimum perfect matching problem. C-MP^2 is a complex combinatorial optimization problem, and its solution is a certain subset of a finite set. Thus this problem can be solved by completely enumerating the subsets. However, this is only applicable when the dataset has a limited scope (Wolsey, 1998). More intelligent methods are required for large datasets.

To get the solution for GMEBCNP, it is natural to choose a similar approach based on GTSP. Hu (2008) extended such a GTSP-based approach with comprehensive consideration of the neighborhood structures. It is obvious that the local arrangement of the connection can provide only approximated solution. Furthermore, Hu's algorithm is suitable for biconnected graph, and it needs extra work for the general graph, which has discretional connection instead of biconnected connection among the nodes.

Combinatorial Optimization (Cook,1998) is the process of finding one or more best (optimal) solutions in a well defined discrete problem space, i.e. a space containing a finite set of possible solutions, that optimize the so-called objective function (Pop, 2002). Being inspired by the definition of GTSP, we also formulate our problem using the linear model.

To model the C-MP^2 problem, we define the possible selection of the edge as a binary variable x_{ij} and the edge cost c_{ij} , where *i and j* respectively denote the sequence of edge in the partner graph $S^{'}=\langle V^{'},Q^{'}\rangle$ and in set E which is the corresponding set for $T^{'}$, and J_i changes according to the edge $t_i^{'}$. The objective function of C-MP^2 is

$$\min \sum_{i\in T^{'}}^{I}\sum_{j=1}^{J_i}c_{ij}x_{ij} \tag{4-10}$$

subject to:

$$\sum_{j=1}^{J_i} x_{ij} \le 1 \quad \forall i = \{1, 2, \cdots, I\} \tag{4-11}$$

$$\sum_{j=1}^{J_i} z_{kj} = 1 \quad \forall k = \{1, 2, \cdots, K\} \tag{4-12}$$

$$z_{kj} = N_{kj}^{x_{ij}} \quad \forall k = \{1, 2, \cdots, K\} \tag{4-13}$$

$$z_{kj} = \begin{cases} 1 & if\ select\ point\ j\ \ from\ super\ junction\ k \\ 0 & otherwise \end{cases} \tag{4-14}$$

$$x_{ij} = \begin{cases} 1 & if\ select\ the\ edge\ j\ from\ edge\ group\ i \\ 0 & otherwise \end{cases} \tag{4-15}$$

This is a typical mixed 0-1 integer programming with the binary variables x and z to be determined. Eq. 4-9 means only one edge will be selected from each edge group that corresponds to edge $T^{'}$ in the partner graph. Eq. 4-11 ensures that only one point can be selected from each super junction N_i. Eq. 4-14 guarantees the corresponding relationship between node and edge, where z_{kj} is the point j from super junction N_k, whereas $N_{kj}^{x_{ij}}$ indicates the binary value of the end point from edge variable x_{ij} in super junction N_k, these topological constraints will ensure the identical connection from other super junctions.

For a general road network which is usually a very large dataset, The linear programming model has the advantage to get efficient solutions for Eq. 4-8 from software packages, e.g. CPLEX, LINGO, 'linprog' from MATLAB, LPSLOV which can deal with large number of variables. In practice, the road network reveals many characteristics which are different from the general graph. Therefore, we will explore these characteristics in the following sections to build up a novel solution to the sparse matching problem of road network.

4.3.3 Matching cost and local optimization

The edge cost c_{ij} in Eq. 4-10 is so far assumed as a known factor reflecting matching performance between the road segment and the extracted line elements. The similarity between the two candidates depends highly on the transformation parameters that include the displacement v, the rotation θ and the scaling factor s.

$$C_i(M_i, L_{ij}) = w_v \cdot v_i + w_\theta \cdot \theta_i + w_s \cdot s_i \tag{4-16}$$

where M is the road segment; L is its counterpart in extracted line elements; w_v, w_θ, w_s are the cost for the transformation parameters; i means the current super junction, and j is intersection sequence in the super junction.

From Eq. 4-16 we know that the cost for the matched candidates is directly proportional to the parameters. The costs for parameters can be empirically determined. Their sum is equal to 1.

According to our experiences the orientation has a higher reliability than other factors. The broken lines and the uncertain location of the end node of the extracted lines may all influence the calculation of parameters. We may verify our experience with an example.

Since the node pairs e.g. $\left\{I_1, n_1^1\right\}$ and $\left\{I_4, \{n_4^1, n_4^2\}\right\}$ in Fig. 4-15 are known from the current correspondence, the parameters in Eq. 4-16 can be simply derived from the end nodes. However, the measures of displacement v, the rotation θ and the scaling factor s, need to be normalized so that we may determine their costs.

The rotation angle reflects the orientation difference between the matched lines. It can be calculated by moving the matched node, as shown in Fig. 4-16. The rotation angle is counted clockwise, and two different conditions are considered: A) normal matching; B) turnover matching.

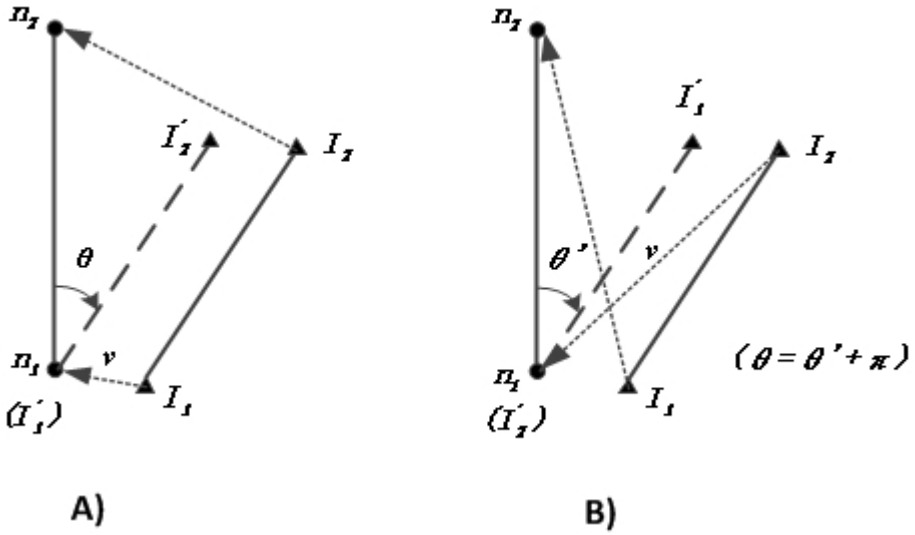


Fig. 4-16 Rotation of the matched lines

Normalization of the rotation is possible for both A) and B). We choose the smaller angle of A) because the angle of B) is associated with a larger cost .

$$A)\ \theta = \begin{cases} \theta / 2\pi & \theta \le \pi/2 \\ (\pi - \theta)/2\pi & \theta \ge \pi/2 \end{cases} \quad B)\ \theta = \begin{cases} (\theta + \pi)/2\pi & \theta \le \pi/2 \\ \theta / 2\pi & \theta \ge \pi/2 \end{cases} \tag{4-17}$$

Similarly, the displacement and the scaling factor can be normalized using Eq. 4-18 and Eq. 4-19 respectively. The displacement is normalized using a threshold of the possible matching area. The larger distance influences the cost function significantly, but it set the upper limit as 1 for the parameter v. The scaling factor is related with the length of the matched lines. We assume that the road segment is usually larger than the broken lines from the image. Similar lengths are associated with low cost.

$$v = \begin{cases} \dfrac{v}{T_v} & v < T_v \\ 1 & v \ge T_v \end{cases} \tag{4-18}$$

where v is Euclidean distance between the nodes as shown in Fig. 4-16; T_v is the empirical threshold.

$$s = \begin{cases} 1 - \dfrac{L_I}{L_n} & L_I < L_n \\ 1 & L_I \geq L_n \end{cases} \tag{4-19}$$

where s is the scale factor for matched lines; L_I is a line element from the image; L_n is the road segment.

The divide-and-conquer strategy is efficient for extensive tasks in GIS community, and it is also suitable to solve the multiple correspondence problem. To be more specific, the isolated multiple correspondences can be optimized separately from the entire dataset if it contains the specific pattern i.e. isolated super junction in the network.

Proposition 4-1 Given a *k*-correspondence node N_i with m neighbors, the minimal cost of the node depends only on the unique correspondence of each neighbor.

Proof. Assume the node N_j which is one of N_i's neighbors has a multiple correspondences n, then there are $k \cdot n$ virtual connections between them. To determine the right candidate from N_i, every correspondence of N_j should be considered, which in turn is connected to its own neighbors. Therefore, the optimization task propagates further and the optimal choice can't be easily decided. If N_j has a single correspondence, the process is much simpler.

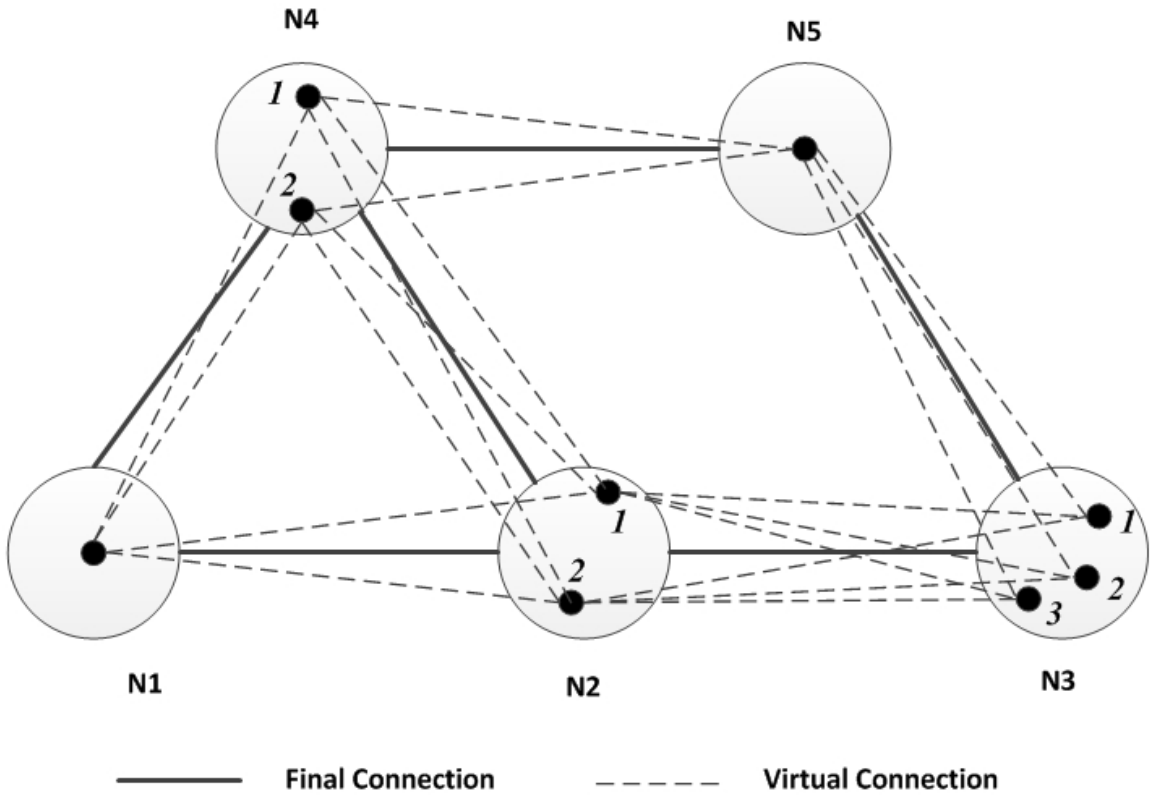


Fig. 4-17 Multiple correspondences of the individual nodes

In Fig. 4-17 N_3 and N_4 contain multiple isolated nodes which can be optimized individually. The solid lines indicate the topological connection between the nodes, whereas

the dash line is the virtual connection between them. The optimal location of the intersection has the minimal cost among the virtual connections and it can be derived from its multiple neighbors. Node N_3 is related with super junctions $\{N_2, N_5\}$, which have simple structure with only one point inside, so the minimal cost is obvious with simple comparison of the edge cost. Likewise, node N_4 is related with super junctions $\{N_1, N_2, N_5\}$, which has the same character as super junctions N_3. In this example, it is the simple case although there are two super junctions with more intersections.

Assume there are *P* possible connections in super junction N_k, the total cost of n_{kj} will be

$$\bar{C}'_{kj} = \sum_{p=1}^{P} c_{kj}^{p} \tag{4-20}$$

where c_{kj}^{p} is the edge cost for the connection *p* to the node n_{kj}.

The transition from the edge cost to node cost in Eq. 4-20 provides an efficient solution of the matching approach for a simple structure. The node selection is straightforward.

$$\min\left\{\bar{C}'_{kj}\right\} \ \forall j = 1, 2, \cdots r_k \tag{4-21}$$

where r_k is the number of candidates in super junction N_k.

It avoids the complex computation in Eq. 4-10. Unfortunately, it is not common in practice because the real road network has complicated structures. Therefore, it is necessary to traverse all individual nodes to get the optimization solution in the entire network.

4.3.4 Multiple correspondences optimization for super junctions

Eq. 4-10 shows the solution to filter out the multiple correspondences by means of the overall minimal connection cost for general networks with large super junctions, but the method is inefficient for the optimization task with small or medium-sized networks which, for instance, have less than 100 nodes. It is therefore necessary to seek an inefficient solution for small or medium-sized networks. We know that the divide-and-conquer strategy is efficient for larger conflation framework (Wu et al., 2007) and many other GIS-based applications. Being inspired by the strategy, we will introduce a novel solution based on Eq. 4-10 but for a fixed number of super junctions N, which is dedicated for our sparse matching problem.

The super junctions are the generalized form of the simple node-edge graph, which have been reported in (Gallo and Pallottino, 1988). Our new idea is rooted on the classical approaches such as the shortest path, GTSP etc. The all-pairs shortest path approaches like Dijskstra (Dijkstra, 1959) algorithm and Floyd-Warshall algorithm (Floyd, 1962) are well known for their capability of computing the minimal cost between two nodes. But the node in their context is limited to only one point. In our context of multiple correspondences, there are many possible points around the same super junction and we cannot get the minimal cost for the edges even though we have the shortest path for each node pair. Moreover, the strategy of determining the optimal cost from the shortest paths is still problematic. GTSP is another

efficient approach and it can compute a tour with the minimal cost which is the optimal Hamilton circle from the super junction.

GTSP visits each node only once with a minimal cost from all the possible combinational tours. GTSP differs from C-MP^2 in two aspects: 1) the GTSP guarantees the minimal cost in its solution, but it is not always the optimal solution if it connects the branch between the super junctions, because the principle of GTSP is to find the minimal cost with the requirement to visit each node only once, whereas the C-MP^2 ensures the minimal cost for all given connections in the road network; 2) the topological connection is a constraint for the GTSP. It restricts the number of the combinational solutions which can be deduced from the symmetry or asymmetry adjacent matrix. In case of C-MP^2, there are many no connections (∞) between the nodes. Table 4-3 shows the adjacent matrix from Fig. 4-15. The ratio between the number of connected nodes (*1*) and no connected nodes (∞) is around *1:1*, which means that it has only 50% of all combinations from the GTSP approach.

Table 4-3 Adjacent matrix for C-MP2

		N_1	N_2		N_3			N_4		N_5
		N_1^1	N_2^1	N_2^2	N_3^1	N_3^2	N_3^3	N_4^1	N_4^2	N_5^1
N_1	N_1^1	∞	1	1	∞	∞	∞	1	1	∞
N_2	N_2^1	1	∞	∞	1	1	1	1	1	∞
	N_2^2	1	∞	∞	1	1	1	1	1	∞
N_3	N_3^1	∞	1	1	∞	∞	∞	∞	∞	1
	N_3^2	∞	1	1	∞	∞	∞	∞	∞	1
	N_3^3	∞	1	1	∞	∞	∞	∞	∞	1
N_4	N_4^1	1	1	1	∞	∞	∞	∞	∞	1
	N_4^2	1	1	1	∞	∞	∞	∞	∞	1
N_5	N_5^1	∞	∞	∞	1	1	1	1	1	∞

The C-MP^2 is based on the topological constraints and the multiple correspondences in each super junction, especially, when each connection i.e. the topological relationship between two super junctions is known from the road network, it provides the important lower and upper bounds for the optimization problem in Eq. 4-10. The lower bound is sum of the edges with minimal cost for each connection, i.e.

$$\sum_{i \in T'}^{I} \min c_{ij} \quad \forall j \in (1, J) \tag{4-22}$$

Likewise, the upper bound can be computed. The optimal cost lies between the lower and upper bound. Bearing in mind the points with minimal cost in the cluster (Eq. 4-21), our idea is to rebuild the topological relationship by selecting the edges which have an overall minimal

cost for the matching problem. At the same time, the related node cost can be minimized locally:

$$\begin{cases} \sum_{i \in T}^{I} \min c_{ij} & \forall j \in (1, J) \\ \min\left\{\bar{C}_{kj}^{'}\right\} & \forall k = 1, 2, \cdots K \end{cases} \tag{4-23}$$

Eq. 4-23 reflects a contradictory proposition because it is not possible to obtain correct topological relationships for the network by selecting the edge with minimal cost for each connection and because the point with the minimal cost to the neighboring node of the super junction worsens the situation. We combine the two constraints in our algorithm and utilize the greedy concept to compute the optimal solution.

Different from the global optimal approach in Eq. 4-10, the new greedy algorithm computes the local optimization for the nodes and connections and then propagates the result till a globally minimal solution is determined. The process is controlled by the derived lower and upper bound of the cost. Since the super junction groups are assumed in small or medium size, the greedy strategy can get suitable result for C-MP^2.

The mandatory selection of the edge with minimum cost contains the fragmental lines at certain super junctions which need to be fixed and keep the cost at a reasonable level. Due to the sparse connection in the adjacent matrix as well the many-to-one correspondence between the super junctions and the intersections of the road network, we name this greedy algorithm as Sparse Matching Algorithm (SMA), which is summarized in Algorithm 4-3.

Algorithm 4-3: SMA

Input: 1) Super junction matrix *N* and edge matrix *E*;

2) Cost matrix *C*;

Output: 1) The matrix of optimal points *P*, $P \in N$;

Steps:

1) Derive the edge with minimum cost for each connection among the super junctions, and the edge group $E_{\min}$ is the basis to optimize the overall cost;

2) Get the sum of the costs for each point connected with *E* in the super junctions and determine the point with the summed minimum cost;

3) Compute the lower bound $T_{\min}$ and upper bound $T_{\max}$;

4) Check if there are any unconnected edges at super junctions, if yes, go on with the next step, otherwise terminate the algorithm if all the nodes are fixed,

5) If it is the first time to visit the network, get the node with maximal valence p_m and build a tree structure which is rooted on p_m . It requires all the neighboring super junctions to be added to the tree; otherwise get the groups (leaves) at the current level of the tree;

6) For each connection at the current level, check if there is a direct connection in $E_{\min}$ between the parent node and current node;

7) If there is no direct connection, select the point from current node by minimizing the cost to the already fixed nodes;

8) Mark the super junction as being fixed in N ;

9) Visit all the leaves at the current tree level;

10) Search the neighbors for the next level and add them to the tree;

11) Go to Step 4 if not all the super junctions have been traversed; otherwise, return the result and terminate the procedure.

In SMA, Step 4 ~10 is the iterative procedure from the specific node and extends to the whole network. In general, the super junctions with fragmental edges are difficult to predict. It is the main factor which influences the computing difficulty of SMA. In the ideal situation, the edge group $E_{\min}$ with minimal cost covers exactly the point in the super junction which has no fragmental lines; whereas in the worst situation each super junction needs to be fixed.

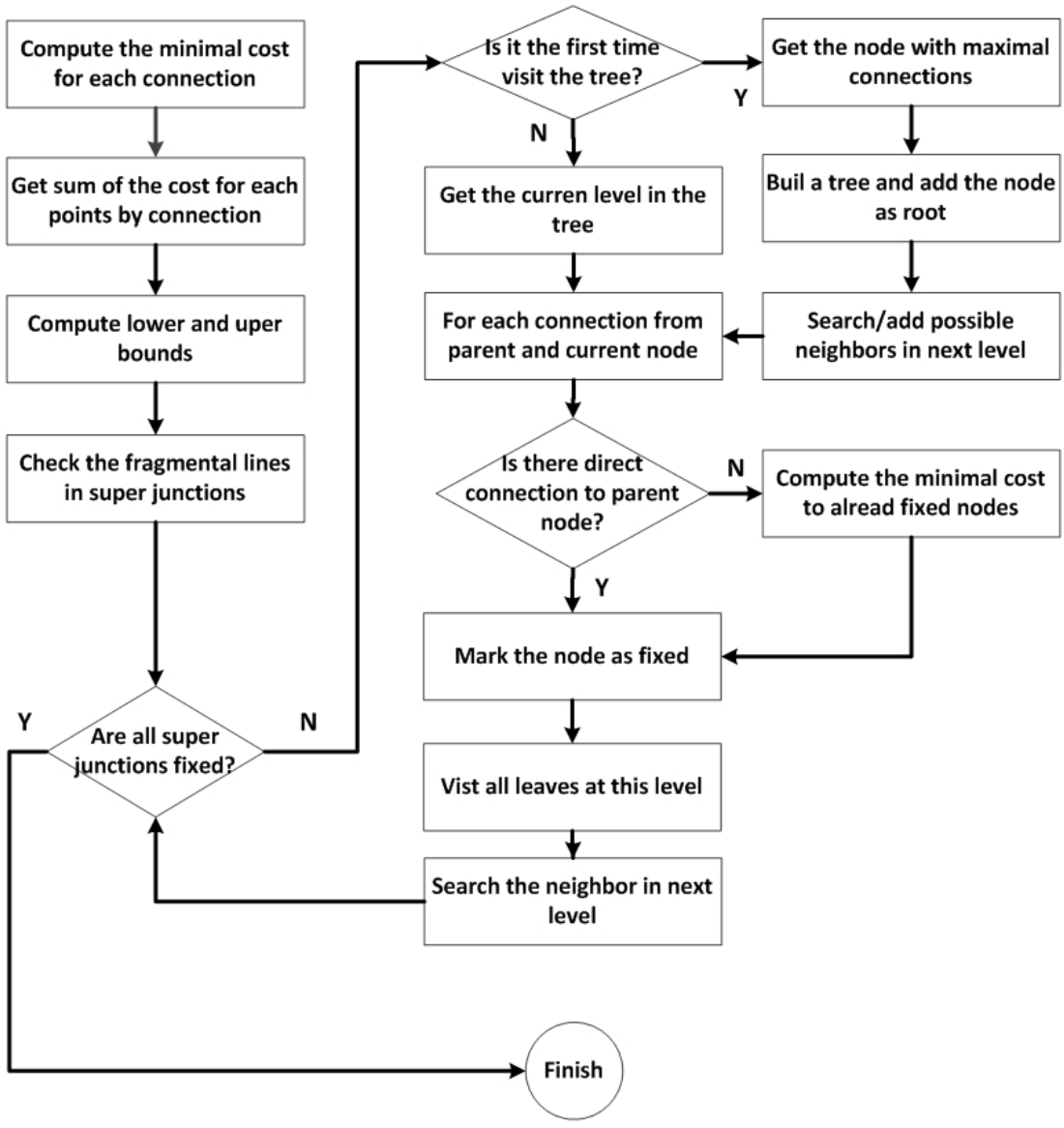


Fig. 4-18 Data flow in SMA

In order to accelerate the searching steps, the tree data structure has been used to track the propagation throughout the entire network: new leaf in the tree data structure was added when the searching process spread from one node to its neighbors. The repartition at the super junction, where more than one point of the super junction have been selected in previous step, is based on the minimal connection cost to the already visited nodes, and the point with minimal cost to all neighbors is chosen as the right candidate. If the point has been decided, it was marked as visited in the super junction matrix N . The operation continues until all nodes in this network have been traversed.

Fig. 4-18 shows the data flow of the algorithm, which has been described in Algorithm 4-3. The local optimal solution is not necessarily the global optimization, but the upper and lower bounds are determined in our problem. Moreover, the point with minimal cost in the super junctions is taken as the reference candidate in the propagation from the first node to the other nodes in the network. The performance of the SMA is discussed in the following section.

4.3.5 Efficiency

To evaluate the efficiency of SMA, we use a group of synthetic datasets. The only difference from the real dataset is the randomly generated costs for the connections between super junctions. The costs are delimited in the range of $[0,1]$ in the test. We choose different sizes of super junctions - 6,10,20,30,40,50,60,70,80 and 90 to verify time efficiency of SMA in the computation of minimal cost.

Table 4-4 Computing times of SMA for synthetic datasets

ID	Number of Nodes	Number of Points	Number of Edges	Lower Bound	Upper Bound	Cost returned by SMA	Time (Sec)
1	6	14	9	1.34	3.76	3.43	0.013
2	10	23	17	3.37	6.76	6.18	0.013
3	20	40	43	10.2	18.9	18.5	0.016
4	30	60	68	15.7	28.0	28.0	0.022
5	40	79	92	22.8	36.8	36.5	0.029
6	50	100	114	26.0	47.3	46.6	0.034
7	60	120	141	33.7	62.6	61.6	0.059
8	70	142	166	38.1	68.7	66.9	0.059
9	80	166	192	45.2	82.3	78.8	0.074
10	90	190	221	44.6	88.9	88.1	0.112

Since SMA is a greedy base algorithm, it cannot guarantee that its optimal cost is the global minimal cost in the dataset, but it delimits the optimal cost within a certain range between the lower and upper bound. Table 4-4 shows the detailed computing results of the

test dataset. The number of points and edges are the main factors which influence the computing time.

Fig. 4-19 shows the optimal result from SMA with respect to the minimal cost of the edges. 4 super junctions with fragmental edge connections (66%) need to be fixed in SMA.

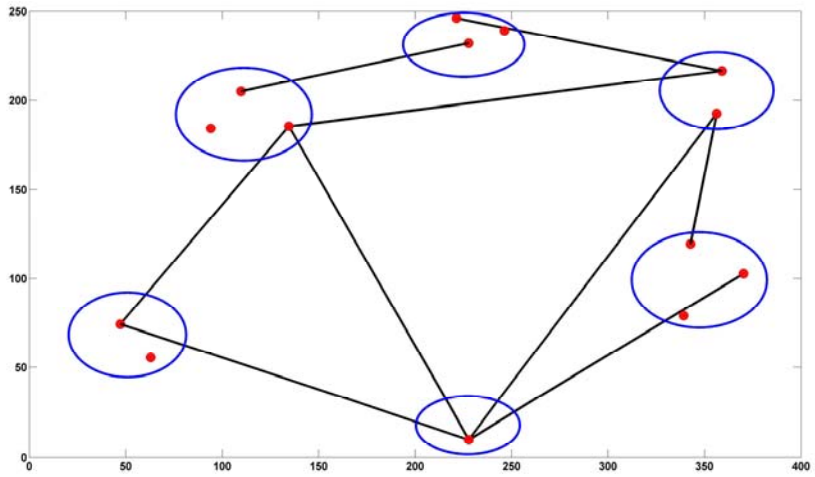


Fig. 4-19 SMA result compared to minimal edges with 6 super junctions

The multiple selections of intersection from super junctions became worse with large network with more super junctions. In Fig. 4-20, there are 30 super junctions in the network, and in some areas, the neighboring super junctions are all broken. Fig. 4-21 shows the reconstructed network where only one intersection has been selected in each super junction.

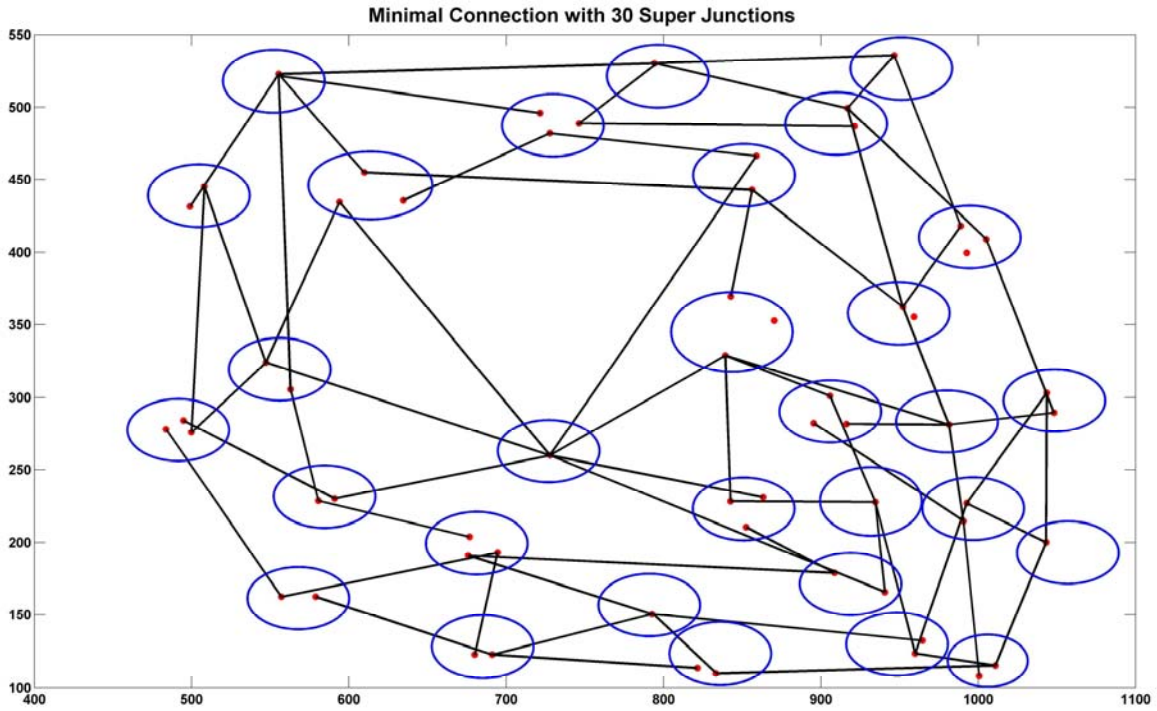


Fig. 4-20 The minimal connections for the synthetic network

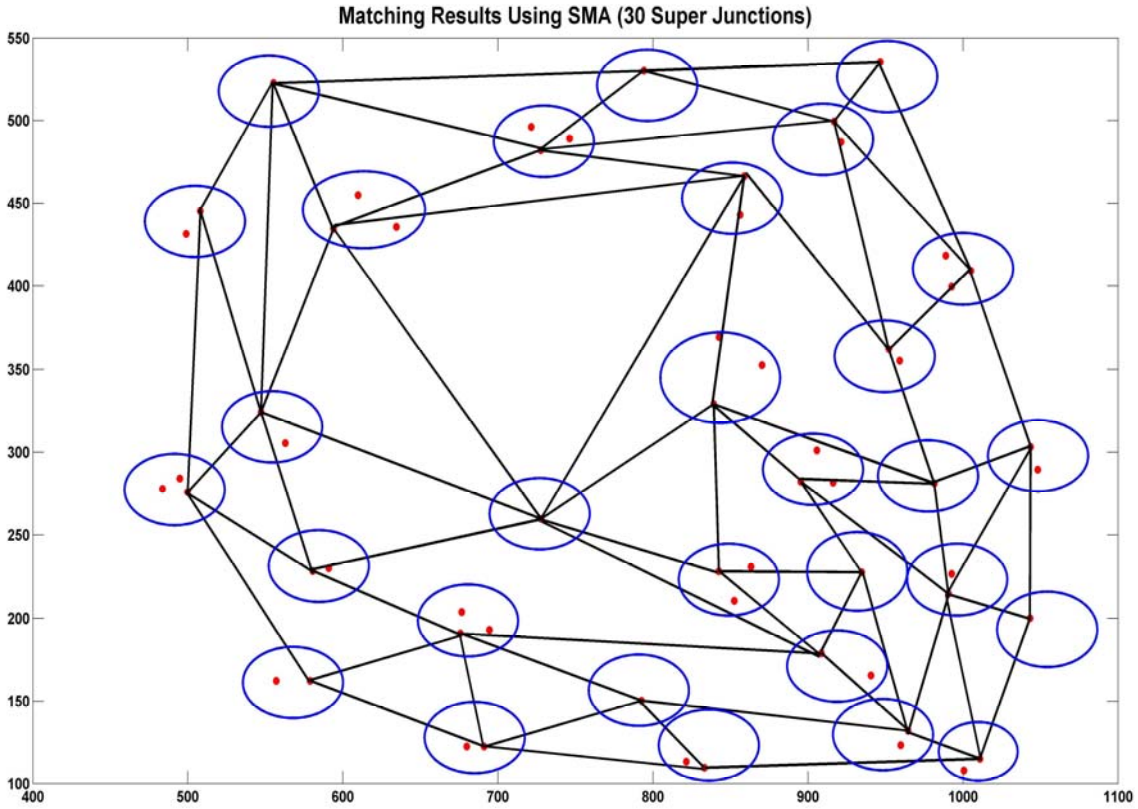


Fig. 4-21 The reconstructed network using SMA

Comparing to the minimal costs by simply selecting the edge or point with minimum cost for each connection, the optimal cost returned by SMA is a compromise. It indicates that the selection of the points from super junctions by the minimal cost of either edges or points is not the possible solution for the problem. The SMA can overcome the limitation and deliver a better solution. Fig. 4-22 shows the comparison of the three different costs for each group in our test. The optimal solution from SMA is close to the connection of the minimal cost of each super junction, which indicates a better solution if a quick computing is required.

The time complexity of SMA relies mainly on the initializing steps (Step *1~3*) and the greedy searching process (Step *4-10*). Assuming we have m edges, k points and n super junctions, the minimal edge computation in Step *1* runs in $O(km)$, and the point with minimal cost in Step *2* is the same as Step *1*, whereas Step 3 has the complexity $O(m)$. In the process from Step *4-10*, the determination of the minimal cost of points for super junctions with broken connections in Step 6-7 needs $O(n^2)$ for the worst situation, the overall complexity for the loop in Step 4-10 is therefore $O(n^2 \log n)$, which is also the complexity of SMA. For the synthetic datasets, the actual computation time is shown in Fig. 4-21 which confirms the efficiency of SMA for smaller sizes of super junctions.

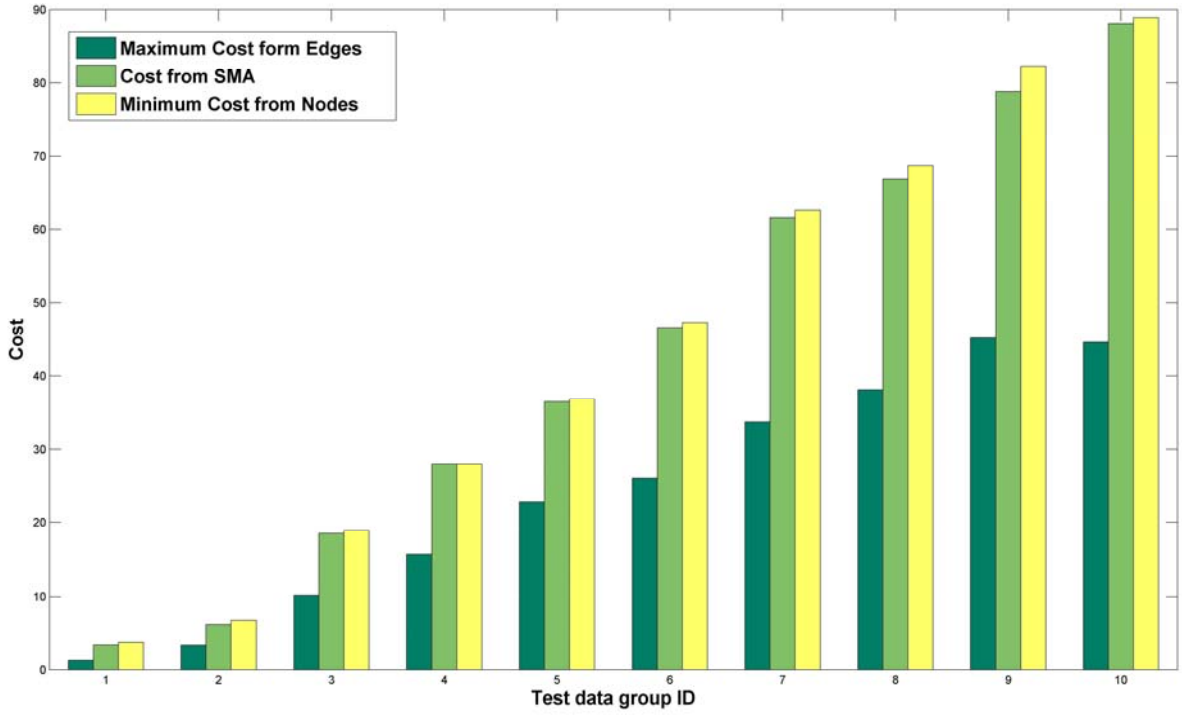


Fig. 4-22 Optimal cost from SMA comparing to minimum cost from edges and nodes

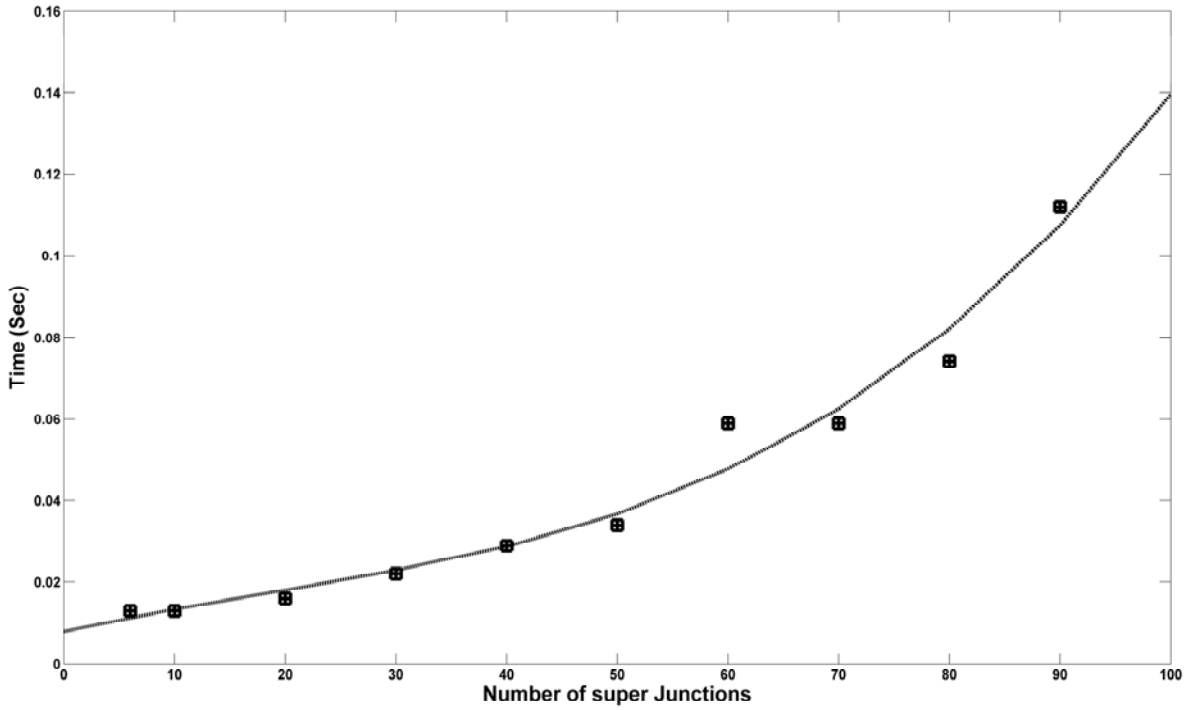


Fig. 4-23 Computing time for super junction groups

In practice, the time complexity for a small number of super junctions is rather low as shown in Fig. 4-23. The time cost in the last experiment from the synthetic datasets where the number of points in the super junctions rises up to 190 and the number of connections up to 200 is till kept below a second. Since the time complexity of GTSP has an exponential nature, it runs usually in several seconds to determine the route with minimal cost (http://www.tsp.gatech.edu/concorde/benchmarks, accessed 04-2012). The greedy idea in

SMA to explore the topologic information from the connections is not the only solution to the problem, it is therefore reasonable to choose a heuristic method to solve C-MP^2.

Besides the synthetic test datasets, we will also test the approach using the real datasets derived from the road network and the image in Chapter 5.

Chapter 5

Implementation and Evaluation of the Conflation Approach

5.1 Spatial Data Initialization

As discussed in previous chapters, the proposed congruent hybrid model for heterogeneous spatial datasets contains three interrelated components: 1) feature detection and grouping by ECM and GA respectively; 2) multiple correspondence matching by SMA; and 3) feature transformation. The feasibility and efficiency of the first two components have been proved with synthetic datasets. In this chapter, we verify and evaluate the entire model step by step with test datasets from the real world.

Fig. 5-1 Locations of the test regions

The environment of the experiment is based on the Matlab (R2009, 32bit) on a normal PC (personal computer) with the CPU of Intel® Core® 2 Duo CPU 1.99GHz and 3.0 GB physical memory. We choose the test data in rural and suburban area with a road density around 0.5 ~1.5 KM/KM^2 near Munich, Germany. The vector road network is an enriched version of road network by Zhang (2009) and contains both topological and semantic information which is useful for feature-to-feature matching approach, but we only focus on the topological characters, i.e. nodes and segments, in the congruent hybrid model. In aerial orthoimage has a resolution of 0.4 meter (Clerico et al., 2009) and a geo-referenced spatial coordinate system. As we already know from Chapter 3, the roads in Very High Resolution image appear as area

objects other than simple linear features, they have no semantic information, but contain certain noises from other man-made objects, e.g. cars and zebra lines.

1)

2)

3)

4)

5)

6)

Fig. 5-2 Different orthoimages of the test areas

It should be pointed out that there are non-systematic geometric displacements between the road network and the image. In the urban areas in our datasets where the correction of the misalignment has been done during the geo-referencing process, the displacement is less serious than in suburban and rural area where more Ground Control Points (GCP) are needed

to correct these errors, but it is expensive to build up the dense GCP in rural area. Manual work is tedious and time consuming. In our congruent hybrid model, the correspondences between the image and road network should serve the same function as the GCP in the geo-referencing process.

The linear feature extraction is the first and inevitable component in the congruent hybrid model. The extraction ratio is usually very low due to the image complexity and the ratio of roads covered by nearby features. We also compare the different transformation functions for the matched features to order to identify the most efficient extraction solution for the geometric features.

Table 5-1 Extensions of test regions

Area ID	Appr. Total road length (KM)	Appr. Road Density (KM/100KM^2)	Location	
			Upper Left	Unter Right
1	2.3	67	$11^{\circ}39'03.18" E$	$11^{\circ}39'30.34" E$
			$48^{\circ}09'01.28" N$	$48^{\circ}08'03.18" N$
2	1.3	50	$11^{\circ}39'11.28" E$	$11^{\circ}39'22.61" E$
			$48^{\circ}10'54.67" N$	$48^{\circ}10'47.11" N$
3	0.5	107	$11^{\circ}39'07.6" E$	$11^{\circ}39'20.96" E$
			$48^{\circ}07'22.95" N$	$48^{\circ}7'15.56" N$
4	11.4	129	$11^{\circ}31'26.24" E$	$11^{\circ}31'06.47" E$
			$48^{\circ}10'02.04" N$	$48^{\circ}10'01.38" N$
5	2.7	85	$11^{\circ}35'07.19" E$	$11^{\circ}35'19.89" E$
			$48^{\circ}09'00.25" N$	$48^{\circ}08'44.01" N$
6	1.2	84	$11^{\circ}06'25.76" E$	$11^{\circ}38'21.31" E$
			$48^{\circ}06'25.76" N$	$48^{\circ}06'14.57" N$

Following the “divide-and-conquer” strategy, the large orthoimage was divided into small tiles in order to preserve the computing efficiency (Wu et al., 2007). We also generated the ground truth for test dataset for the evaluation purpose. The selection of the experimental area is guided by our intension of testing the proposed algorithm for various scenarios, e.g. rural as well as suburban areas, although we can include only limited test regions in this dissertation. In test area 1, roads and buildings have similar radiometric characters, it is therefore

impossible to discriminate them from each other if we only consider the radiometric information. Test area 2 and 3 are suburban regions which have curved or straight road segments. With test area 4 and 5, we want to demonstrate the algorithm with pedestrian roads which usually include high noise ratio from the nearby trees. Test area 6 is a typical intersection in a suburban region. Its roads reveal different radiometric characters because of the different materials of road surfaces, indicating amendments at different times.

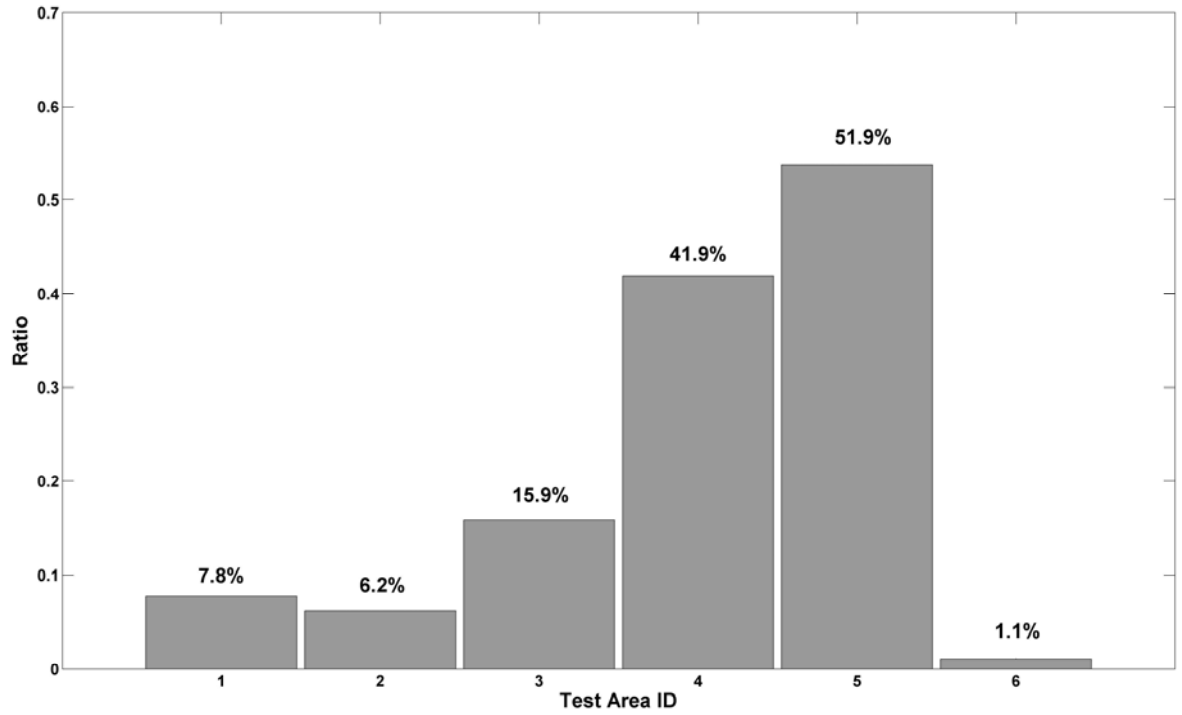


Fig. 5-3 Ratio of the invisible roads in selected regions

The noise in the images varies from image to image, and we also manually approximate the covered road comparing to the road ground truth as shown in Fig. 5-3. The maximal noise ratio in our test is up to 50% in the test area 5, although higher noise ratio is also possible in the images.

5.2 Evaluation Methodology

A series of measures for each component is adopted and combined in the tests to acquire the overall efficiency for the congruent hybrid model. For the feature extraction and grouping, two favorite quality measures have been used in many road extraction approaches (Hinz and Baumgartner, 2003). One is the completeness which is the ratio between the extracted road segment and the reference ground truth as expressed in Eq. 5-1. The other is the correctness which is the ratio between the length of correctly extracted road features and the total length of extracted features as expressed in Eq. 5-2.

$$\mathrm{K}_{completeness} = \frac{Length\ of\ \ extracted\ road\ \ features}{Length\ of\ \ ground\ truth} \tag{5-1}$$

$$\Lambda_{correctness} = \frac{Length\ of\ \ correctly\ extracted\ road\ \ features}{Total\ \ length\ of\ \ extracted\ \ features} \tag{5-2}$$

The completeness and the correctness indicate the efficiency of ECM detector with which linear segments are extracted from the edge maps. Another measure is needed to indicate the efficiency of the grouping method with which the neighboring linear features are chained by means of their geometric and radiometric characters into longer objects. The grouping procedure aims to reduce the number of extracted lines, therefore, its efficiency can be reflected by the ratio between the number of lines after and before grouping as expressed in Eq. 5-3.

$$R = \frac{Number\ of\ \ grouped\ \ features}{Number\ of\ \ extracted\ \ features} \tag{5-3}$$

As the key component in the congruent hybrid model, the efficiency of SMA influences the overall performance of the model. The sparseness expressed in Eq. 5-4 reflects the unbalance between the two datasets to be matched:

$$S_p = \frac{Number\ of\ \ grouped\ \ features}{Number\ of\ \ road\ vectors\ in\ the\ database} \tag{5-4}$$

SMA may create various correspondences between the two data sources: 1) correct matching (Eq. 5-5); 2) false matching (Eq. 5-6); and 3) no matching. The third type of correspondence can be derived from the first and second case, because the road network has high quality in our test.

$$T = \frac{Number\ of\ \ correctly\ matched\ road\ vectors}{Total\ number\ of\ \ road\ vectors} \tag{5-5}$$

$$F = \frac{Number\ of\ \ wrongly\ matched\ road\ vectors}{Total\ number\ of\ \ road\ vectors} \tag{5-6}$$

The final measure is the distance between the transformed feature and the ground truth. We choose the Root-Mean-Square Error (*RMSE*) as the distance measure. The smaller this distance, the better the performance of the proposed model. Although the SMA creates the matching results, the residual displacement is still influenced by the transformation function. The *RMSE* from different transformation functions can be compared. In the calculation of *RMSE* as expressed in Eq.5-7, not only the control points, but also the shape points which are usually located inside the segments, for instance, the curve road, are considered to estimate the overall displacements in the results.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_i^{'})^2 + (y_i - y_i^{'})^2}{N}} \tag{5-7}$$

herein N is the number of points whose locations after transformation deviate from the ground truth.

As an integrated model, it is important to know the propagation of the error in one component to the next one. In the experiments, we demonstrate the reliability of the model in different situations, especially the robustness of the proposed model even for lower extraction ratio. The correct matching of SMA is also the focus of the experiments.

5.3 Performance of Feature Extraction and SMA

5.3.1 Extraction of linear features

The extraction of features, especially man-made features, from image with high extraction ratio is a challenging task that has engaged many researchers in the image processing community because man-made objects such as roads and buildings are often influenced by other man-made or natural objects in their vicinity. Many currently available detection algorithms have rather limited applications due to their low recognition ratio, and the extracted features are usually accompanied by some uncertainty. In our conflation approach, the prior information in the existing road databases is used to improve the reliability of the extracted feature. In the following section, we will discuss the efficiency of our linear feature detector as well as the grouping approach based on the classical edge operator.

As already discussed in Section 3.2, the proposed ECM detector and the GA-based grouping method is developed to detect linear features from VHR orthoimage. Similar to the classical road detection approach, our proposed method also includes a series of parameters, and some of them depend on the complexity of the image itself. Unlike the traditional approach of empirically optimizing the parameters, we set up the general parameters in our experiments to demonstrate the robustness of our approach.

Among many edge detectors with various advantages as introduced in (Martin, et al., 2004) we chose Canny operator to extract edges from image because of its generality. In the tests, we slightly change the standard deviation σ of the Gaussian filter and the threshold T_c for Canny operator as shown in Table 5-2 because the default parameters of Canny operator in Matlab are suitable only for general images to generate even tiny edges which are not important for the whole road network. To reduce the influence from the short edges, an edge map was generated by linking the individual edges from the edge operator with the pixel threshold $\lambda = 10$. The Canny operator generates the general edges for all kinds of objects in the image, and it cannot discriminate objects by their shape. In the VHR image, roads are area objects instead of simple lines, and the homogeneous regions encompassed by the linked edges are the possible area of the road. The ECM is a detector of centrelines. It can evaluate the vacant area between two edges from the edge map. The road width can be assumed with help of the prior knowledge about lanes. A threshold of $T_r = [3, 60]$ meters in the tests is chosen.

In Table 5-2, the number of the estimated line elements has been dramatically reduced in comparison to the initial number of edges from the Canny operator, but they are still too many as compared with the number of segments in the ground truth. Moreover, the extracted lines contain edges from objects other than roads. In many test areas, there is a tendency of over extraction, i.e. the ratio between the extracted length comparing to the length of the ground truth is larger than 1.0. The extracted lines from ECM are composed of two components: the initial extracted line elements as well as their extensions.

Table 5-2 The results of linear feature extraction from test areas

Area ID	Parameters for Canny		Number of Canny Edges	Number of Line Elements	Number of segments in ground truth	Extracted total length (KM)	Ratio of over extraction
	Sigma	Threshold					
1	2.0	[0.01,0.15]	3,885	548 (14.1%)	25	3.9	1.70
2	2.5	[0.02,0.35]	381	94 (22.1%)	13	2.28	1.75
3	3.0	[0.10,0.20]	426	38 (8.9%)	14	0.78	1.56
4	2.0	[0.01,0.15]	10,921	1,130 (10.3%)	217	13.1	1.15
5	2.5	[0.01,0.20]	888	344 (10.3%)	28	2.6	0.97
6	2.0	[0.01,0.20]	744	143 (19.2%)	15	2.3	1.9

The initially extracted line elements are the centrelines from the homogenous areas between two edges which are connected by the circular mask. The ECM has the simplicity of being dependent only on the geometry of the extracted edges, so the feature with linear edges is covered by the mask; furthermore, the centre of the mask is connected with its neighbors along the edges. The ECM is suitable for detecting the centrelines of linear objects, but it also suffers from the problem with fragmental edges. Its estimated centreline is separated from the reference edges. Another challenge is related with the complexity of objects. Large road intersection areas, for instance, may not be detected by the ECM because of their larger width than normal roads. Moreover, the isolated facilities around or in the road intersection may be detected as linear objects. The estimated centreline is quite simple as compared with the original edges because a large number of short edges don't contain any linear regions. Fig. 5-4 and 5-5 show the circular masks as well the Canny edges overlapped with the image, and the color of the circular mask is the ratio between the radius of the mask and the maximal radius in the study area.

Fig. 5-4 Circular masks detected by ECM in the test area 2

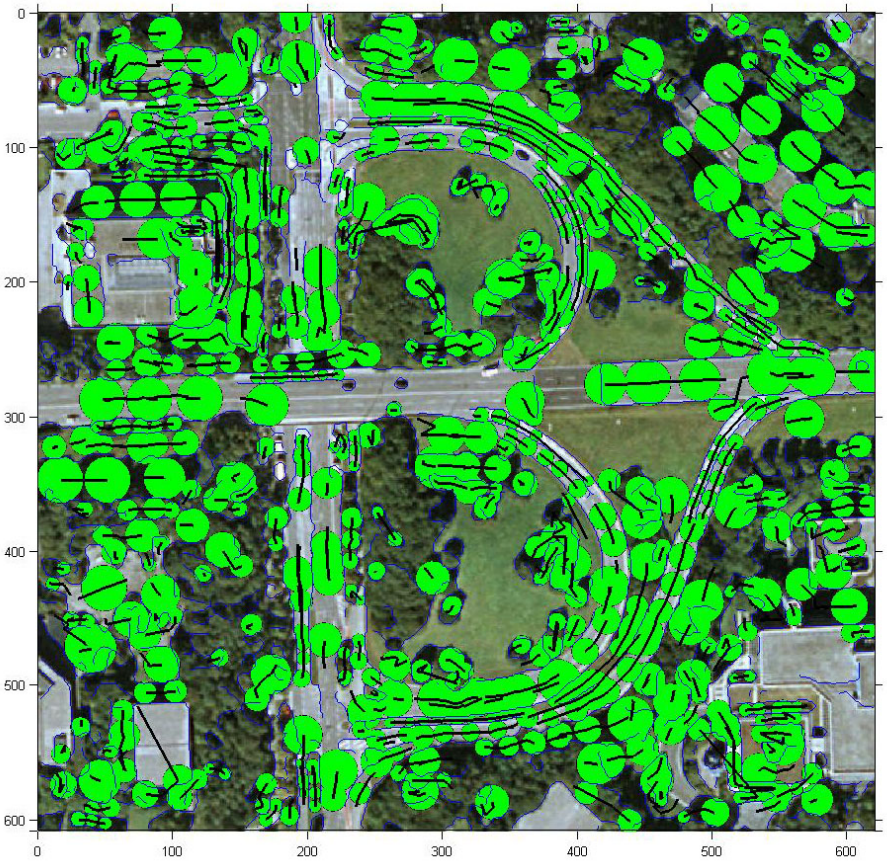


Fig. 5-5 Circular masks detected by ECM in the test area 6

The broken centrelines are formed due to broken extracted edges. This problem may be

alleviated by considering the radiometric intensity on both sides of the estimated centrelines. Following the direction of the curvature, a broken centreline can be extended till it meets other edges or touches the border of the test area. The spatial extension not only covers more features, but also brings together some broken centrelines. However, the extension is not always stable because the two criteria of terminating the extension are not suitable for all the conditions in the edge maps. In spite of this drawback, the extension works well in our tests. Fig. 5-6 shows the extended regions as well as the estimated centrelines. The color tones indicate the average intensities of these regions.

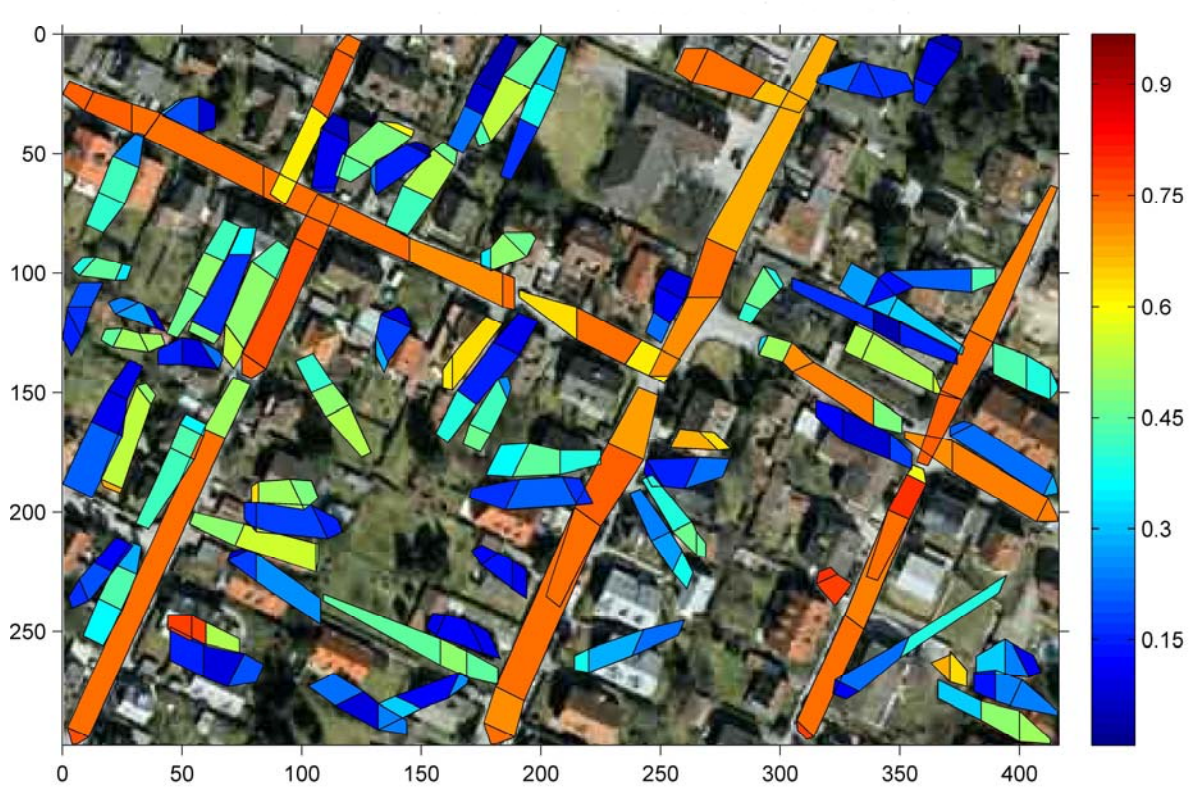


Fig. 5-6 Extracted and extended areas from the centrelines

The centrelines from ECM are still quite fragmental even after the extension on both sides, moreover, they can hardly be used by the SMA because of the high noise ratio. Therefore, it is necessary to reduce the fragmentation and connect the extended centrelines that are separated by tiny noise in the image because the Canny operator is sensitive to any kinds of radiometric difference. GA is a suitable grouping algorithm and has the advantage of generating the optimal results without searching throughout the whole space.

In this work, we use GA to connect extracted lines using a fitness function, which has been discussed in Chapter 3. The result of GA-based grouping approach is shown in Table 5-3 in comparison to the results of ECM and the segments in the reference road dataset.

Table 5-3 Efficiency of GA-based Grouping Approach

Area ID	Number of Edges from ECM	Number of Edges after GA	Ratio (N_{GA} / N_{ECM})	Number of segments in road network	Ratio (N_{GA} / N_{RV})
1	316	185	58.5%	25	7.4
2	94	66	70.2%	13	5.1
3	38	24	63.2%	14	1.7
4	858	686	79.9%	217	3.2
5	124	73	58.9%	27	2.7
6	143	103	72.0%	17	6.1

The GA is effective in reducing the number of centrelines. In our tests, it can reduce around 40 percent of centrelines in number. However, the grouping results are still fragmental compared to the road segments in the road network. One main reason is the existence of non-road features along the centrelines which have very different radiometric values from those defined in the fitness function for GA. Fig. 5-7 shows two types of connections: 1) two separate lines have similar characters as defined in the fitness function. Their connection is visible; 2) two lines are quite near to each other. If they have similar characters as defined in the fitness function, their end nodes are simply merged. Therefore, their connections are no longer visible. Solid blue lines in Fig. 5-7 are the estimated road centrelines, and the solid red lines are the reconstructed connections.

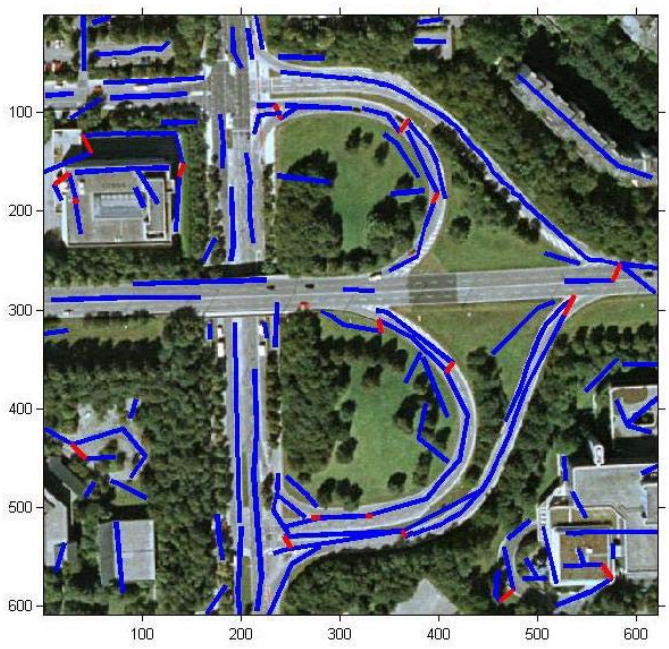


Fig. 5-7 New generated connections by GA

The premature problem of GA algorithm was discussed in Chapter 3. Here we check the fitness values in our test areas as shown in Fig. 5-8. In the test areas, the fitness values become stable after 50 iterations, although there is still some oscillation. It means the convergence in the algorithm.

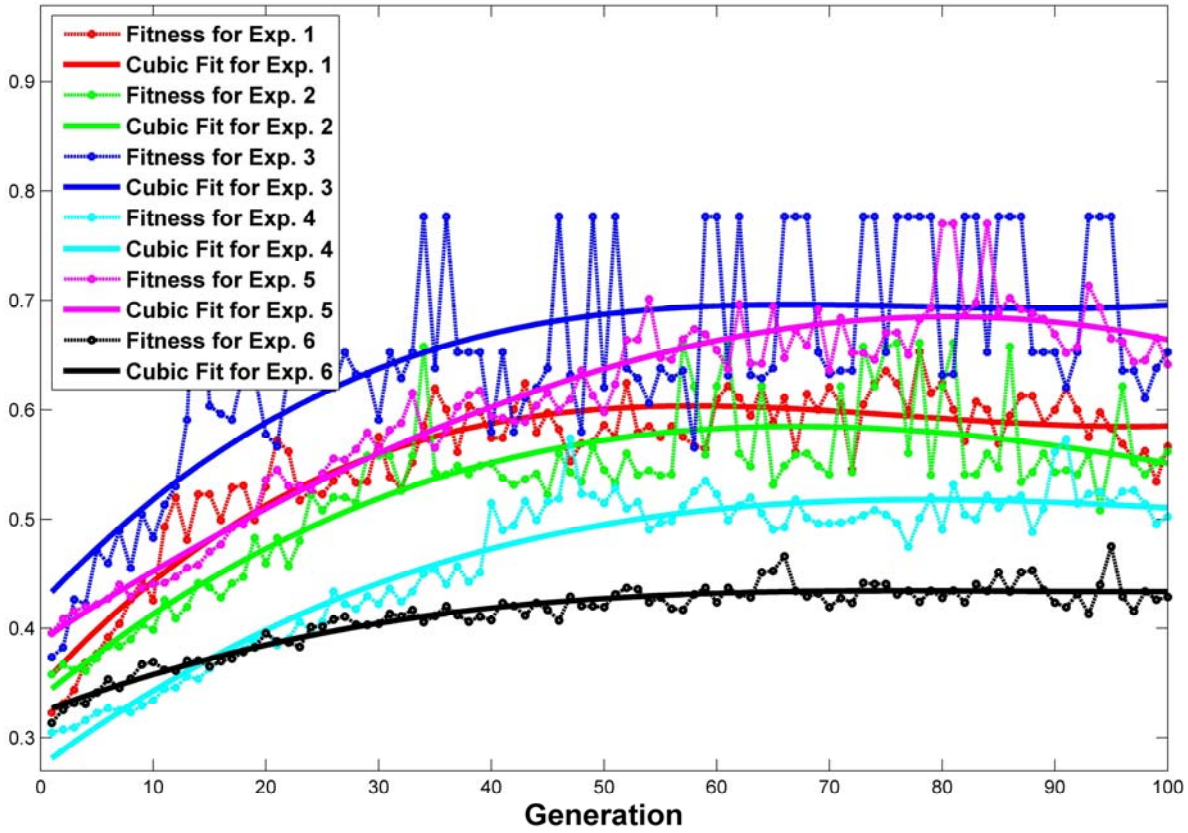


Fig. 5-8 The fitness values of GA algorithm in the test areas

In principle, the grouping of edges only leads to the reduction of the number of edges or longer edges, the completeness and the correctness of the extracted features are not changed. The radiometric character is useful information for the elimination of the non-relevant features. We chose the intensity range [100, 255] to filter out the noise in the extracted features. Table 5-4 shows the correctness and completeness of the centrelines created by ECM and by GA. The correctness after GA could be lower than after ECM if the new connections are interfered by non-road features, but the completeness always becomes larger when new connections within the road area are established. The overall improvement of the correctness and completeness after GA is not so significant in comparison to the grouping effect of centrelines, but GA itself verifies to a certain extent that the extensions of the initial centrelines detected by ECM are in good quality in terms of both geometry and radiometry. With help of GA, the sparse matching ratio can be reasonably increased.

The efficiency of the ECM in terms of completeness and correctness as compared with the grouping results after GA is shown in Table 5-4.

Table 5-4 Comparison of Correctness and Completeness

Area ID	Correctness (Λ)		Completeness (K)	
	Centrelines by ECM	Grouped Lines from GA	Centrelines by ECM	Grouped Lines from GA
1	0.45	0.56	1.68	1.54
2	0.63	0.67	1.69	1.71
3	0.65	0.58	1.57	1.61
4	0.63	0.65	1.14	1.07
5	0.50	0.67	0.97	0.71
6	0.59	0.60	1.93	1.86

5.3.2 Results of sparse matching

The feature matching has the essential task to find correspondence between homologous features. The multiple correspondences *m*-to-*n* in complex situations need to be clarified in different ways using different matching algorithms. In SMA, the complex correspondences are firstly decomposed into individual road cliques with each containing *n* grouped lines, and then the *1-to-n* correspondences are computed for the super junctions to get the exact relationship.

The reference road networks from the navigation database are completely topologized. And the road network can be decomposed into smaller road cliques with each containing a certain number of road segments that are connected with neighboring segments following the good continuation principle of the gestalt. The gestalt decomposition of the road network has the same effect as the GA-based grouping approach for the extracted features, which generates longer lines by connecting the neighbors with sufficient fitness; furthermore, both of them reduce the number of fragmental segments. Using the long road segments, the possibility to get the matched lines from extracted features dramatically increases. The ratio between the decomposed cliques and the existing road segments is

$$R_m = \frac{Number\ of\ decomposed\ groups}{Number\ of\ road\ segments} \tag{5-8}$$

As shown in Algorithm 4-2, finding the preliminary correspondences between the cliques, i.e. the decomposed segments, and the grouped extracted centrelines is straightforward: using a buffer to detect possible counterparts in corresponding datasets. The threshold B_f of the buffer is empirically defined, depending on the offset between the datasets which usually lies in the range of $[0, 200]$ meter. We choose 60m in our experiments because the displacements between both test datasets are rather small. The change of the buffer threshold

does not influence the final quality of the congruent hybrids model, but can cause an increase of the computation time of the SMA. The sparse ratio between the super junctions and the estimated intersection around each super junction is a critical parameter of SMA. Table 5-5 shows its values in the test areas. In test area 4, there are 23 times so many estimated intersections as super junctions in the reference road network. Fig. 5-9 shows an example of the super junctions and their corresponding estimated intersections in our experiment.

Table 5-5 Sparse ratio of SMA

Area ID	Number of super junctions	Number of estimated intersections	Sparse Ratio
1	21	240	11.4
2	14	50	3.57
3	15	30	2.0
4	145	3,337	23.0
5	23	97	4.2
6	14	202	14.4

Fig. 5-9 Super junctions in green within the circular region and their estimated intersections in red in test area 1

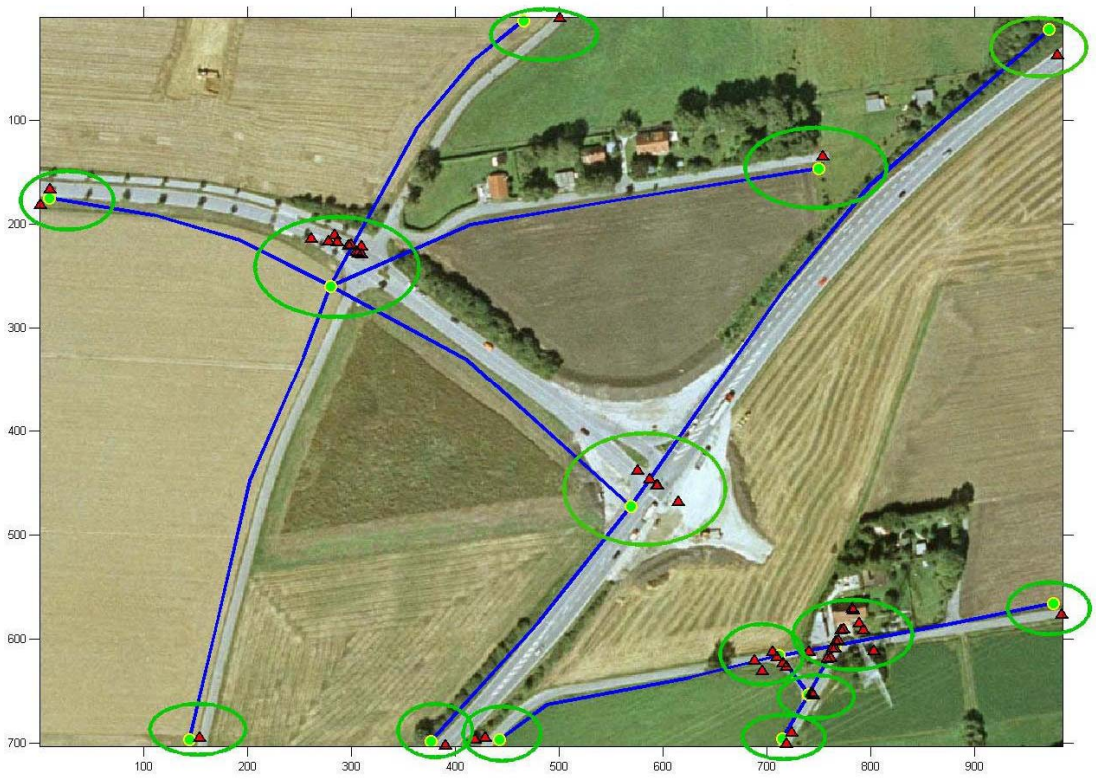


Fig. 5-10 Super junctions and the estimated intersections in test area 2

We have discussed the computing complexity of SMA in Section 4.3 for the synthetic datasets. Now we focus on the matching quality represented by the effectiveness of the matched vector in Eq. 5-5 and false matched vectors in Eq. 5-6 using the real datasets. The false matched vectors include the incorrectly matched vector as well as the unmatched vectors.

One tricky issue in the test is the computation of the matching cost which is formulated in Eq. 4-14 with three parameters: displacement, orientation and scale, but the weight for each parameter is hard to be pre-defined. In the test for synthetic datasets, the weights have been set equal to each other. According to experiences, the orientation has more influence than the other two parameters, but it is not always true for road features in different images. Due to the subtle relationship with the image, we change the weights slightly in different test areas as shown in Table 5-6, where the orientation has been set to be twice so important as the displacement and the scale in most test areas, but the weights in the third test area where the features are almost straight lines are kept the same value for all parameters.

The SMA creates the matched vectors, and the matched nodes in both datasets are connected. Three types of matching situations are included in these vectors. The first two have been discussed in Eq. 5-5 and Eq. 5-6, and the third one deals with the missed matching for the nodes from the reference road network, and can be derived from the first two types. The unmatched nodes influence the final result, but the influence may be reduced to a certain extent by properly selecting parameters for the feature transformation.

Table 5-6 Efficiency of SMA

Area ID	Decomposition Ratio (R_m)	Correct matching rate (T)	False matching rate (F)	Weight		
				w_v	w_θ	w_s
1	44%	89.3%	12.3%	1/4	1/4	½
2	53.8%	100%	0	1/4	1/2	¼
3	50%	92.3%	13.2%	1/3	1/3	1/3
4	45.3%	72.6%	19.5%	1/4	1/2	1/4
5	66.7%	78.3%	16.7%	1/4	1/2	1/4
6	47.1%	83.6%	14.8%	1/4	1/2	¼

As shown in Table 5-6, ca. 50% of reference road segments are regrouped into cliques, depending on the geometry of the reference road networks: the areas containing more straight road segments have a higher regrouping ratio. The SMA performs better in test areas with lower invisible ratios of the road surface.

Fig. 5-11 ~Fig. 5-14 show the road network (in blue), the line elements (in green), the estimated intersections (red cross in red circle) and the final matched results (in pink). Obviously, the estimated intersections became more complicated when the road crossing is surrounded by trees or buildings etc. Fig. 5-14 show an example of the complexity of a road crossing with ramps and overpass.

Fig. 5-11 Matched vectors by SMA in test area 2

Fig. 5-12 Matched vectors from SMA in Test area 3

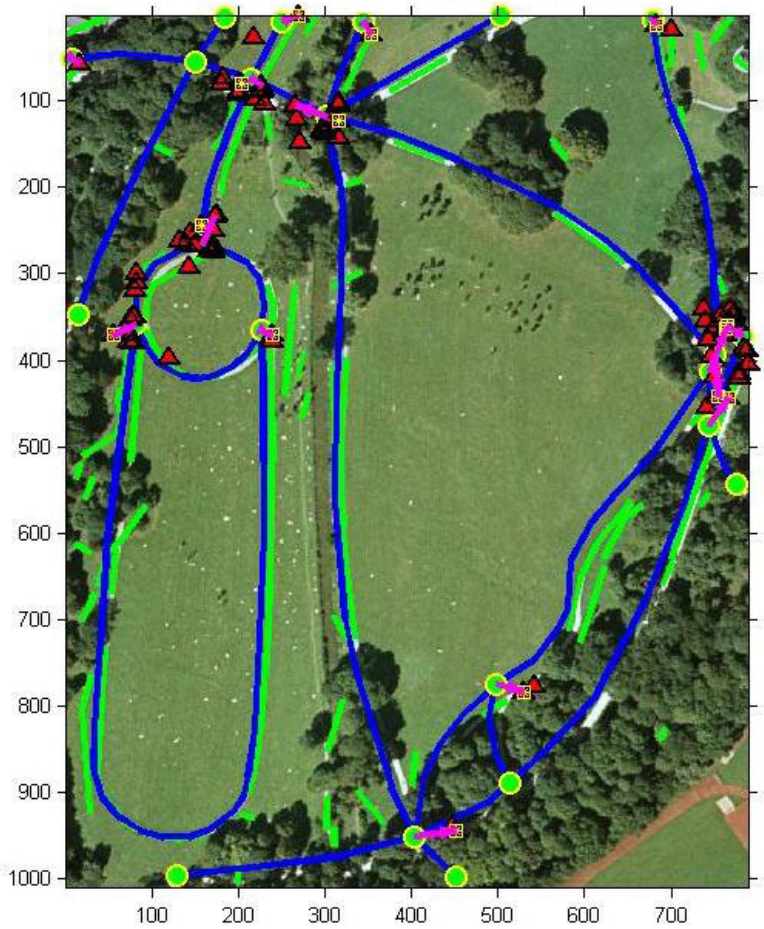


Fig. 5-13 Matched vectors from SMA in Test area 5

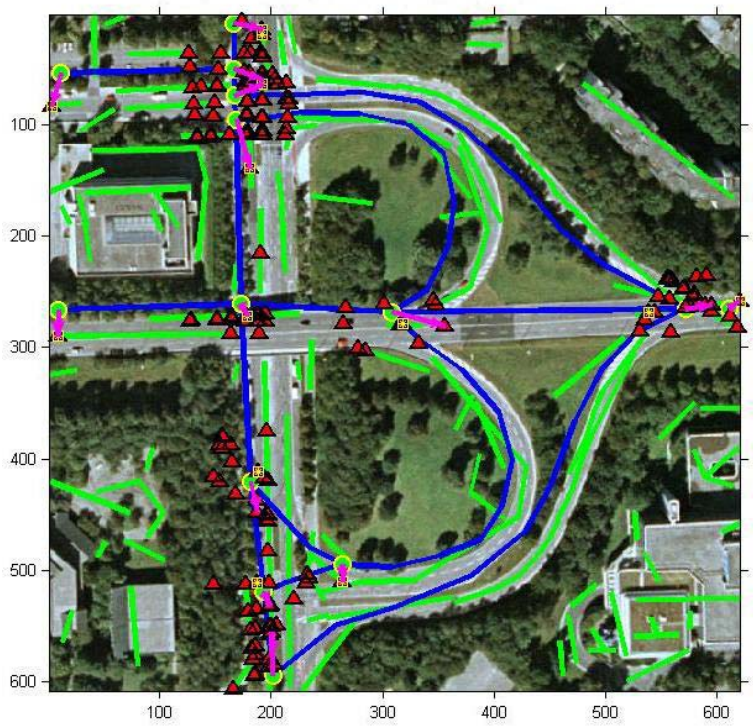


Fig. 5-14 Matched vectors from SMA in Test area 6

Fig 5-15 illustrates the SMA costs in comparison with the minimum and maximum costs of the reconstructed road network from the approximated intersections. The minimum and maximum costs were computed by choosing the intersections which had the minimum and maximum values from the super junctions. In our experiment, the SMA costs are close to the maximal cost or even higher if there is too much noise in the test dataset e.g. in area 2 and 5. The SMA result is an optimal and efficient solution, because the minimum and maximum cost can include broken nodes in the reconstructed road network.

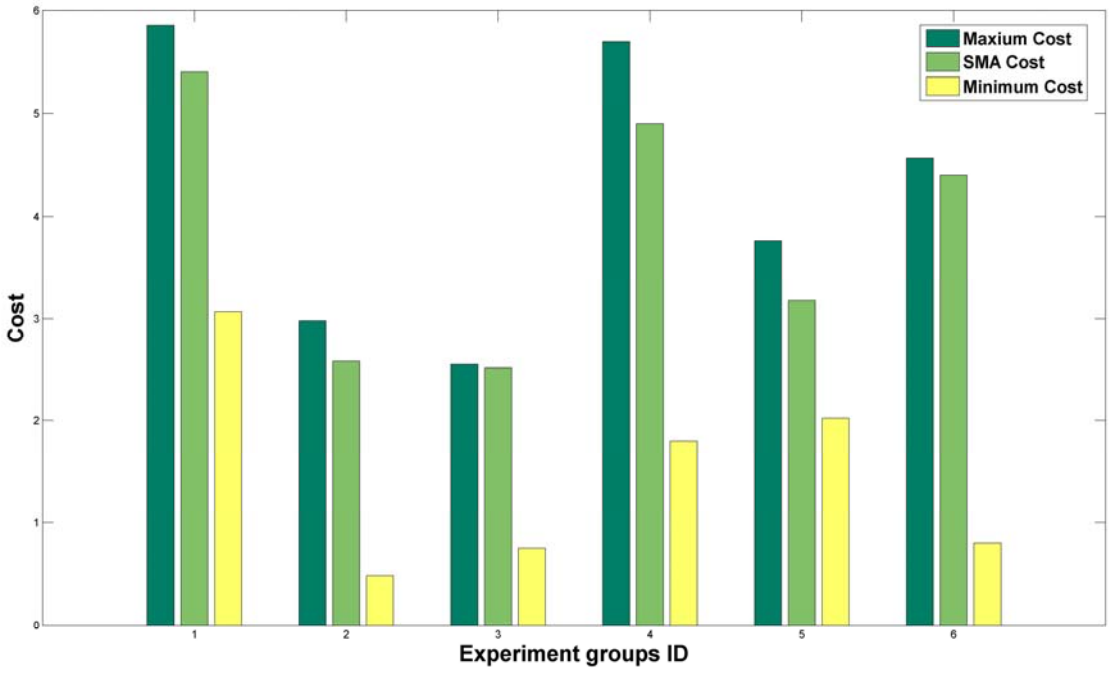


Fig. 5-15 Matching costs for six test areas

We have already verified the computation efficiency of SMA with different sizes of super junctions in Chapter 4. The performance of SMA is influenced by the complexity of the image, in particular the invisible ratio of the road. However, there is still no effective criteria to evaluate the complexity of the image. Another limitation of SMA is the required measure for the matching cost. We have formulated three measures for the matching costs as well as their weights, but it is still not enough in the practice, and there are some wrongly matched nodes even if the cost is minimal in the cluster. More measures or the modification of the weights are required for these wrongly matched nodes.

Another shortcoming of the SMA is its dependency on the parameters describing the input spatial datasets. Some parameters, for instance the Sigma and threshold in ECM, are related with the quality of the spatial data from the image. We manually estimated the matching quality, because it is still impossible to automatically estimate the matching quality in the SMA. To get better results, a rough estimation of the quality of extracted line elements is necessary, for instance, to allow a suitable setup of the weights for the fitness function.

5.4 Performance of Transformation Functions

5.4.1 Continuous transformation

The performance of congruent hybrids model depends on two factors: 1) the quality of the matched pairs created by SMA and 2) the performance of the transformation function which utilizes the matched results to transfer the road segments to its corresponding target, i.e. the extracted line elements from image. We focus on the second factor in this section. It is necessary to analyze the characteristics such as rigid or non-rigid deformations or the geometric offsets between the image and the reference road network before choosing the proper transformation function (Zagorchev and Goshtasby, 2006). The non-rigid deformation usually requires a continuous transformation function, for instance, the transformation used in medical image registration which contains the smooth curves of organic information, whereas the rigid deformation does not have such a requirement. The geometric offsets in our test areas are non-systemic non-rigid deformation. We will verify that the non-smooth transformation function is also suitable under certain conditions.

In our experiments, two favorite transformation functions, the piecewise Rubber-Sheeting (RUBS) and the Thin Plate Splines (TPS) are compared. The former is a non-smooth transformation function whereas the latter is a smooth one and one of the most widely used transformation functions for both rigid and non-rigid deformations.

The general TPS approach has been introduced in Section 2.1.2 when we discussed the conflation strategy, and we reformulate the generalized formulation for the geometric deformation. TPS follows the principle to generate meshes among the matched points, and individually transfer the features one mesh after another with minimized bending energy as expressed in Eq. 5-9.

$$f(x,y) = a_0 + a_1 x + a_2 y + \sum_{i=1}^{N} F_i d_i^2 \ln d_i^2 \qquad (5\text{-}9)$$

herein a_0, a_1, a_2 and F_i are the unknown coefficients; $d_i = \sqrt{(x-x_i)^2 + (y-y_i)^2}$, where (x_i, y_i) are the position of matched point *i*.

There are $N+3$ unknowns for the N control point pairs in Eq. 5-9; therefore, it requires additional constraints in Eq. 5-10 to compute the solution:

$$\sum_{i=1}^{N} F_i = \sum_{i=1}^{N} x_i F_i = \sum_{i=1}^{N} y_i F_i = 0 \tag{5-10}$$

The first constraint in Eq. 5-10 ensures that the plate would not move under the imposition of any loads in the plate and remain stationary. The second and the third constraint require that moments with respect to x and y axes to be zero, which ensures that the plate would not rotate under the imposition of the loads. The linear equation in Eq. 5-9 is solvable for given matched pairs.

The advantage of TPS is characterized by the radial basic function which monotonically increases and is spanned over the space domain. The logarithmic function in Eq. 5-9 is radially symmetric, so it prefers the regular grids or dense matched points in the input datasets. However, TPS is not suitable for correcting the local deformation and it may generate large errors when the matched pairs are irregularly or sparsely spaced because the radial basic function requires in principle symmetric input data to derive the parameters for global interpolation.

5.4.2 Piecewise Rubber-Sheeting (RUBS) transformation

The piecewise RUBS transform the feature piece by piece, where each piece has an independent linear function. Although so far only triangular regions have been used, the regions can take any shape and size. The piecewise RUBS function is continuous but not smooth. The Delaunay Triangulation is a popular approach to divide the area into sub-regions. Therefore, the transformation parameter is defined by the triangle vertices. Saalfeld (1985) presented a fast Rubber-Sheeting transformation using simplified (local barycentric) coordinates, which was derived from the nodes from two related triangles.

The RUBS is computing efficient and robust in comparison to the TPS approach which may induce the ill-posed problem if the pairs are sparsely located. When the regions are small or local deformations between road segments and the feature in images are small, RUBS proves sufficient in terms of matching quality and computation efficiency. However, if local deformations are large, the two triangles containing the reference and target may become quite different which will introduce large transformation errors because RUBS does not consider the information from the neighboring triangles.

5.4.3 Comparison of results from both transformation functions

In order to compare the performance of the two transformation functions, we use the ground truth as the reference dataset, and the Root Mean Square Error (RMSE) between the matched points and the transformed feature has been computed as the principal measure. The input of transformation functions is the matched points from SMA which contains inaccurate matching pairs as shown in Table 5-6. A manual correction of the transformed road network is unnecessary because the results of SMA may be improved by e.g. the snake-based algorithm as discussed in Chapter 2, which can automatically correct the gentle misalignments.

Besides the matched points, the end nodes of the extracted line which correspond to the matched intersections are also used as important information for the transformation because

the geometric structure of the road segments may be different in both datasets to be matched. The consideration of the geometric property in the matched results can increase the accuracy for the correctly matched features, but it primarily depends on the accuracy of matched points. If some prior knowledge of the road network in ground truth such as the curvature is known, we may decide whether the additional end nodes should be used. In our experiment, we used all the related end nodes by assuming that the prior knowledge is not available. In Fig.17~18, we compare the transformed results with and without the related end nodes from the extracted features. Large differences occur at curved road segments especially when the geometry of reference road networks deviate much from that of the extracted features. Without additional end nodes, the transformed road network will keep the same geometry as in the reference network, which is the typical solution of a classical point-based conflation approach.

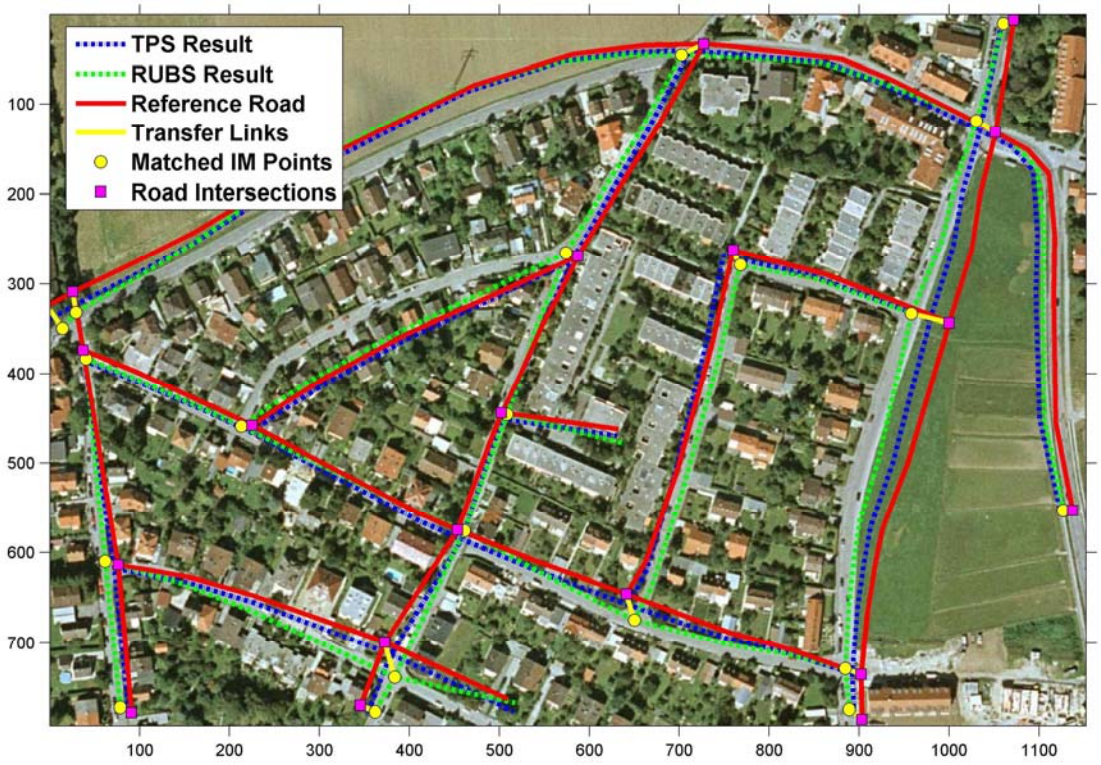


Fig. 5-17 Transformed road network with the matched points in test area 1 (Image Point, IM)

Fig. 5-19~5-23 show the transformed results from both algorithms overlaid with the reference road networks in test area 2-6. Each point used in the transformation function has been considered in the calculation of the RMSE which reflects the final quality of the entire model. Since the RMSE depends only on the discrete points used in the transform function, long road segments may reveal some misalignments compared to the ground truth. Moreover, additional factors may influence the matching quality, especially when the image is complex and only part of the road objects is visible.

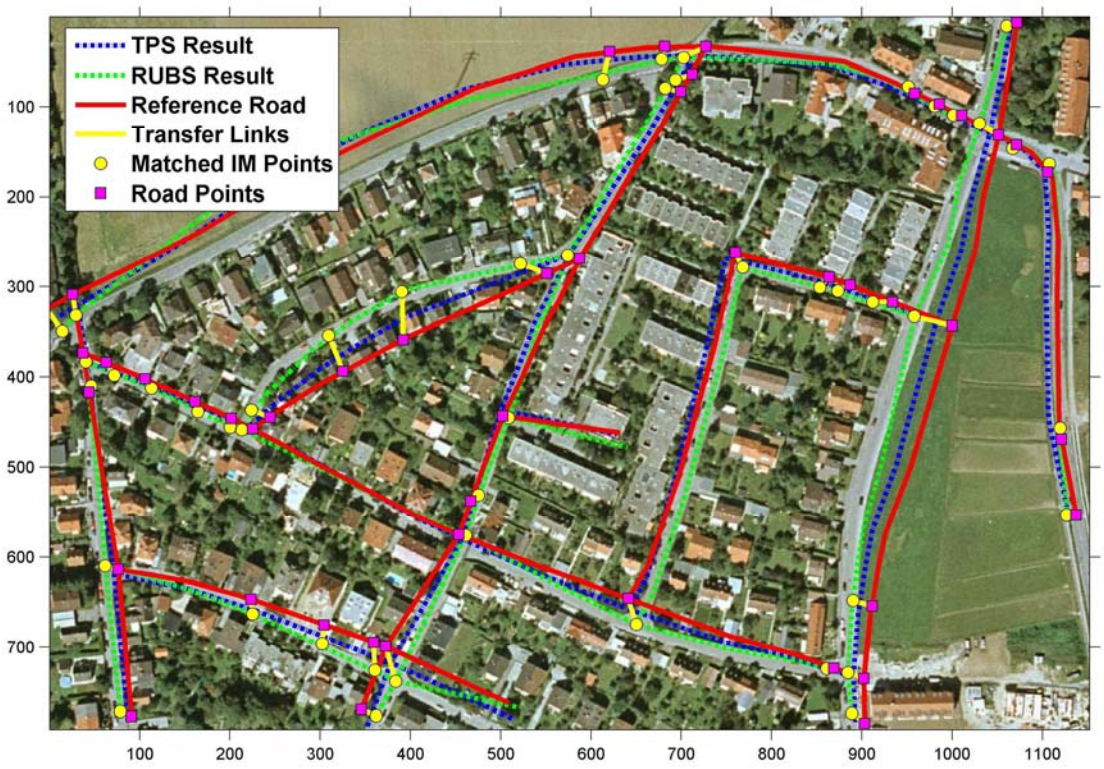


Fig. 5-18 Transformed road network with links to additional end nodes from extracted features in test area 1

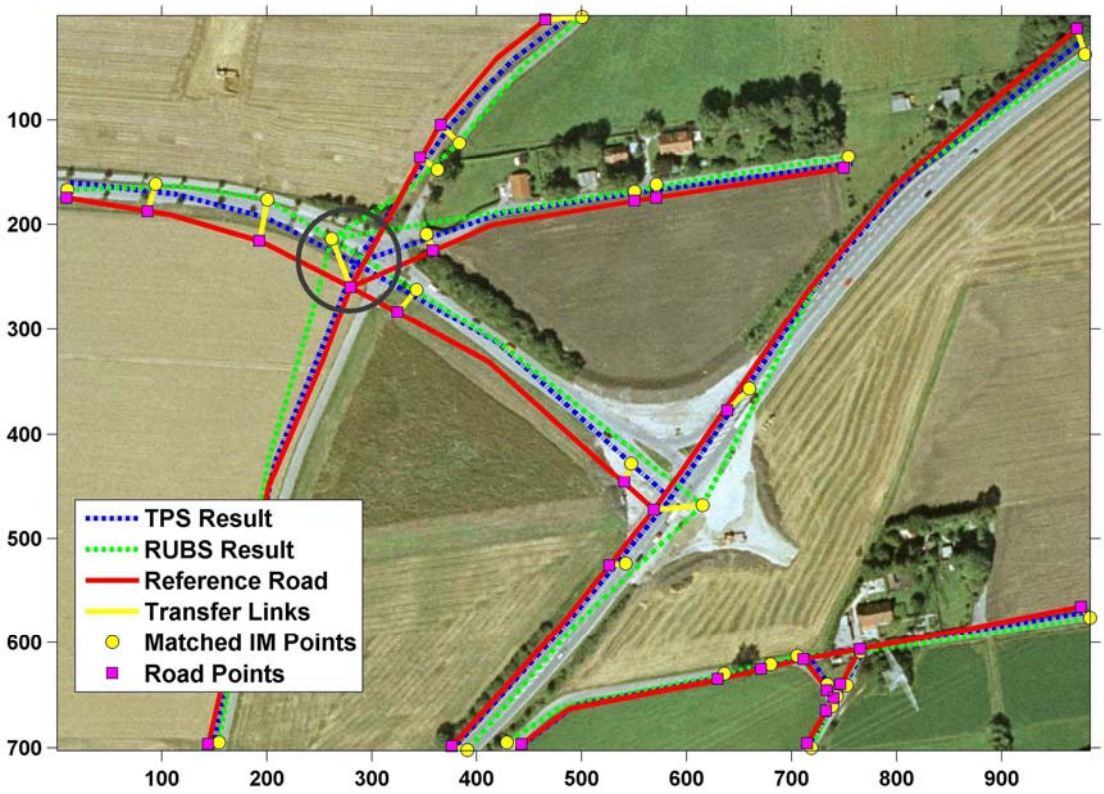


Fig. 5-19 Transformed results in test area 2

In Fig. 5-19, the transformation functions have different performances around the inaccurately matched points, for instance, the intersection in the black circle at (270,230). The

RUBS has bluntly transformed the road network to its counterpart in the image and the matched nodes have been snatched to the counterpart; whereas the TPS reveals a smoother transformation, where the transformed features are located between the reference road network and the results from RUBS. Around the accurately matched nodes, no large differences exist between the results from the two transformation algorithms.

Fig. 5-20 Transformed results in test area 3

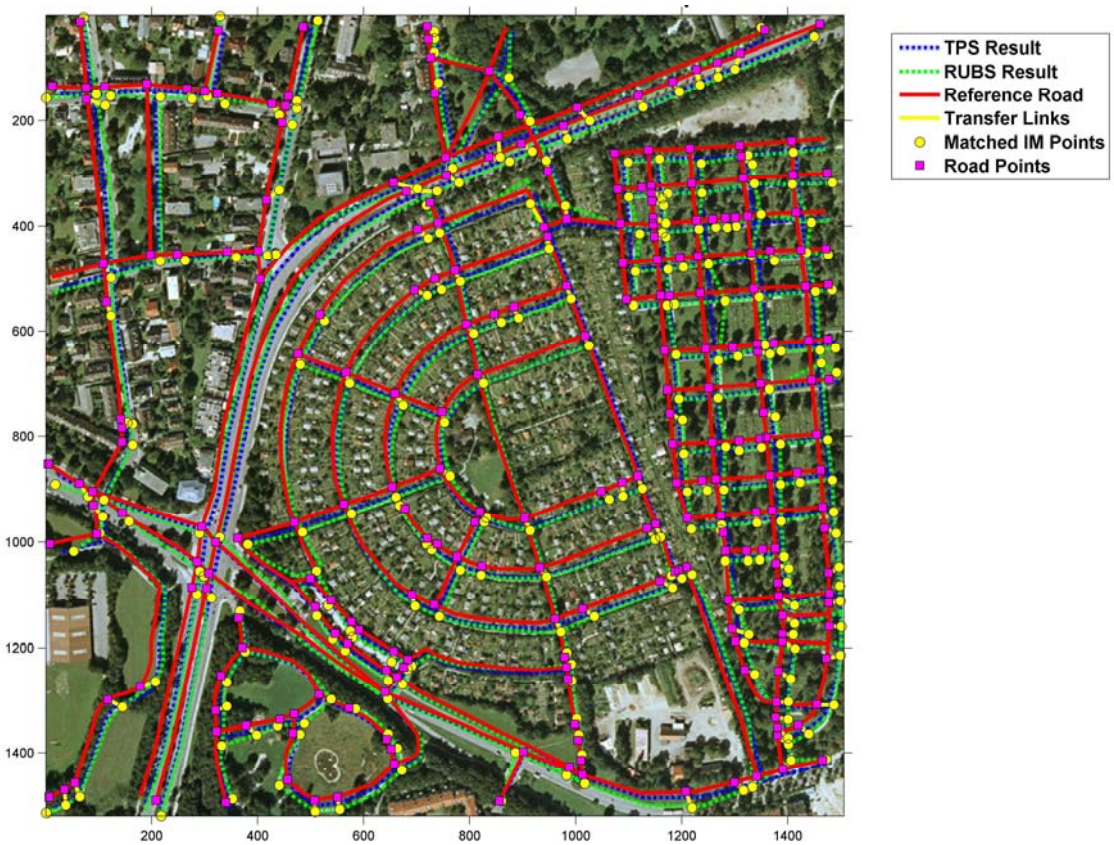


Fig. 5-21 Transformed results in test area 4

Test area 3 in Fig.5-20 shows a simple geometric pattern with straight road segments.

This leads to a high correctness of estimating the intersections from the grouped lines. Apart from the undetected segments by Canny operator, the overall quality is acceptable even without further modification.

Test area 4 in Fig.5-21 is the most complex area in our experiments. The reference road network has a high fidelity to the ground truth, but the features are really multifarious. It includes both the expressway and the pedestrian way, and the trees are dominating in certain area. It has a lower visible ratio comparing to test area 1~3. The curve road segment has been approximately transformed to the matched nodes, but the error of the transformation is quite obvious on the right side, where the extraction ratio is also limited because of the trees.

Fig. 5-22 Transformed results in test area 5

Fig.5-22 reveals a visible ratio of only around 50% of all road segments. Furthermore, the visible road segments are sheltered by the trees, making the grouping of extracted lines difficult. The road segments which cannot be extracted take over a similar geometry and the same topology from the reference road network.

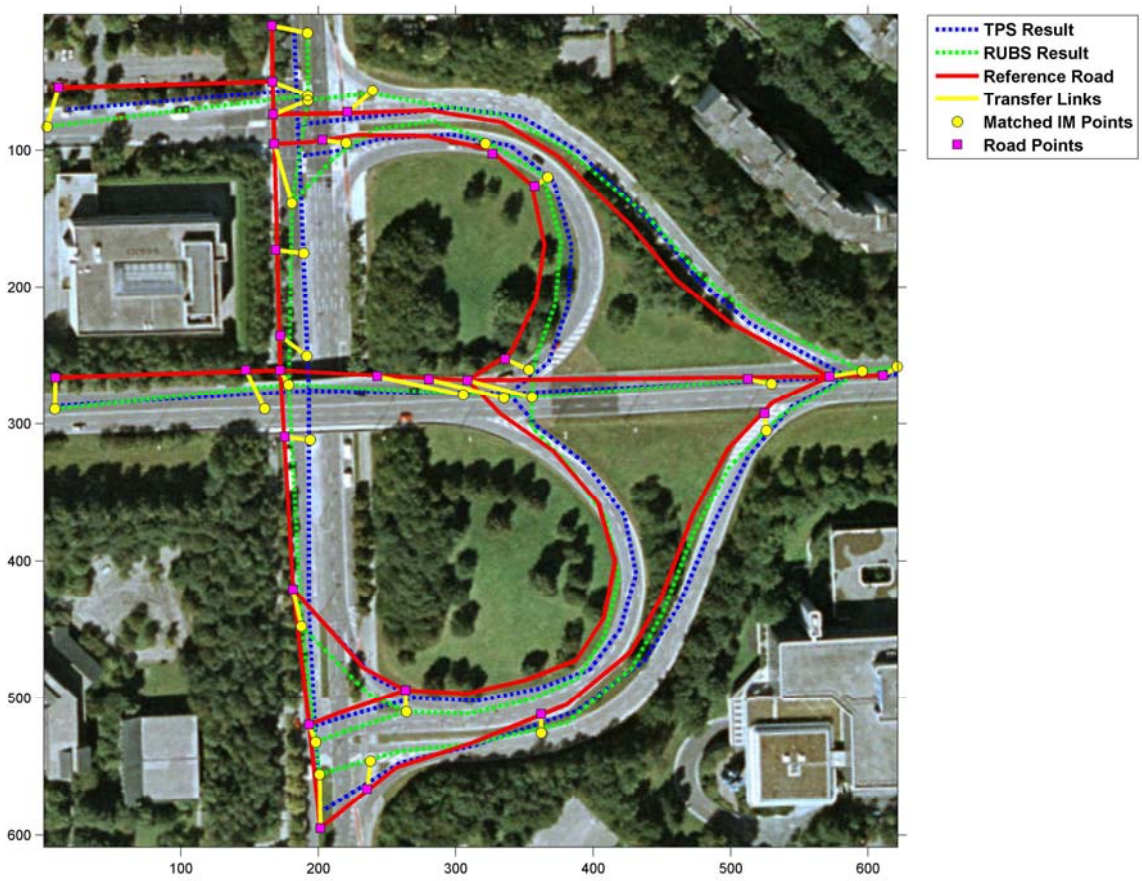


Fig. 5-23 Transformed results in test area 6

Fig. 5-23 contains a complex urban interchange near Munich. The zebra lines as well as other road-related features are detected by Canny operator as linear objects, moreover, a part of the road surface that has a different radiometric property has been also detected as linear objects. This means that not only the visibility of the road, but the structure of the road itself can influence the matching quality. Obviously, the transformed result needs additional improvements, e.g. by means of a snake-based algorithm.

Table 5-7 summarizes the RMSE errors i.e. original RMSE between the reference road network and the ground truth, and the RMSE errors from the matched data in the experiments. The results from the approach with or without the additional extracted nodes are compared to show the importance of including geometric information for the conflation approach.

In the experiments, all the RMSE value were reduced in the transformed results except for the test area 5, which was with a low visible ratio of the road surface, so the matched points contain much inaccurate information which can be enhanced by the transformation functions. The additional information from the extracted lines is helpful to generate high quality spatial data, but the inaccurate additional points (in test area 6) and the undetected road object (in test area 3) produce even worse results than the approach only with the matched points.

Table 5-7 Transformation results from RUBS and TPS

Area ID	Original RMSE (m)	RMSE (m)			
		Only matched Pts		With additional end nodes	
		RUBS	TPS	RUBS	TPS
1	15.4	7.0	8.6	5.9	7.8
2	8.5	6.4	5.6	6.0	4.9
3	7.7	5.2	5.0	5.3	5.5
4	9.4	6.2	5.0	6.0	4.9
5	5.1	8.4	7.1	8.1	6.4
6	10.3	8.2	6.1	8.6	8.7

The experiments have verified the ability of both RUBS and TPS to wrap the original road network to its counterpart feature in the image using the matched point from SMA, and the transformed results are acceptable for further process. RUBS has better performance when it has accurately matched points, but it is error prone if the matched feature is not accurate or even false; whereas TPS considers more global information from the matched features and therefore provides smoother results and better performance even when the matched points are not sufficiently accurate, so the TPS is a desirable choice to deal with areas with more wrongly matched pairs.

As a whole, our approach proves robust in dealing with noise, e.g. the cars and shadows in the test areas, because the SMA does not require the complete detection of road features. The complexity of the image is an important influence factor for the image-to-vector conflation, but difficult to quantify. The visibility of the road surface in the image and the road structure are further influence factors. The road objects after transformation are close to the centre of road features in the image, but small modifications are still necessary, e.g. snake-based method could be applied to optimize the final results.

5.5 Discussion

We presented and tested the congruent hybrid conflation model for a reference road network and a geo-referenced image where the misalignments of the road features are corrected. The proposed approach is capable of identifying corresponding road features in image and in road network. Additional preprocessing is triggered to decrease the noise ratio for the extracted road features, and thus to assure a stable performance of the sparse matching procedure. We choose GA to exclude the wrong candidates of road features. The correspondences have been computed with the sparse matching algorithm which demonstrate the *1-to-n* cardinalities and reflect the perceptual factors of geometry, topology and intensity.

In the experiments, the noise ratio in the detected edges has been dramatically reduced after matching operation with ECM and grouping operation with GA, but it remains a challenge for the classical matching algorithm to get proper matching results from a highly noisy image.

One important reason is the diversity of the radiometric property of the road in the image. For instance, if a road segment is divided into two parts, they can be hardly grouped by the GA algorithm because of the large radiometric difference. If the tolerance of radiometric difference is increased, more noise may interfere the matching algorithm.

Test regions with straight road segments have better matching performance from the proposed model, but it is a rare situation to have regions full of straight segments. Our model has also considered the curvature of roads and is able to generate better transformed results than the classical point-based conflation approach. However, the estimation of the curvature is not always stable, especially when the extracted features contain many fragmental pieces and the additional end nodes from the extracted features may be inaccurate, thus do not lead to improvement of matching accuracy.

Two transformation functions RUBS and TPS have been compared, and both of them are efficient in our experiments, and the average RMS value has been decreased. The transformed road network keeps the original topology from the reference spatial dataset, but the geometry of the reference road segments have been snatched to the matched features from the image. TPS is generally more efficient than RUBS which works well only with homogeneously distributed and accurate matched points. However, in despite of the additional end nodes from the extracted features, the density of the matched points is still relatively low for TPS which prefers the denser distribution to determine the reliable transformation parameters. Because of the global consideration of the matched points, TPS is less sensitive to matching errors and provides smoother results than RUBS. We summarize the pros and cons for the two functions in Table 5-8.

Table 5-8 Transformation results from RUBS and TPS

ID	Factors	Priority	
		RUBS	**TPS**
1	Deformation type	Rigid Deformation in both datasets	Non-rigid Deformation in both datasets
2	Density of matched points	For both dense density or sparse density	Suitable for dens density with grid matched points
3	Accuracy of matched points	No requirement	No requirement
4	Errors propagation in transformation	Local influence	Global influence
5	Computation Complexity	Low	High

The overall performance of the proposed congruent hybrid conflation model is slightly influenced by the quality of the input datasets i.e. the complexity of the image as well as the reference road network. In the test, we choose the visibility of the road in the image as a measure of the image's complexity, and the maximal visible ratio in the tests is around 50%

which could be even smaller in urban or forest area. The correctness (Table 5-6) and the transformation error (Table 5-7) show the relation between the quality of the input data and the quality of the results. In the test, the reference road network has reliable topology, although it contains deformed road segments. Our conflation model can generate high quality results if one of the two spatial datasets has a reliable quality.

Chapter 6

Conclusions and Outlook

6.1 Conclusions

The thesis focuses on the spatial data conflation problem between a reference vector database and its image data which have some obvious misalignments. The essential task involved in the conflation is the correct identification of the inherent correspondences which may take various cardinalities such as *1-1*, *1-0*, *1-n* to *m-n* etc. We comprehensively presented and discussed a Congruent Hybrid Model (CHM). Unlike the classical point-based conflation approach, CHM works on the linear features from the image. It is composed of four modules characterized by the functions of extraction, grouping, matching and transforming.

The main contributions of the CHM include: 1) a linear feature extraction approach with two algorithms ECM as well as the GA-based grouping; 2) a novel SMA approach which can deal with multiple correspondences; and 3) two comparative transformation functions RUBS and TPS. The feasibility and performance of the CHM has been verified using synthetic datasets and tested on real world datasets. The accomplished work can be summarized in four aspects:

1) Linear feature extraction and grouping

Patterns of linear features in the image can be detected to a certain extent. Most of cartographic objects e.g. roads, buildings, rivers, can be decomposed into clusters of individual lines. The road elements in the VHR geo-referenced image appear as elongated area objects. The ideal parallel borders of a road are difficult to recognize with computer algorithms, especially when much noise is involved. In this work, we adopt the edge operator Canny operator to derive edges from the image. The areas between the extracted edges form interesting regions which may likely belong to road objects. The ECM algorithm is developed to search the linear features from the extracted edges in the whole test region. The common-sense knowledge about the road, such as the road width, can be fed into the algorithm. Although the noise level in the results of ECM is still quite high compared with the ground truth, up to 90% redundancy in the extracted edges can be reduced as shown in Table 5-2. The ECM suffers from a delicate problem with the fragmental lines inherited from the edge operator which only recognizes the potential edges based on the difference between neighboring pixels. Another problem of the ECM is reflected in the over extraction of linear features which not only belong to road objects, but other objects such as roof outlines. The subsequent GA-based grouping and the SMA are respectively devoted to solving these two problems.

The extracted lines are very fragmental and 10 times in number as those of ground truth in our tests. Since the matching algorithm can hardly handle the outliers, the fragmental lines should be processed. In fact, the extracted centrelines that are sequentially aligned near to each other or have the similar radiometric property can be chained into longer lines. Due to the huge searching space, however, we use GA as a more efficient means to generate

connections. Without having to search throughout the whole region, GA relies on a fitness function which considers the geometric, the topologic and the radiometric characters of extracted centrelines. In the experiments, up to 40% of fragmental centrelines can be chained into longer lines, although the grouping effort does not improve the correctness and completeness of the results from the ECM.

2) Matching quality from sparse correspondence

The up-to-date point-based matching algorithm is rather sensitive to outliers, as discussed in Chapter 2. With 20% outliers or more, the matching algorithm will not properly work. However, the image data for our conflation task has a very high noise ratio. The extracted linear features contain many features other than roads, although some unrelated segments have been excluded after GA-based grouping operation. The existing algorithms are unable to completely discriminate the road objects from other linear features which reveal the same radiometric and geometric patterns as the roads. For this reason, the SMA is developed as a novel matching algorithm to handle the multiple correspondences in the matching process.

It is assumed that the reference road network has a reliable topology. Previous studies have proved that the gestalt decomposition of the reference road network into cliques or super junctions can increase the matching efficiency. A preliminary correspondence between the decomposed road network and linear features from the image is established by an adaptive buffer around the individual super junctions. Each super junction represents one node from the reference road network and corresponds to one or more intersection nodes among grouped centrelines extracted from the noisy image data. Estimating the fittest intersection among the multiple correspondences is termed as sparse matching. In certain test areas, one node may correspond up to 14 intersections. To deal with the extreme sparse ratio, we define the matching problem as a Constrained Minimum Perfect Matching Problem (C-MP^2). The simplest form C-MP^2 is the 1-1 correspondence between the super junction and the intersection. Unfortunately, the C-MP^2 is an NP-hard problem for which optimal solutions can be created only for large datasets.

Instead of directly solving the objective function in Eq. 4-10, we choose the greedy strategy in SMA to compute the optimal cost between the super junctions and the estimated nodes using a similarity measure. The greedy approach has the advantage of a predicable optimal cost in the optimization problem, and it avoids the time-consuming procedure in searching globally minimal cost. The efficiency and the false ratio of SMA have been also computed to indicate the matching quality compared with the ground truth.

3) Influences of the transformation function

Besides the accurate matching points, the transformation function also plays an important role for the correction of misalignments between the two datasets. We applied two favourite transformation functions RUBS and TPS. Both functions have the ability to transform the matched feature to its correspondence properly. The topology of the reference road network was kept in the transformed results. However, RUBS and TPS have different characteristics: RUBS transformed features using the neighbors in the triangulation, whereas TPS considered the global influence of all the matched points in the test region. Moreover, these two functions have different abilities to deal with the noise. The transformed results from RUBS are close to the ground truth if the matching quality is high, and the results from TPS have better performance if there are inaccurate matched points.

4) Performance and the crucial parameters of CHM model

The CHM is an integrated solution for the vector-to-image conflation problem, although the individual modules can be operated separately. The efficiency of each module can be determined using various measures. The CHM has reached high correctness and completeness in the test areas with varying image complexity which is quite hard to estimate. There are many different measures to indicate the image complexity. The visibility of the road surface is one of these measures and has been manually calculated in these test areas. A low visibility ratio may lead to some unexpected results with larger RMSE than the reference road network.

We have analyzed the computing efficiency of SMA with synthesized datasets as well as the real spatial datasets. ECM and GA are both computing efficient in test areas. As compared to the point-based conflation approach, the CHM has a better geometric solution. In areas where the road intersections are invisible, the CHM can still generate results if part of the road segments can be seen. Moreover, the CHM has the potential to correct some wrong geometry in the reference road network, although the additional information from the extracted feature is not always as stable as the super junctions.

6.2 Future Works

In spite of the extensive consideration and convincing performance of the CHM in test regions, we anticipate a number of further improvements and extensions in order to keep pace with rapid technological evolutions:

1) Optimization of the required parameters

It should be noted that the CHM is dependent on many parameters in each module, and some of them are related with the selected datasets. This may restrict the universal matching quality of the CHM. A compromise could be the optimization of several key parameters. For instance, some prior knowledge about the reference road network can be integrated with the fitness functions so that the GA-based grouping approach could deliver more intelligent results.

2) Design of effective similarity measure

In our experiments, we have considered only some simple similarity measures for SMA. Furthermore, the interaction among the measures is determined in terms of weights. Different weight combinations are intuitively assigned in our experiments and it is hard to prove that they bring about the optimal results. The matching procedure can be improved in two aspects. First, designing more effective similarity measures which consider geometric and radiometric properties of the road network could lead to a higher matching ratio. Second, automating the estimation of the interactions among the similarity measures may also enhance the matching performance. Both SMA and GA need additional iterations to get more information from the reference road datasets. The new similarity measures should be based on the derived characteristics from the image using preliminary image processing algorithms.

3) Matching based on Spatial Point Process (SPP)

The estimated intersections in SMA are treated as the real intersections, and the matching cost is based on the subjective assumption. In fact, an estimated intersection represents only a probable node which has a comprehensive geometric, topological and radiometric relation. It

should be also possible to incorporate more high-level knowledge such as the contextual information from the image with the matching process. All the related information can be integrated as a special mark in SPP which is a stochastic process based on the current knowledge, and the SMA can be reformulated as a model selection problem. The SPP-based matching approach may reveal a high robustness against noise, even if the test area has a very low invisible ratio.

4) Conflation of image from other high accurate sensors

The CHM can be further improved with the development of new technologies in two aspects: the advanced pattern recognition algorithm and the highly accurate sensor other than the optical sensors. In CHM, a classical edge operator is adopted to extract the comprehensive linear edge. From more accurate sensors, either optical or radar sensors, some pre-processing procedures such as edge detection can be integrated into object detection. That means, road objects can be directly extracted using more context information from high resolution image. The current research works in the pattern recognition community are more increasingly dedicated to advanced pattern recognition algorithms which may support discrimination of specific objects from a noisy image. The performance of CHM could become more reliable if the road features can be separated from other linear objects before the matching procedure starts.

Highly accurate active sensors, such as SAR and laser scanners, have been rapidly developed in the recent decade. They can deliver geo-referenced image data with higher resolutions. Moreover, some sensors have the ability to penetrate surface features such as trees. For instance, if the high resolution LIDAR DEM can be incorporated with SAR measurement (Stilla, et al., 2004), the visibility of road features can be increased, leading to improved matching quality.

Bibliography

Adluru, N.; Latecki, L.J. (2009): Contour Grouping Based on Contour-Skeleton Duality. In International Journal of Computer Vision, Vol. 83, No.1, pp. 12~29.

Amini, A.A; Tehrani, S.; Weymouth, T.E (1988): Using Dynamic Programming For Minimizing The Energy Of Active Contours In The Presence Of Hard Constraints. In : [1988 Proceedings] Second International Conference on Computer Vision: IEEE, pp. 95~99.

Andriy, M. (2010): Non-rigid Image Registration: Regularization, Algorithms and Non-rigid Image Registration: Regularization, Algorithms and Applications. Oregon Health & Science University.

Anthony, E.L.; William, H.M. (Eds.) (1987): A general approach to map conflation. Proceedings of the International Symposium on Computer-Assisted Cartography. Maryland. Available online at http://mapcontext.com/autocarto/proceedings/auto-carto-8/index.html.

Baihua, L.; Qinggang, M.; Holstein, H. (2003): Point pattern matching and applications - A review. In : SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483): IEEE, pp. 729~736.

Baltsavias, E.P. (2004): Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No.3-4, pp. 129~151.

Barbara, Z.; Jan, F. (2003): Image registration methods: a survey. In Image and Vision Computing, Vol. 21, No.11, pp. 977~1000.

Barrow, H.G. Tenenbaum, J.M Bolles R.C. Wolf H.C (1977): Parametric correspondence and chamfer matching: two new techniques for image matching. In : Proceeding IJCAI'77 Proceedings of the 5th international joint conference on Artificial intelligence, Vol. 2, pp. 659~663. Available online at http://ijcai.org/Past%20Proceedings/IJCAI-77-VOL2/PDF/024.pdf, checked on 21/05/2012.

Barsi, A.; Heipke, C. (2003): Detecting road junctions by artificial neural networks. In : Proceedings 22nd Digital Avionics Systems Conference. (Cat. No.03CH37449): IEEE, pp. 129~132.

Besl, P.J; McKay, H.D. (1992): A method for registration of 3-D shapes. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No.2, pp. 239~256.

Bonnefon, R.; Dherete, P.; Desachy, J. (2002): Geographic information system updating using remote sensing images. In Pattern Recognition Letters, Vol. 23, No. 9, pp. 1073~1083.

Bookstein, F.L. (1989): Principal warps: thin-plate splines and the decomposition of deformations. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 6, pp. 567~585.

Bordes, G.; Giraudon, G.; Jamet, O. (1997): Road Modeling Based on a Cartographic Database for Aerial Image Interpretation. In In: Grün, A., Baltsavias, E.P., Henricsson, O. (Eds.)Automatic extraction of man-made objects from aerial and space images (II), pp. 23~139.

Borgefors, G. (1988): Hierarchical chamfer matching: a parametric edge matching algorithm. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 10, No. 6, pp. 849~865.

Brown, L.G. (1992): A survey of image registration techniques. In ACM Computing Surveys, Vol. 24, No. 4, pp. 325~376.

Butenuth, M. (2008): Network snakes. Deutsche Geodätische Kommission, Reihe C, Nr. 620, Verlag der Bayerischen Akademie der Wissenschaften, ISBN 978-3-7696-5034.

Byoung-Ki, J.; Jeong-Hun, J.; Ki-Sang, H. (2002): Road detection in spaceborne SAR images using a genetic algorithm. In IEEE Transactions on Geoscience and Remote Sensing, Vol. 40, No. 1, pp. 22~29.

Canny, J. (1986): A Computational Approach to Edge Detection. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679~698.

Chiang, Y.; Knoblock, C.A.; Shahabi, C.; Chen, C.C. (2009): Automatic and Accurate Extraction of Road Intersections from Raster Maps. In Geoinformatica, Vol.13, No. 2, pp. 121~157.

Chen, C.C. (2005): Automatically and and Accurately Conflating Road Vector Data Street Maps and Orthoimagery. University of Southern California.

Chen, C.C.; Knoblock, C.A.; Shahabi, C. (2006): Automatically Conflating Road Vector Data with Orthoimagery. In Geoinformatica, Vol. 10, No. 4, pp. 495~530.

Christine, P.; Stefano, S. (2000): Database Integration: the Key to Data Interoperability. Edited by M. P. Spaccapietra S. &. Z. Tari Papazoglou. The MIT Press, checked on 12/12/2011.

Chui, H. (2001): Non-Rigid Point Matching: Algorithms, Extensions and Applications. Dissertation, Yale University.

Chui, H.; Rangarajan, A. (2003): A new point matching algorithm for non-rigid registration. In Computer Vision and Image Understanding, Vol. 89, No. 2-3, pp. 114~141.

Clerico, Y.; Stefan, S.; Schleder, D. (2009): Arbeitshilfe Geodaten in der Praxis. Bayerische Vermessungsverwaltung.

Cobb, M.A.; Chung, M.J.; Foley, I.H.; Petry, F.E.; Shaw, K.B.; Miller, H.V. (1998): A Rule-based Approach for the Conflation of Attributed Vector Data. In Geoinformatica, Vol. 2, No. 1, pp. 7~35.

Cohen, E.H.; Singh, M. (2007): Geometric determinants of shape segmentation: Tests using segment identification. In Vision Research, Vol. 47, No. 22, pp. 2825~2840.

Conte, D.; Foggia, P.; Sansone, C.; Vento, M. (2004): Thirty Years of Graph Matching in Pattern Recognition. In International Journal of Pattern Recognition and Artificial Intelligence, Vol. 18, No. 3, pp. 265.

Conte, D.; Foggia, P.; Sansone, C.; Vento, M. (2003): Graph matching applications in pattern recognition and image processing. In : Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429).

Cook, W.J.; Cunningham, W.H.; Pulleyblank, W.R.; Schrijver, A. (1998): Combinatorial optimization. New York: Wiley.

Cordella, L. P.; Stefano, C.; Frucci, M. (2000): A Robust Shape Decomposition Method. In Atul K. Chhabra, Dov Dori (Eds.): Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 219~227.

Cui, M.; Femiani, J.; Hu, J.; Wonka, P.; Razdan, A. (2009): Curve matching for open 2D curves. In Pattern Recognition Letters, Vol. 30, No. 1, pp. 1~10.

David, P.; DeMenthon, D.(2005):Object recognition in high clutter images using line features. Proceedings of the 10th IEEE International Conference on Computer Vision, pp. 1581~1588.

Deschênes, F.; Ziou, D. (2000): Detection of line junctions and line terminations using curvilinear features. In Pattern Recognition Letters, Vol. 21, No. 6-7, pp. 637~649.

Desolneux, A.; Moisan, L.; Morel, J. (2008): From Gestalt theory to image analysis. A probabilistic approach. New York, NY: Springer.

Dijkstra, E.W. (1959): A note on two problems in connexion with graphs. In Numerische Mathematik, Vol. 1, No. 1, pp. 269~271.

Feremans, C.; Labb, M.; Laporte, G. (2002): A comparative analysis of several formulations for the generalized minimum spanning tree problem. In Networks, Vol. 39, No. 1, pp. 29~34.

Feremans, C.; Labb, M.; Laporte, G. (2004): The generalized minimum spanning tree problem: Polyhedral analysis and branch-and-cut algorithm. In Networks, Vol. 43, No. 2, pp. 71~86.

Ferrari, V.; Fevrier, L.; Jurie, F.; Schmid, C. (2008): Groups of Adjacent Contour Segments for Object Detection. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 1, pp. 36~51.

Fiset, R.; Cavayas, F.; Mouchot, M.C.; Solaiman, B.; Desjardins, R. (1998): Map-image matching using a multi-layer perceptron: the case of the road network. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 53, No. 2, pp. 76~84.

Floyd, R.W. (1962): Algorithm 97: Shortest path. In Commun. ACM, Vol. 5, No. 6, pp. 345.

Fortier, A.; Ziou, D.; Armenakis, C.; Wang, S. (2001): Survey of work on road extraction in aerial and satellite images. In Canadian Journal of Remote Sensing, Vol. 27, No.1, pp. 76~89.

Fortier, A.; Ziou, D.; Armenakis, C.; Wang. S. (2001): Automated Correction and Updating of Road Database from High Resolution Imagery. In Canadian Journal of Remote Sensing, Vol.27, No. 1, pp. 76~89.

Gabay, Y.; Doytsher Y. (1994): Automatic adjustment of line maps. Phoenix, USA (Proceedings of the GIS/LIS094 annual convention).

Gallo, G.; Pallottino, S. (1988): Shortest path algorithms. In Annals of Operations Research, Vol. 13, No. 1, pp. 1~79.

George, L.F.; Henry, F.S. (Eds.) (1987): Improvement of GBF/DIME file coordinates in a geo-based information system by various transformation methods and rubber sheeting based on triangulation. Proceedings of the 8th International Symposium on Computer-Assisted Cartography. Maryland. http://mapcontext.com/autocarto/proceedings/auto-carto-8/pdf/improvement-of-gbfdime-file-coordinates-in-a-geobased-information.pdf, checked on 14/11/2011.

George, V.; Marlies, de G. (1997): Updating Road Maps by Context Reasoning. In In: Grün, A., Baltsavias, E.P., Henricsson, O. (Eds.)Automatic extraction of man-made objects from aerial and space images (II), pp. 267~276.

Gerke, M.; Heipke, C. (2008): Image‐based quality assessment of road databasesX. In International Journal of Geographical Information Science, Vol. 22, No. 8, pp. 871~894.

Gerke, M.; Butenuth, M.; Heipke, C.; Willrich, F. (2004): Graph-supported verification of road databases. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No. 3-4, pp. 152~165.

Gillman, D. (1985): Triangulations for rubbersheeting. Washington (Proceedings of Autocarto 7). Available online at http://mapcontext.com/autocarto/proceedings/auto-carto-7/pdf/triangulations-for-rubber-sheeting.pdf, checked on 2011-12.

Gisela, K. (2003): A comparative discussion of distance transforms and simple deformations in digital image processing. In Machine Graphics & Vision International Journal, Vol. 12, No 2.

Goldberg, D.E. (1989): Genetic algorithms in search, optimization, and machine learning. Repr. with corr. Reading, Mass: Addison-Wesley.

Goodchild, M.F.; Hill, L.L. (2008): Introduction to digital gazetteer research. In International Journal of Geographical Information Science, Vol. 22, No.10, pp. 1039~1044.

Gunst, M.E de; den Hartog, J.E. (1994): Knowledge-based updating of maps by interpretation of aerial images. In : Proceedings of 12th International Conference on Pattern Recognition: IEEE Comput. Soc. Press, pp. 811~814.

Haunert, J.H.; Sester, M. (2008): Area Collapse and Road Centerlines based on Straight Skeletons. In Geoinformatica, Vol. 12, No. 2, pp. 169~191.

Heiner, H.; Fritsch, D. (1998): Integration of Vector Data and Satellite Imagery for Geocoding. In proceedings of GIS-Between Visions and Applications, Vol. 32, No. 4, pp. 246~251.

Hill, D.L.G.; Batchelor, P.G.; Holden, M.; Hawkes, D.J. (2001): Medical image registration. In Physics in Medicine & Biology, Vol. 46, No. 3.

Hinz, S.; Baumgartner, A. (2003): Automatic extraction of urban road networks from multi-view aerial imagery. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No.1-2, pp. 83~98.

Hinz, S. (2003): Automatische Extraktion urbaner Straßennetze aus Luftbildern. PhD Thesis. Technischen Universität München.

Hu, B. (2008): Hybrid Metaheuristics for Generalized Network Design Problems. PhD thesis. Vienna University of Technology. Institute of Computer Graphics and Algorithms, checked on 28/05/2012.

Hu, J.; Razdan, A.; Femiani, J.C.; Cui, M.; Wonka, P. (2007): Road Network Extraction and Intersection Detection From Aerial Images by Tracking Road Footprints. In IEEE Transactions on Geoscience and Remote Sensing, Vol. 45, No. 12, pp. 4144~4157.

Huang,X.Q.; Xion, Y..; Liao, M.; Tang, Y.Y. (2007): Skeletonization of ribbon-like shapes based on local minimum modules of wavelet transform. In : 2007 International Conference on Wavelet Analysis and Pattern Recognition: IEEE, pp. 1247~1251.

Huygens, D. (2002): Version generalisee du probleme de conception de reseau 2-areteconnexe. Master's thesis. Universite Libre de Bruxelles.

Jain, A.K. (2010): Data clustering: 50 years beyond K-means. In Pattern Recognition Letters, Vol. 31, No. 8, pp. 651~666.

Jia, C.; Zhao, L.; Wu, Q.; Kuang, G. (2008): Automatic Road Extraction from SAR Imagery Based on Genetic Algorithm. In Journal of Image and Graphics, Vol. 13, No. 6.

Jian, B.; Vemuri, B.C (2005): A robust algorithm for point set registration using mixture of Gaussians. In : Tenth IEEE International Conference on Computer Vision (ICCV'05) Vol. 2, pp. 1246~1251.

Ju Z.; Hong, Y. (2001): Skeletonization of ribbon-like shapes based on regularity and singularity analyses. In IEEE Transactions on Systems, Man, and Cybernetics. Part B, Vol. 31, No. 3, pp. 401~407.

Kass, M.; Witkin, A.; Terzopoulos, D. (1988): Snakes: Active contour models. In International Journal of Computer Vision, Vol. 1, No. 4, pp. 321~331.

Kimia, B.B.; Tannenbaum, A.R.; Zucker, S.W. (1995): Shapes, shocks, and deformations I: The components of two-dimensional shape and the reaction-diffusion space. In International Journal of Computer Vision, Vol. 15, No. 3, pp. 189~224.

Klang, D. (1998): Automatic Detection of Changes in Road Databases Using Satellite Imagery. In : IAPRS:GIS-Between Visions and Applications, Vol. 32, pp. 293~298, http://www.isprs.org/proceedings/XXXII/part4/klang62.pdf , checked on 7/05/2012.

Lacroix, V.; Acheroy, M. (1998): Feature extraction using the constrained gradient. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 53, No. 2, pp. 85~94.

Laporte, G.; Mercure, H.; Nobert, Y. (1987): Generalized travelling salesman problem through n sets of nodes: the asymmetrical case. In Discrete Applied Mathematics, Vol. 18, No. 2, pp. 185~197.

Li, L., Goodchild, M.F. (2010): Optimized Feature Matching in Conflation. In: Sixth international conference on Geographic Information Science, checked on 2010.

Li, L.; Goodchild, M.F. (2011): An optimisation model for linear feature matching in geographical data conflation. In International Journal of Image and Data Fusion, Vol. 2, No. 4, pp. 309~328.

Li, H.; Huang, X.; He, L. (2012): Object Matching Using a Locally Affine Invariant and Linear Programming Techniques. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No. 2, pp. 411~424.

Liu, Y. (2004): Developing structural constraints for accurate registration of overlapping range images. In Robotics and Autonomous Systems, Vol. 47, No. 1, pp. 11~30.

Martin, D.R; Fowlkes, C.C; Malik, J. (2004): Learning to detect natural image boundaries using local brightness, color, and texture cues. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 5, pp. 530~549.

Mathias, J.P.M.L.; Koldo, F.V. (1988): Automatically GIS Updating from Classified Satellite Images Using GIS Knowledge. In In International Archives of Photogrammetry and Remote Sensing, Vol. 27, No. 4, pp. 198~206.

Mayer, H. (2008): Object extraction in photogrammetric computer vision. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 63, No. 2, pp. 213~222.

Mena, J. (2003): State of the art on automatic road extraction for GIS update: a novel classification. In Pattern Recognition Letters, Vol. 24, No. 16, pp. 3037~3058.

Mena, J.; Malpica, J. (2005): An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery. In Pattern Recognition Letters, Vol. 26, No. 9, pp. 1201~1220.

Modersitzki, J. (2007): Image Registration with Local Rigidity Constraints. In Bildverarbeitung für die Medizin, pp. 444~448.

Mustière, S.; Devogele, T. (2008): Matching Networks with Different Levels of Detail. In Geoinformatica, Vol. 12, No. 4, pp. 435~453.

Myronenko, A.; Xubo S. (2010): Point Set Registration: Coherent Point Drift. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No.12, pp. 2262~2275.

Negri, M.; Gamba, P.; Lisini, G.; Tupin, F. (2006): Junction-aware extraction and regularization of urban road networks in high-resolution SAR images. In IEEE Transactions on Geoscience and Remote Sensing, Vol. 44, No. 10, pp. 2962~2971.

Papari, G.; Petkov, N. (2011): Edge and line oriented contour detection: State of the art. In Image and Vision Computing, Vol. 29, No. 2-3, pp. 79~103.

Peggy, A.; Anthony, S.; Sotirios, G. (2001): Differential Snakes for Change Detection in Road Segments. In Photogrammetric Engineering and Remote Sensing, Vol. 67, No. 12, pp. 1391~1399.

Peng, T.; Jermyn, I.H.; Prinet, V.; Zerubia, J. (2010): Extended Phase Field Higher-Order Active Contour Models for Networks. In International Journal of Computer Vision, Vol. 88, No. 1, pp. 111~128.

Peteri, R.; Celle, J.; Ranchin, T. (2003): Detection and extraction of road networks from high resolution satellite images. In : Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429).

Petr, F.; Stepan, O. (1998): Straight Skeleton Implementation. In : Proceedings of Spring Conference on Computer Graphics, checked on 4/01/2012.

Pop, P.C.; Kern, W.; Still, G. (2006): A new relaxation method for the generalized minimum spanning tree problem. In European Journal of Operational Research, Vol. 170, No. 3, pp. 900~908.

Poullis, C.; You, S. (2010): Delineation and geometric modeling of road networks. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 65, No. 2, pp. 165~181.

Ravanbakhsh, M.; Heipke, C.; Pakzad, K. (2007): Knowledge-based road junction extraction from high-resolution aerial images. In : Urban Remote Sensing Joint Event: IEEE, pp. 1~8.

Rochery, M.; Jermyn, I.H.; Zerubia, J. (2006): Higher Order Active Contours. In International Journal of Computer Vision, Vol. 69, No. 1, pp. 27~42.

Rottensteiner, F. (2009): Status and further prospects of object extraction from image and laser data. In : 2009 Joint Urban Remote Sensing Event: IEEE, pp. 1~10.

Rueda, S.; Udupa, J.K.; Bai, L. (2008): Local curvature scale: a new concept of shape description. In : Proceedings of SPIE: SPIE, pp. 69144Q.

Ruiz, J.; Rubio, T.; Urena, M. (2011): Automatic extraction of road intersections from images in conflation processes based on texture characterization, Vol. 43, No. 321, pp. 212~225.

Ruiz, J.J.; Ariza, F.J.; Ureña, M.A.; Blázquez, E.B. (2011): Digital map conflation: a review of the process and a proposal for classification. In International Journal of Geographical Information Science, Vol. 25, No. 9, pp. 1439~1466.

Saalfeld, A. (1985): A Fast Rubber-Sheeting Transformation Using Simplicial Coordinates. In Cartography and Geographic Information Science, Vol. 12, No. 2, pp. 169~173.

Saalfeld, A. (1988): Conflation Automated map compilation. In International journal of geographical information systems, Vol. 2, No. 3, pp. 217~228.

Safra, E.; Kanza, Y.; Sagiv, Y.; Doytsher, Y. (2006): Efficient integration of road maps. In : Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems - GIS '06: ACM Press, pp. 59.

Sharp, G.C; Lee, S.W; Wehe, D.K. (2002): ICP registration using invariant features. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, pp. 90~102.

Siddiqi, K.; Kimia, B.B. (1995): Parts of visual form: computational aspects. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 3, pp. 239~251.

Song, W.; Keller, J.M.; Haithcoat, T.L.; Davis, C.H. (2009): Automated Geospatial Conflation of Vector Road Maps to High Resolution Imagery. In IEEE Transactions on Image Processing, Vol. 18, No. 2, pp. 388~400.

Steger, C. (1998): An unbiased detector of curvilinear structures. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 2, pp. 113~125.

Stilla, U. (1995): Map-aided structural analysis of aerial images. ISPRS Journal of photogrammetry and remote sensing, Vol. 50, No. 4, pp. 3-10.

Stilla, U., Michaelsen, E., Soergel, U., Hinz, S., Ender, H.J. (2004): Airborne monitoring of vehicle activity in urban areas. In: Altan MO (ed) International Archives of Photogrammetry and Remote Sensing. Vol 35, Part B3, pp. 973-979.

Tang, Y.Y.; You, X. (2003): Skeletonization of ribbon-like shapes based on a new wavelet function. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 9, pp. 1118~1133.

van Cleynenbreugel, J.; Fierens, F.; Suetens, P.; Oosterlinck, A. (1990): Delineating road structures on satellite imagery by a GIS-guided technique. In Photogrammetric Engineering and Remote Sensing, Vol. 56, pp. 893~898.

Veltkamp, R.C. (2001): Shape matching: similarity measures and algorithms. In: Proceedings International Conference on Shape Modeling and Applications: IEEE Computer Society, pp. 188~197.

Volker, W. (1996): Zuordnung von raumbezogenen Daten - am Beispiel der Datenmodelle ATKIS und GDF. Universität Stuttgart, Stuttgart. Bauingenieur- und Vermessungswesen.

Waldo, R.T. (1970): A computer movie simulating urban growth in the detroit region. In Economic Geography, Vol. 46, No. 2, pp. 234~240.

Wen, G. (2006): A Global Algorithm for Straight Line Stereo Matching Based on Feature Grouping. In Journal of Software, Vol. 17, No.12, pp. 2471.

Wiedemann, C.; Hinz, S. (2000): Automatic extraction and evaluation of road networks from satellite imagery. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIII, pp. 980~986.

Wolsey, L.A. (1998): Integer programming. New York: Wiley.

Woods, R.P., Grafton, S.T.; Watson, J.D.G.; Sicotte, N.L.; Mazziotta, J.C. (1998): Automated image registration: II. Intersubject validation of linear and nonlinear models. In Journal of Computer Assisted Tomography, Vol. 22, pp. 153~165.

Wu, X.; Rodrigo, C.; Hui, F.; Steve, Z.; Andrew, K. (2007): Automatic alignment of large-scale aerial rasters to road-maps. In : Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems - GIS '07: ACM Press.

Xinquan, S.; Palmer, P. (2000): Uncertainty propagation and the matching of junctions as feature groupings. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 12, pp. 1381~1395.

Xiong, D.; Sperling, J. (2004): Semiautomated matching for network database integration. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 59, No. 1-2, pp. 35~46.

Yoo, T.S. (2004): Insight into images. Principles and practice for segmentation, registration, and image analysis. Wellesley, Mass: A K Peters.

Zagorchev, L.; Goshtasby, A. (2006): A comparative study of transformation functions for nonrigid image registration. In IEEE Transactions on Image Processing, Vol. 15, No. 3, pp. 529~538.

Zhang, M.; Meng, L. (2006): Implementation of a generic road-matching approach for the integration of postal data. In : Proceedings of the 1st ICA Workshop on Geospatial Analysis and Modeling, pp. 141~154. Available online at http://129.187.175.5/lfkwebsite/fileadmin/user_upload/publications/Zhang/Zhang_Meng20060316.pdf, checked on 7/05/2012.

Zhang, M. (2009): Methods and Implementations of Road-Network Matching. Technische Universität München, München. Institut für Photogrammetrie und Kartographie.

Zhang, C. (2004): Towards an operational system for automated updating of road databases by integration of imagery and geodata. In ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No. 3-4, pp. 166~186.

Zhang, C. (2003): Updating of Cartographic Road Databases by Image Analysis. Institute of Geodesy and Photogrammetry, Zürich.

Zhang, J.; Zhu, Y.; Krisp, J.; Meng, L. (2010): Derivation of road network from land parcels. In : Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '10: ACM Press, pp. 418.

Zhang, J.; Zhu, Y.; Meng, L. (2011): Conflation of road network and geo-referenced image using sparse matching. In : Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '11: ACM Press, pp. 281.

Abbreviations

Coherent Point Drift (CPD)

Computer Tomography (CT)

Congruent Hybrid Model (CHM)

Constrained Minimum Perfect Matching Problem (C-MP^2)

Digital Elevation Model (DEM)

Digital Surface Model (DSM)

Elastic Circular Mask (ECM)

Generalized Minimum Edge-Biconnected Network Problem (GMEBCNP)

Generalized Travelling Salesman Problem (GTSP)

Genetic Algorithm (GA)

Geographic Information Systems (GIS)

Iterative Closest Point (ICP)

Light Detection and Ranging (LIDAR)

Medial Axes (MA)

Normalized Difference Vegetation Index (NDVI)

Nuclear Magnetic Resonance (NMR)

Remote Sensing (RS)

Rubber-Sheeting (RUBS)

Shape Decomposition Algorithm (SDA)

Skeleton Pruning Algorithm (SPA)

Sparse Matching Algorithm (SMA)

Straight Skeleton (SS)

Sum-of-Squared-Differences (SSD)

Synthetic Aperture Radar (SAR)

Thin Plate Spline (TPS)

Thin Plate Spline Robust Point Matching (TPS-PRM)

Topology Reconstruction Algorithm (TRA)

Triangulation Based Skeleton (TBS)

Very High Resolution (VHR)

Acknowledgements

I got an exclusive research time between March 2009 and October 2012 to conduct this work at the Department of Cartography, Technische Universität München, Germany. I am very grateful to the support and encourage from my supervisors, my colleagues, my family and my friends. Without their support, this dissertation would not have been successful.

I would like to express my sincere gratitude and appreciation to my supervisor, Prof. Dr.-Ing. Liqiu Meng, who not only provided me with the opportunity to work within an excellent team, but also allowed me to freely choose my favourite research topic. From the fruitful discussions, my elementary idea was pruned into a scientific question with a better theoretical formulation; furthermore, her profound knowledge and patience coached me from one success to another. Besides the professional knowledge, I was also deeply influenced by her rigorous ethics, which inspired me always in pursuing the best quality in my research. No doubt I will benefit a lot from the experiences in future.

My gratitude is also devoted to my co-supervisor, Prof. Dr.-Ing. Dieter Fritsch, for reviewing my work and giving me valuable comments.

My special thanks and respects also belong to all of my colleagues at the Department of Cartography. The friendly and pleasant atmosphere has made my research work a joyful undertaking. The open discussions in the internal seminars or at the coffee breaks helped me a lot to get more fresh ideas. Moreover, I got the feeling of being at home due to their readiness for help. For instance, Mr. Fritz Meier changed the comfortable chair and table for me; Mrs. Luise Fleißner always reminded me to fill the application form before each business travel; she also helped me to correct the grammatical mistakes in my German. I cannot imagine my life in Munich without their supports.

I very much appreciate the endless support and understanding from my family. My parents agreed with me and supported me to study abroad. I also greatly thank my wife Dr.-Ing. Yueqin Zhu for her patience and love, especially when she also did her PhD study at the same time. To my lovely daughter, Jiayi Zhang, you are so cute and so pretty, and thank you for reminding me to keep my research schedule because you otherwise preferred to play with me after each working day. I'm so sorry for having left you too little time before you went to bed every day, but I really hope it is not the reason that you cannot lose your baby fat.

My research work was partially supported by United Maps Company, and I would like to thank Mr. Andreas Wiedmann, Mr. Carsten Recknagel, and Mr. Stefan Knecht for their novel suggestions and timely providing me with the orthoimage and road network in the test area.

Last but not the least, I want to thank all my friends in Germany and China, and my life has become colourful because of the activities and the conversations we have been involved.

Curriculum Vitae

Personal Data

Name	Jiantong Zhang
Date of Birth	October 29, 1977
Nationality	Chinese

Education

03.2009~	PhD student
	Department of Cartography, Technische Universität München
10.2006~10.2008	Master of Science
	Department of Cartography, Technische Universität München
09.1999~07.2002	Master of Science
	Wuhan University & Technology, Wuhan, China
09.1995~07.1999	Bachelor of Engineering
	Taiyuan University & Technology, Taiyuan, China
09.1992~07.1995	High School
	High school of Huayou, Renqiu, Hebei, China
09.1989~07.1992	Middle School
	Dongfeng school of Huayou, Renqiu, Hebei, China
09.1983~07.1989	Primary School
	Dongfeng school of Huayou, Renqiu, Hebei, China

Experiences

03.2009~	Scientific Assistant
	Department of Cartography, Technische Universität München
07.2002~10.2006	Manager of software development department
	National Intelligent Transport Systems (ITS) Center, China