TECHNISCHE UNIVERSITÄT MÜNCHEN
Institut für Photogrammetrie und Kartographie
Fachgebiet Photogrammetrie und Fernerkundung

Extraction and Velocity Estimation of Vehicles in Urban Areas from Airborne Laserscanning Data

Wei Yao

Dissertation

2010

TECHNISCHE UNIVERSITÄT MÜNCHEN
Institut für Photogrammetrie und Kartographie
Fachgebiet Photogrammetrie und Fernerkundung

Extraction and Velocity Estimation of Vehicles in Urban Areas from Airborne Laserscanning Data

Wei Yao

Vollständiger Abdruck der von der Fakultät für Bauingenieur- und Vermessungswesen der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr.-Ing. Liqiu Meng

Prüfer der Dissertation:

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

2. Univ.-Prof. Dr.-Ing. habil. Stefan Hinz
   Universität Karlsruhe (TH)

Die Dissertation wurde am 14.10.2010 bei der Technischen Universität München eingereicht und durch die Fakultät für Bauingenieur- und Vermessungswesen am 29.11.2010 angenommen.

Abstract

In this work a two-step strategy for traffic monitoring in urban areas by analysis of single-pass airborne laser scanning (ALS) data is presented and investigated. In the first step vehicles are extracted, and their states of motion are analyzed in the following step.

For vehicle extraction, two methods are proposed. For the first method it is assumed that all road sections in the studied scene are component of the ground surface. The laser data are transformed from the point cloud into a grid representation. Based on analysis of the height distribution the ground surface included the vehicles is first separated from other objects such as buildings and vegetation by an iterative process. Then, a morphological segmentation is carried out to isolate the cars from the ground surface. The second method assumes that road sections are also located on bridges or overpasses. Through an adaptive "mean shift" a 3D segmentation of point clouds is performed. Based on local structure, point clouds, which could represent the vehicles, are immediately separated from all other objects. The distinction between vehicle and background is performed by a classification using a support vector machine. In scenes with dense placement of vehicles such as occurring in parking lots, a grouping of larger objects by "normalized cuts" is further conducted to enable a combination with the first method.

In the step of motion analysis, based on the extracted point clouds of the vehicles a motion state is initially determined and subsequently for those classified as moving vehicles the velocity is estimated. To determine the motion state the shape of the vehicle point cloud is fitted by a parallelogram and classified on the basis of the parameters of the aspect ratio and shearing angle. The classification consists in a binary decision made by the evaluation with Lie Group metric. Finally, the velocity of moving vehicles is estimated based on the deformation structures. Fundamentally the moving direction of vehicles indicated by the road orientation can be considered as prior knowledge in the velocity estimation. With this information, three methods have been analyzed to determine the velocity. When it is lack of information about the road orientation, the velocity and direction can be simultaneously determined by solving a system of linear equations.

The approaches were analyzed by four laser datasets of three different cities. For evaluation of the detection results reference data have been created manually. To evaluate the motion estimation video sequences concurrently recorded for two scenes were examined. The results of both methods for vehicle extraction have shown that a high completeness (up to 87%) of the detection of vehicle objects is reached by using the first method while the second method provides a high accuracy with respect to the vehicle geometry. Moreover, for the motion detection the recognition was investigated in dependence on the point density, intersection angle and vehicle velocity by simulated point clouds. Studies on the accuracy of velocity estimation show a strong dependence on the ratio of flight velocity to vehicle velocity and the intersection angle in between. The best estimate from the experiments show a deviation of about 10% compared to the video sequences.

Kurzfassung

In dieser Arbeit wird eine zweistufige Strategie zur Verkehrsüberwachung in urbanen Bereichen durch Auswertung von Single-Pass Airborne Laserscanning (ALS)-Daten vorgestellt und untersucht. Dabei werden in der ersten Stufe zunächst die Fahrzeuge extrahiert und in der folgenden Stufe deren Bewegungszustand analysiert.

Für die Fahrzeugextraktion werden zwei Methoden vorgeschlagen. Bei der ersten Methode wird davon ausgegangen, dass alle Straßenabschnitte in der untersuchten Szene Bestandteil der Bodenfläche sind. Die Laserdaten werden von der Punktwolke in eine Rasterdarstellung gewandelt. Basierend auf Analyse der Höhenverteilung wird zunächst durch ein iteratives Verfahren die Bodenfläche inklusiv der Fahrzeuge von anderen Objekten wie Gebäuden und Vegetation separiert. Anschließend erfolgt eine morphologische Segmentierung um die Fahrzeuge von der Bodenfläche zu isolieren. Bei der zweiten Methode wird davon ausgegangen, dass Straßenabschnitte auch auf Brücken oder Überführungen liegen. Durch einen adaptiven „Mean Shift"-Ansatz wird eine 3D-Segmentierung der Punktwolken durchgeführt. Dabei werden auf Basis der lokalen Struktur Punktwolken, die Fahrzeuge darstellen könnten, direkt von allen anderen Objekten getrennt. Die Unterscheidung von Fahrzeug und Hintergrund erfolgt durch eine Klassifikation mit einer Support Vektor Maschine. Bei Szenen mit dichten Fahrzeuganordnungen wie sie bei Parkplätzen auftreten, wird weiterhin eine Gruppierung größerer Objekte durch „Normalized Cuts" durchgeführt und eine Kombination mit der ersten Methode angewendet.

Bei der Bewegungsanalyse wird basierend auf den extrahierten Punktwolken der Fahrzeuge zunächst ein Bewegungsstatus bestimmt und bei den als bewegt klassifizierten Fahrzeugen nachfolgend die Geschwindigkeit geschätzt. Zur Bestimmung des Bewegungsstatus wird die Form der Fahrzeugpunktwolke durch ein Parallelogramm approximiert und aufgrund der Parameter aus Längen/Breitenverhältnis und Scherwinkel klassifiziert. Die Klassifikation besteht in einer Binärentscheidung die durch Auswertung mit einer Lie Group Metrik erfolgt. Schließlich wird die Geschwindigkeit der bewegenden Fahrzeuge auf Grundlage der Deformationsstruktur bestimmt. Prinzipiell kann bei dieser Schätzung die Bewegungsrichtung aus der Straßenanordnung als Vorkenntnis berücksichtigt werden. Mit dieser Information wurden drei Ansätze zur Geschwindigkeitsbestimmung untersucht. Liegen keine Information zur Straßenausrichtung vor, werden Geschwindigkeit und Richtung durch Lösung eines linearen Gleichungssystems ermittelt.

Die Ansätze wurden mit vier Laserdatensätzen von drei verschiedenen Städten untersucht. Für die Bewertung der Detektionsergebnisse wurden Referenzdaten manuell erstellt. Um die Schätzung der Bewegung zu bewerten, wurden die bei zwei Szenen gleichzeitig aufgenommenen Videosequenzen ausgewertet. Die Ergebnisse der beiden Methoden zur Fahrzeugextraktion haben gezeigt, dass bei der ersten Methode eine hohe Vollständigkeit (bis 87%) bezüglich der Erkennung von Fahrzeugobjekten erreicht wird, während das zweite Verfahren eine hohe Genauigkeit bezüglich der Fahrzeuggeometrie liefert. Für die Bewegungsdetektion wurde durch simulierte Punktwolken die Erkennung in Abhängigkeit der Punktdichte, dem Beobachtungswinkel und der Geschwindigkeit untersucht. Untersuchungen zur Genauigkeit der Geschwindigkeitsschätzung zeigen eine starke Abhängigkeit von dem Verhältnis der Fluggeschwindigkeit zur Fahrzeuggeschwindigkeit und dem Beobachtungswinkel. Die besten Schätzungen aus den Experimenten zeigen eine Abweichung der Geschwindigkeit von ungefähr 10% im Vergleich zu den Geschwindigkeitsschätzungen aus den Videosequenzen.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Transportation is one of indispensable pillars of modern society and economy. Traffic, however, is also a source of environmental pollution and a cause of injuries and deaths. Smart traffic management attempts to optimize vehicle throughput while still mitigating these adverse effects. And, traffic keeps on changing on short or long time scales. Road construction, major events, or disasters require rapid reaction and response. Political developments, like the expansion of the European Union have significantly changed the traffic scenarios in the member states. Therefore, data and models for traffic management have to be continuously updated. Today's road systems are equipped with a set of sensors for monitoring traffic status: induction loops, overhead radar sensors, video camera systems are the most prominent examples. They all deliver accurate, reliable, timely, yet merely point-wise ground-based measurements. Remote sensing systems on the other hand give us synoptic and extended views of complex traffic situations and the associated context (Hinz et al., 2006,). These data are complementary to the ones of the road sensors and can be used in research for improving traffic models and complement the accurate but sparsely sampled measurements of local fixed/floating sensors. The means of data acquisition via mobile remote sensing platforms can be rapidly deployed as needed and it does not interfere with the traffic on the road.

With the recent advances in sensor technology, the automatic detection, characterization and monitoring of traffic using remotely sensed data from airborne/spaceborne platforms has become an emerging field of research (Stilla et al., 2005; Hinz et al., 2006; Stilla et al., 2009). Approaches for vehicle detection and motion estimation include not only video cameras but nearly the whole range of available sensors such as optical aerial and satellite sensors, infrared thermal cameras, Synthetic Aperture Radar (SAR) systems, and Airborne Laser Scanning (ALS), also referred to as airborne LIDAR. Although imaging sensors including optical cameras and SAR systems are already in use and seem to be an obvious and intuitive choice to provide data sources, airborne LiDAR systems have entered the sub-metric resolution era required for the detection and characterization of small targets such as vehicles. Airborne LiDAR is a young technology, which even can work in the night time. Yet there are also inherent difficulties in the ALS scanning process that must be overcome to design a reasonably reliable approach for traffic monitoring.

LiDAR systems first came into operation at the beginning of 1990s. By virtue of the accurate and direct sensor orientation using GPS and INS, it allows to acquire detailed 3D data of the nature environment and man-made objects actively by emitting and receiving near infrared pulses. Apart from information about the geometric structure of the scanned earth's surface, laser scanners can also record the pulse intensity of backscattered echoes, which represents physical properties of object's surface. Consequently, laserscanning has become well established surveying techniques of the acquisition of geospatial information, as also summarized by several books (Vosselman and Maas, 2010; Shan and Toth, 2008).

In last decade, major advances have been made in technological capabilities of commercial ALS

systems, such as sampling density, footprint size, multiple pulses reflection, and positional accuracy. Nowadays, the high quality of obtained 3D point clouds are routinely used for a diversity series of purposes including the production of digital terrain models (Chen 2010), 3D city models (Sohn and Dowman, 2007; Sampath and Shan, 2007), forestry management and monitoring (Reitberger et al., 2009; Rutzinger et al., 2008), corridor mapping (Alexander et al., 2010), and documentation of cultural heritage (Doneus et al., 2008). Usually, such applications build upon preprocessed rather than raw ALS data, which requires the raw point clods to undergo the physical (Kaasalainen et al., 2009) and geometric calibration (Skaloud and Lichti, 2008), sometimes also the registration (Bae and Lichti, 2008) operation due to multiple datasets acquired at differ time or from different perspectives.

A key feature in current generation ALS systems is the possibility for digitizing and recording of the full-waveform of the echo signal of reflected laser pulses. Such so-called full-waveform ALS systems are helpful to extract and interpret quasi 3D vertical structure information of natural and artificial objects that are illuminated as volume-scattering objects, just like forested terrain, and building boundary. Since its first introduction in spaceborne platforms (Blair et al., 1999), the full waveform system opens up new possibilities for a more detailed description of object structures and more accurate range measurements by incorporating multiple intermediate pulse reflections and new physical feature of illumined surfaces such as pulse form and width (Lin and Mills, 2010). Moreover, according to Jutzi and Stilla (2006), capturing the complete waveform of laser pulses allows discriminating differences in a range smaller than the length of a laser pulse. It is especially useful for such critical measurement situations that surfaces are close to their edges or small objects within the beam footprint are closely located in range. Yao and Stilla (2010) has recently proposed a new method of pulse detection by exploiting mutually enhanced information hiding in a group of LiDAR waveforms to recover or enrich 3D point clouds of severely occluded object surfaces.

Recently, traffic monitoring applications using airborne LiDAR have received an increasing attention in our community. Toth and Grejner-Brzezinska (2006) initiated a research work which adopted an airborne laser scanner coupled with a digital frame camera to analyze transportation corridors and acquire traffic flow information. The traffic flow information is successfully obtained alongside with the topographic spatial data as an affiliated product. The testing of this system is limited to a section of motorway and its narrow surroundings; it is interesting to investigate the same problem in more challenging and broader regions using the system equipped solely with LiDAR sensor. Toth and Grejner-Brzezinska (2005) have compared airborne LiDAR with optical frame camera in view of the motion estimation task and suggested to ideally use a combined dataset, namely the fusion of simultaneously acquired imagery and laser data could produce better results. But, there is still the need to deepen the understanding concerning the capability and performance of the usage of airborne LiDAR to carry out the task of motion indication and estimation.

Moreover, the research work dealing with the traffic analysis by airborne LiDAR in urban areas presented in this thesis are particularly motivated by following points:

- the penetration ability of laser rays towards volume-scattering objects allows to improve the completeness of vehicle detection(Chang et al., 2010, Chevalier et al., 2007), see vehicles in red ellipses of Figure 1 as example

- the motion artifacts generated (Figure 2) by the linear scanning mechanism used in airborne LiDAR allows to determine the object motion
- the explicit extraction and modeling of vehicles can refine the result of operations such as DTM filtering and road detection where vehicles are regarded as stubborn disturbances. (Sithole and Vosselman, 2005; Clodea et al., 2007)
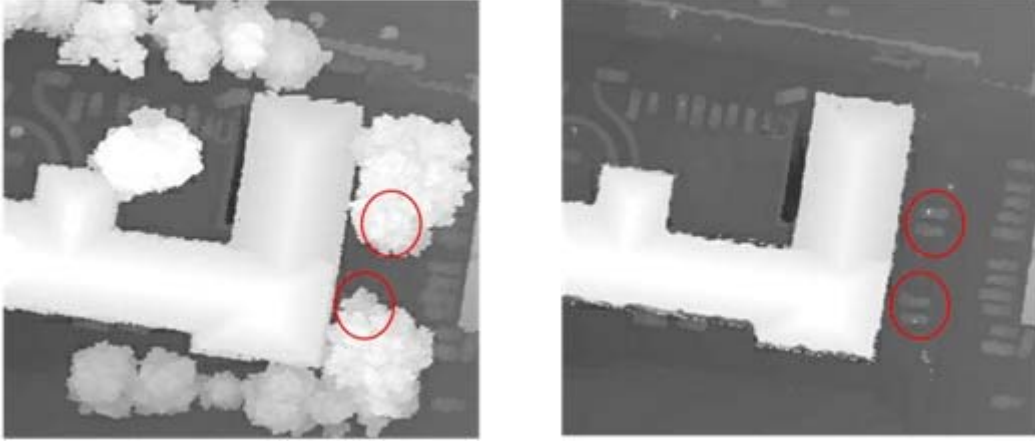
Figure 1: Airborne LiDAR range image illustrating the penetration ability of laser pulses through tree crowns to hit vehicles beneath, (a) first-pulse range image, (b) last-pulse range image (Stilla 2003)
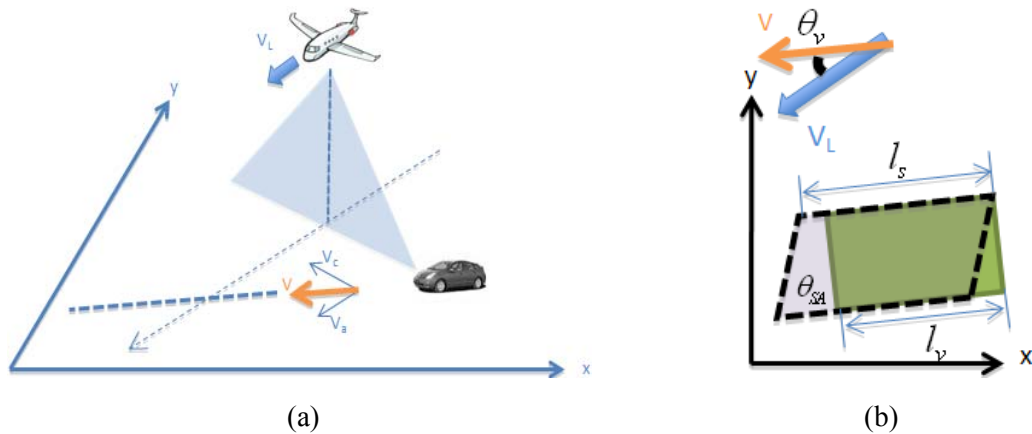
Figure 2: Moving objects undergo the scanning process of airborne LiDAR, (a) a scenario of moving object captured by ALS, (b) generated motion artifact effects, green rectangle indicates the original shape and position of the vehicle

1.2 Related works concerning traffic monitoring from ALS data

Automatic traffic monitoring has evolved to an important and active research issue in the remote sensing community during the past years (Stilla et al., 2004; Hinz et al., 2006; Hinz et al., 2008). Transportation represents a major segment of the economic activities of modern societies and has been keeping increase worldwide which leads to adverse impact on our environment and society, so that the increase of transport safety and efficiency, as well as the reduction of air and noise pollution are the main task to solve in the future. On the one hand, today's traffic monitoring systems are mainly equipped by a series of sensors like induction loops, overhead radar sensors and stationary video cameras, etc. They all deliver accurate, reliable, timely, yet merely point-wise measurement.

On the other hand spaceborne and airborne sensors can complement the sparsely sampled ground-based collection and give us a synoptic view of complex traffic situations. With the recent advances in sensor technology, a number of approaches for automatically detecting vehicles, tracking vehicles and estimating velocity have been developed and intensively analyzed, using different air-and spaceborne remote sensing platforms (Lenhart et al., 2008), e.g. Synthetic aperture radar (SAR) (Meyer et al., 2006, Suchandt et al., 2010), thermal infrared(IR) cameras (Stilla and Michaelsen, 2002; Hinz and Stilla, 2006; Kirchhof and Stilla, 2006; Yao et al., 2009), frame and linear pushbroom optical cameras (Zhao and Nevatia, 2003; Eikvil et al., 2009; Larsen et al., 2009; Leitloff et al., 2010). However, so far there have been few works conducted in relation to traffic analysis from laser scanners due to the special acquisition mechanism and data characteristics. The traffic monitoring remains as one of few fields which are still not intensively analyzed in the LiDAR community compared to other modern sensors used in remote sensing. The relevant works concerning traffic analysis from LiDAR data can be generalized into two aspects:

1) Vehicle (target) extraction

It is related to one of important research fields in laser data processing - 3D object recognition, which is primarily dedicated to Automatic Target Recognition (ATR) in military applications (e.g. Grönwall et al., 2007; Steinvall et al., 2004; Grönwall, 2006; Ahlberg et al., 2003). The scene can be scanned from different platforms and perspectives, such as terrestrial or airborne platforms. The point which significantly distinguishes the use of laser sensor for urban traffic analysis from for military applications lies in data coverage and the application objective. The military applications often feature small field of view (FOV) and very high-resolution (very high density of laser points) with respect to data recording. The data acquisition process is target-orientated and limited to a relative small coverage, the interest region or object is scanned with very high resolution and concentrated energy. Most of algorithms developed within this scope aim at recognition of the object type (e.g. classification of military vehicles) and pose estimation (e.g. orientation of tank); someone even tried to detect fine sub-structures of object (e.g. barrel and turret of a tank). Among these algorithms, model-based shape matching or fitting strategies have been most frequently applied to the LiDAR data in order to find and recognize the corresponding object class and its state (Koksal et al., 1999; Zheng and Der, 2001; Johansson and Moe, 2005). More recently, Nieves and Reynolds, (2010) have proposed an efficient 3D spatial transformation that preserves the geometrical attributes of the LIDAR data in all its dimensions. This transformation permits the utilization of well established statistical and shape-based descriptors for the implementation of automatic target recognition. Grönwall et al, (2010) have utilized geometric features, shadow analysis and height-based detection in flash LiDAR data to detect partly occluded objects such as vehicles. It has been found that the range information is valuable for detection of small objects that are typically represented by 5-10 pixels in the dataset. Range information is also valuable in tracking problems when tracked objects are occluded under parts of its movement even if there are several objects in the scene.

2) Motion indication and estimation

The most relevant and up-to-date research is, to the best of our knowledge, contributed from Toth & Grejner-Brzezinska (2006) and Grejner-Brzezinska et al., (2004). In the work an airborne laser scanner coupled with digital frame imaging sensor was adopted to analyze transportation corridors

and acquire traffic flow information automatically. They have tried to extract traffic-related static and dynamical information as part of the regular topographic mapping. Vehicle velocity can be estimated either by analyzing motion artefacts in LiDAR data or by vehicle tracking in image sequences acquired with reasonable acquisition rate. The experiences gained so far by their test flying campaigns showed that the two sensors have different strength and weakness for various data processing tasks and, in most cases, they complement each other. It is declared that the combination of airborne LiDAR and imaging sensors can provide valuable traffic flow data that can effectively support traffic monitoring and management. However, the extensive testing of this system is limited to highway, freeway and other heavily travelled roads where occlusions cast by buildings, vegetations and some other anomalies (e.g. guild rails) are rare. In addition to directly related research topics devoted to traffic monitoring, there are some indirectly related research works reported in recent years, which deal with moving object detection and tracking using LiDAR data. For instance, Zhao and Shibasaki, (2005) presented a system for tracking pedestrians in a wide and open area. A simplified pedestrian's walking model based on the typical appearance of moving feet are defined and a tracker utilizing the Kalman filter is used to track pedestrian's trajectories and analyze the movement behavior pattern. Szarvas et al, (2006) used LiDAR-based detector and a convolutional neural network-based classifier to detect moving targets in real time. LIDAR-based detector can reduce the number of false positives by a factor of 2 and reduce the processing time by a factor of 4 thanks to its ability to automatically learn a small number of highly discriminating features. Mahlisch et al., (2006) proposed to incorporate multiple sensor data into detection and tracking framework for real-time vehicle detection and motion indication. It is intended for emergency breaking or ACC applications, which can benefit from the low level fusion and alignment of LiDAR and vision sensor measurements in discrimination performance and computational complexity.

The existent approaches deal with the vehicle (target) tracking and velocity estimation process, but often manual work or the initial position of target is required to assist the processing or enable the tracking process in imaging data to work. The motion detection and estimation method used by these approaches are mostly built upon the combination of LiDAR and other data sources or only allow the motion analysis at single-object level. For urban traffic monitoring, in order to ensure the system efficiency and derive accurate traffic flow information, a much broader area is needed to be covered by laser scanner surveying and multiple instances of the vehicle object and their motion states have to be located and recognized simultaneously. It requires more advanced algorithms to separate 3D point sets of vehicles from complex surroundings and a new moving object model to allow extracting object motion directly and accurately from ALS data. Under this consideration, the operations used for pose estimation or geometric inference are not crucial as semantic decision of whether a vehicle exists or not (vehicle counting). Moreover, it must be examined that how the newly proposed method of motion detection and estimation can capture and exploit the essential shape features of moving objects in LiDAR data to enable and even improve the results regarding the traffic monitoring.

1.3 Goals of the thesis

The main goal of the thesis is the development of a framework for the detection and estimation of vehicle motion from single-path airborne LiDAR data of urban areas. It enables not only the extraction of as complete vehicles as possible but also the distinction and estimation of their motion

state and the velocity of classified moving vehicles. Potential applications are the traffic monitoring using airborne LiDAR, even if the flight campaign is originally designed to accomplish the task of city mapping and modeling, and the quantitative and qualitative modeling and analysis of the impact of the transportation system on the urban environment.

The research work in this thesis has been carried out with following sub-goals:

- derivation of the mathematic model for motion artifacts effect in single-path airborne LiDAR data. The Model establishes the analytic functional relations between shape deformation parameters and velocity of moving objects to enable the qualitative and quantitative determination of object motion.

- development of a hybrid vehicle extraction strategy from ALS data which combines a 3D point cloud analysis method and a gridded data based method to achieve results. The 3D analysis method runs directly on 3D data space of LiDAR so as to avoid the unreliable determination of geometric features depending on local neighborhood selection and can cope with the undesirable complexity of urban scenes. The gridded data based method is designed to aim at the utilization of advanced methods from image processing and the detailed vehicle (context) model to attain the purpose. The control mechanism of the overall strategy can be switched between these two methods depending on the global context property of test scenes in order to optimize vehicle extraction results by fusion.

- investigation and development of a classification framework of vehicle motion state based on parameterizing the shape deformations described in the motion artifacts model. The classifier should cope with the nonlinear variability of vehicles shapes represented in 3D point cloud, the high sensitivity to noise and uncertainty induced by the vehicle extraction as well as the parameterization process.

- build up a velocity estimator for classified moving vehicles which contains different estimation methods derived based on analyzing respective shape deformations by solving the inverse problem of the motion artifacts model. The derivation of the theoretic accuracy of velocity estimates is investigated based on applying the error propagation principle to the functional model for deriving the velocity. It is needed to examine a criterion based on which the best estimation method is automatically selected for certain road sections in accord with conditions of flight campaign and indications of theoretic accuracy analysis.

Consequently, the presented work in this thesis mainly deals with general methodological research and performance analysis of urban traffic monitoring based on analyzing airborne LiDAR data. By defining the moving model in ALS data analytically, the motion detection and estimation of traffic related objects becomes feasible and applicable. This work tends to demonstrate a possibility to improve the result of vehicle extraction and recover those vehicles that are partially occluded by tree leaves in point cloud and cannot be detected by optical sensors (camera). The potential information about traffic scenarios are expected to be gained through the proposed algorithm alongside with other common products for topographic mapping, even though configurations of ALS data acquisition are designed far from the optimization of traffic monitoring applications.

1.4 Structure of the thesis

The thesis is structured as follows:

After this introduction, Chapter 2 explains the basics of vehicle and context models in short to provide the fundamentals required for the development of a vehicle extraction method. Then the basics and essential characteristics concerning Lie group theory and the principle geodesics analysis is introduced to outline the framework of 3D shape description and classification towards motion detection.

Chapter 3 describes the complete strategy for monitoring urban traffic from airborne LiDAR in detail. At first, it gives an overview of the proposed strategy in which three modules corresponding to three different sub-tasks in the strategy are described while three routines towards vehicle extraction are defined. Then, the scheme of vehicle extraction is presented in which two extraction methods are developed and described. The first method builds upon gridded data and makes use of local contexts and the vehicle model to concentrate the extraction on the ground surface and potential objects. The second method is performed based on an adaptive 3D segmentation based classification method which runs directly on raw point clouds. The global context characteristics of data are analyzed to determine the control mechanism of the module i.e. when and where these two methods are to be called respectively or even combined to perform the extraction. Afterwards, the mathematic model of the motion artifacts in airborne LiDAR data is defined and presented. Based on single vehicle points extracted by the last step, the development of a classification method of vehicle motion state by 3D shape categorization is described. Finally, for obtained moving vehicles, methods for velocity estimation with theoretic accuracy analysis are given where the decision to select the proper estimation methods is made according to data acquisition conditions and the numerical performance prediction.

In Chapter 4 the used recording systems, flight campaigns and characteristics of the various ALS datasets are summarized. The method used to evaluate the results is introduced. Four different datasets are selected to undergo the experiment. Apart from reference data for the verification of vehicle extraction results for all the datasets, for certain datasets there also exist reference data for motion detection and velocity estimation which were provided by video sequences data acquired during the same flights.

Chapter 5 is to apply the strategy proposed in Chapters 3 to the data of Chapter 4 to examine the effectiveness of these methods for traffic analysis from airborne LiDAR platform. The results obtained by performing experiments on the datasets are presented and evaluated. The parameters and thresholds need to be determined adaptively for every intermediate step, which are empirically determined by training data chosen from datasets of the same flight campaigns. The training and evaluation processes are based on objects where the interpretation was made unambiguously with the help of both geometric and spatial context information. For the detection of vehicle motion, the training data samples for moving and stationary vehicle classes are collected by areas chosen from the same flight campaigns as positive and negative samples, respectively. The velocity estimation method by combining two motion components is selected due to its reliability and generality. The results of three steps within the whole strategy – vehicle extraction, motion detection and velocity estimation are assessed by external evaluation whenever possible.

Chapter 6 deals with the discussion of results of the performed experiments to analyze and predict the performance of the strategy. To this end, results obtained by two vehicle extraction methods is first assessed and compared to provide useful comments on the influence of the vehicle extraction method on motion analysis. Then, the performance of the designed motion detector is analyzed in both analytic and experimental manners based on simulations. Finally, the accuracy prediction for velocity estimation on real-life road networks is conducted by analytical calculations based on road maps and LiDAR flight configurations.

In Chapter 7 the developed strategy of traffic monitoring using airborne LiDAR data is discussed in details to derive concluding remarks. Finally, it makes an outlook on open problems and research directions in future.

2 **Basics**

In this chapter the basic knowledge concerning vehicle extraction and motion estimation from airborne LiDAR data is given to provide the fundaments required for the developments of the complete strategy for traffic monitoring presented later in this thesis. First, the vehicle model in airborne LiDAR data are constructed to formalize and represent available knowledge about the object structure so as to reconstruct semantics and properties of the object using the stored data values solely based on algorithms. Additionally, it is necessary to incorporate the knowledge about the influence of surrounding objects and their appearance in the model, particularly for the treatment of complex scenes. This is done by a context model. Second, a nonlinear 3D shape classification framework based on the Lie group manifold is introduced and described. The mathematical basics and structure of Lie group theory and principal geodesics analysis are discussed in view of shape description and classification.

2.1 Modeling

2.1.1 Vehicle model

The modeling of vehicles plays a central role in the extraction of traffic-related objects, regardless of the strategies and the data sources adopted. Research works on vehicle detection using the imaging sensors, such as optical camera, IR camera or SAR, usually are distinguished based on the underlying type of modeling. There are generally two types of vehicle models – appearance-based implicit model and explicit model. The latter one is, for instance, represented by 2D or 3D represented by a filter or wire-frame (Hinz, 2004). For the purpose of better understanding of the sensor scanning mechanism und data characteristics, it is assumed that the vehicle modeling is equally required for vehicle detection in the context of traffic monitoring from ALS systems, although the consistent object modeling in LiDAR data is very difficult.

Here the vehicle model refers to summarizing all the vehicle instances appearing on the ground and road in the scene including parking vehicles, temporarily motionless ones (mainly cars in urban areas) and moving vehicles, which comprise an important vehicle category in terms of the motion state for deriving traffic information. The typical object model usually compiles knowledge about geometric, radiometric, and topological characteristics. To be in accordance with a diversity of motivations vehicle detection methods developed so far are very heterogeneous in their models and strategies (Bogenberger et al., 1999). Vehicles are geometrically modeled in remotely sensed imagery (Tan et al., 1998; Michaelsen et al., 1998; Haag and Nagel, 1999, Michaelsen and Stilla, 2000; Zhao and Nevatia, 2001) as 3D objects using a wire-frame representation where their substructures, such as windshield, hood, roof, etc, are included and described. The level of detail of such a description usually has the consequence that a large number of models are required to model different vehicle types and sizes. Sullivan et al. (1995) shows, however, by an eigenvalue-based investigation of geometric variations of car models that in the case of typical aerial recognition, a large majority of passenger cars can be described with relatively few models (e.g., four in Zhao and

Nevatia, (2001)). Michaelsen and Stilla (2002) also reported a vehicle recognition system using articulated polyhedral models which is implemented in a data-driven voting scheme to search edge line in the image. Flexible or hierarchical models as shown in (Dubuisson-Jolly et al., 1996; Olson et al., 1996; Hinz and Baumgartner, 2001) are not contained here.

For airborne LiDAR data, the geometric property is considered to be the essential part of a vehicle model (Figure 3), which is used to support the recognition task in point clouds. The intensity of received laser pulses is so far hardly utilized due to lack of the calibration and the insight into physical background. The model represents the *standard case*, i.e. the appearance of vehicles is not affected by relations to other objects, e.g. shadow cast by buildings, vegetation occlusion or laser reflections of clutters. Moreover, since the detection of vehicles beneath the vegetation is also important, the modelling scheme for such vehicles is needed to cope with a variety of appearance.

For ALS data the implicit model can be regarded as 3D point pattern (set) of vehicles (Figure 3(a)), whereas the explicit vehicle model makes use of the surfaces plus their boundaries or the height discontinuity as 3D representation (Figure 3(b) and (c)). It is difficult to strictly distinguish two models and to make a choice on their performances at first glance. Unfortunately, the shade silhouette which complements the geometric model by significantly contributing to inference of the presence of small objects such as vehicles is not always available in ALS data. Both models focus on the geometric shape features without radiometric properties, and in terms of our research objectives and data characteristics, fundamental and robust features of cars cannot always be summarized by only using the vehicle model due to random reflection property of laser pulses against vehicle surfaces and unfavourable contamination by clutters nearby. It demands the incorporation of more advanced knowledge, such as context relations to ground and road, intensity or global context model, into the detection strategy.

(a)

(b)   (c)
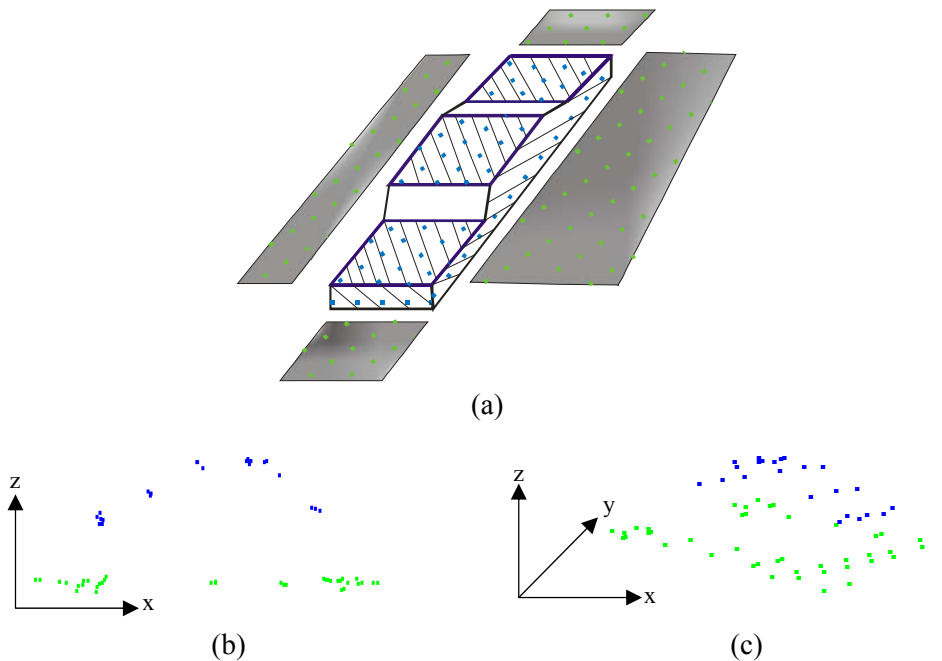
Figure 3: Vehicle model and point cloud (green: ground, blue: vehicle), a) schematic 3D vehicle model. Measured point cloud in b) side view and c) oblique view.

2.1.2 Context model

2.1.2.1 Local context

Local context describes geometric and topological interactions of a vehicle with objects in their immediate surroundings. The need for local context modeling results from the fact that many objects of urban areas are in a confined space and thereby local relationships between vehicles and their so-called context objects influence the interpretation of the scene strongly. The vehicle model above can be extended by knowledge about local context objects that contain background objects like roads, ground surfaces, buildings, trees or external data to substantially support vehicle extraction if their mutual relations are modeled appropriately. The model of local context relations between vehicles and urban primary objects are shown in Figure 4. This kind of knowledge can be incorporated into the extraction process to assist automatic inference in guiding weighting the order of processing steps. In dense settlements, many footprints of buildings are close to roadsides and give, therefore, strong hints for vehicles. Buildings or other high objects like trees potentially occlude parts of a vehicle or cast shadows on it. Hence, a context relation "occlusion" gives rise to the examination of possibly existent vehicles beneath thanks to the partial penetration ability of laser pulses against the volume-scatter objects, whereas a context relation "shadow" can tell an extraction algorithm to neglect these areas. Also, vehicles occlude the pavement of a lane segment and leave data gaps on it, but if an extraction system is able to identify them as such, they can be directly treated as vehicle. The ground surface is always under the vehicles which lie 0.3m- 3.5m higher and occlude the ground surface. Hence, calling a module that separate out ground level information is of top priority in the situation of intending to extract vehicles from ALS data. Furthermore, external GIS data like DTM is also able to be introduced into the context model since it usually almost approximates the ground level. Note, however, that the exploitation of specific context relations will be sensible to data acquisition configurations, and in most cases, be feasible in single-path ALS data with sufficient point density only (denser than about 4 points/m^2), because the context objects like cars with sub-structures and other topological features which contribute to the local context relations are usually not reliable and consistent in low density or co-registered data yet. Therefore, the modeling of local contexts is more tightly connected with ALS data of high point density.
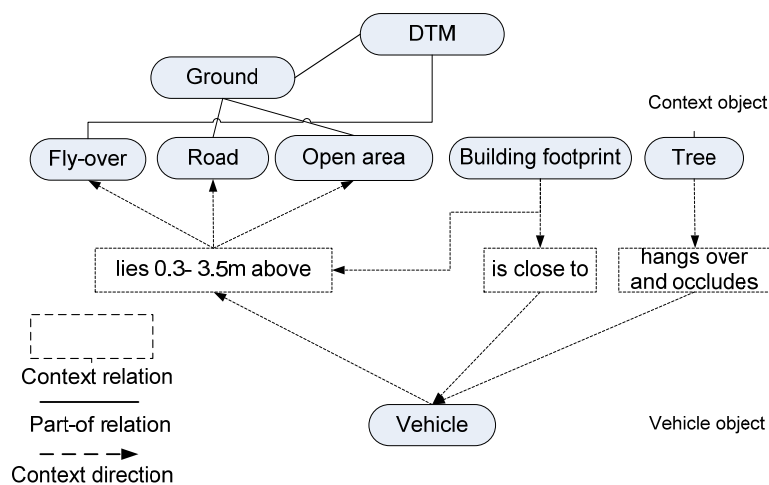
Figure 4: Local context-relations model in urban areas

2.1.2.2   Global context

The coarse impression of vehicles can be recognized by the global context. It enables a focus on the typical characteristics in the respective global scene, local contexts and model of vehicles. Modeling this kind of context supports high flexibilization of a vehicle extraction strategy, since depending on different global contexts the extraction strategy can provide various alternative routines and thus allows a dynamic control of the extraction process. This can lead to that for each region an individual extraction system with specially adapted model and methods is available, so that large scenes of various characteristics can also be processed with manageable models and efficient strategies.

Depending on respective coarse environments with particular features the vehicle model could vary, e.g. there are normally a relative large distance between two vehicles on the road, whereas vehicles at parking lots in dense urban areas are placed very close to each other and become inseparable by modeling vehicles explicitly only. The closest mutual relation between vehicle and adjacent context objects could have a different focus confronted with different areas. The modeling of the global context provides a classification of the urban scene into area of different properties. In this work it is distinguished between areas of two global properties - area without elevated roads where all road networks adhere to ground surface and area with elevated roads where road sections are extended over ground surface such as bridge and overpass. For the vehicle extraction in areas without elevated roads, we could accordingly resort to bugled object models and blob detection strategies, as developed in (Soille and Pesaresi, 2002; Hinz, 2004; Stilla and Michaelsen, 2002; Hinz and Stilla, 2006). In contrast to areas without elevated roads local contexts appear in urban areas with elevated roads need to be traced back to the spatial relation between vehicle and more generic background objects – road surface other than ground surface (Yao et al., 2008). Once the knowledge about the global property of an investigated area is acquired, the strategy to be adopted to extract vehicles will be adjusted in accordance to the corresponding global context.

2.2   Lie Group metric for 3D shape categorization

The vehicle database used in this thesis was constructed from the LIDAR scan of urban areas consisting of vehicles. The sparsity of points in point clouds can be observed from Figure 3 which makes it difficult to extract useful information, for example, learning the appearance model mentioned in Ferryman et al, (1998). Instead, the vehicle points are fitted with a shape representation called the spoke model in an automatic fitting procedure. The spoke models are subsequently used in a robust classification framework for distinguishing the motion state which is presented in Section 3.3.3. It introduces a 3D representation of vehicles as a space of scale and orientation transformations that define the shape space of individual vehicle instances. The property of this shape space is usually studied by the statistical shape analysis framework which is well understood when the parameters of the objects are elements of a Euclidean vector space. Although the shape description by the spoke model has proven to be effective, the model parameters are not naturally elements of a Euclidean space. It has been shown that spoke description of 3D vehicle points are in fact elements of a Lie group, since the airborne LiDAR can capture the nonlinear variability of the vehicle shape (Yalagadda et al, 2008).

This section is to introduce the mathematical foundation for a representation framework of 3D vehicle shape – Lie group, based on which the distinction of the motion state is performed in the sense of shape classification and categorization. The basics of the classification scheme using the principle geodesics analysis (PGA) are also presented to facilitate the understanding of advantages using Lie group based metric to distinguish 3D vehicles points of two motion states.

2.2.1 Lie Group theory

Lie group theory is the fundamental representation of a space of transformations. A Lie group G is a differentiable manifold that also forms an algebraic group (Duistermaat and Kolk, 2000), where the two group operations,

$$
\begin{aligned}
\mu : (x,y) &\mapsto xy \quad : \quad G \times G \to G \quad \text{Multiplication} \\
\tau : \ x &\mapsto y \qquad : \quad G \to G \qquad \text{Inverse}
\end{aligned}
\tag{1}
$$

are differentiable mappings. Many common geometric transformations of Euclidean space form Lie groups. For example, rotations, translations, magnifications, and affine transformations of all form Lie groups. More generally, Lie groups can be used to describe transformations of smooth manifolds. The group action could be thought of as a transformation of the manifold, whereas matrices are transformations of Euclidean space.

A Lie algebra g is a vector space with a bilinear product $[\ ,\ ]: g \times g \to g$, called the Lie bracket, for all $X;\ Y;\ Z \in$ g satisfying

$$
\begin{aligned}
Antisymmetry \ &\to [X;\ Y\,] = \ -[Y;X] \\
Jacobi\ Identity \ &\to [[X;\ Y\,];Z] = [X;\ [Y;Z]] - [Y;\ [X;Z]]
\end{aligned}
\tag{2}
$$

Let e denote the identity element of a Lie group G. The tangent space at e, T_eG, forms a Lie algebra, which we will denote by g. The exponential map, exp: g $\to G$, provides a method for mapping vectors in the tangent space T_eG into G. Given a vector $v \in$ g, the point $\exp(v) \in G$ is obtained by flowing to time 1 along the unique one-parameter subgroup emanating from e with initial velocity vector v. When the Lie group is given a compatible Riemannian metric, this one parameter subgroup is the unique geodesic at e with velocity v. The exponential map is a diffeomorphism of a neighborhood of 0 in g with a neighborhood of e in G. The inverse of the exponential map is called the log map. The geodesic distance between two points $g;\ h \in G$ is given by $\left\| \log(g^{-1}h) \right\|$, where $\left\| . \right\|$ is the Frobenius norm of the resulting algebra element(Rossmann 2002). The spoke transformation matrix T is the Cartesian product of transformation matrices M_i acting on individual spokes. Each M_i is a Cartesian product of $\Re \times SO(2)$ which forms a Lie group. Since the Cartesian product of Lie group elements is a Lie group, T forms a Lie group und can be viewed as a point on a Lie Group space.

2.2.2 Shape classification based on Lie group distance

Examining the similarity between 3D vehicle shapes can be formulated as a classification task on the manifold based on Lie group distance metric. This vehicle shape space forms a group, where the similarity of different vehicle observations can be evaluated using a distance measure defined by Lie group theory. A generic class of vehicles (e.g. Sport Utility Vehicle (SUV), van or passenger car) is represented by a set of curves on the Lie group manifold, called geodesics. The classification of any

given vehicle instance is achieved by finding the class with the smallest Lie distance between the geodesics and the specific category of vehicle shape. To adapt common classification methods used in Euclidean feature space we need to extend statistical measures and methods to manifolds.

A primary goal of the statistical shape analysis is to describe the variability of a population of geometric objects. The standard linear techniques, namely linear averaging and principle component analysis (PCA), cannot be applied. The standard shape analysis techniques can be generalized to handle manifold data by computing averages on a manifold and using a new method named principal geodesic analysis (PGA), a generalization of PCA, for describing the variability of data on a manifold. A principal geodesic is one that accounts for the maximum variation in a set of transformations along the path, analogous to principal component of a covariance matrix. Fletcher et al. (2003) used Lie group metric to measure the statistics of shapes on a Lie group manifold. They encode the shape of organs such as the human kidney as a set of 3D surface elements called an M-rep. The intrinsic variability in organ shape is represented by shortest curves (geodesics) on the Lie group manifold. Klassen et al. (2004) also used geodesic paths to span shape spaces and demonstrate their approach to generate a continuous family of shape prototypes within a category.

The three central statistical elements of the Lie group framework are: (i) Lie distance; (ii) intrinsic mean; (iii) principal geodesics. Lie distance is a measure of the similarity of two transformations. The intrinsic mean represents the average of a set of transformations, i.e. the transformation that minimizes the Lie distance to all the transformations in the set, which is defined as the minimizer in group G of the sum-of-squared Riemannian distances to each point on the Lie manifold (Srivastava and Klassen, 2002 and Grenander et al., 1998). The methods for computing geodesics and distances on them arise naturally from Lie group actions on manifolds. The averaging methods on manifolds have previously been studied very well (Fletcher et al., 2003); principal component analysis has recently been developed for manifolds and is to be reviewed here, as it is of great importance to the classification scheme on the Lie group manifold.

Considering a set of points $x_1, \ldots x_n$ on a manifold M. Our goal is to describe the variability of the x_i in a way that is analogous to PCA. Thus, we will project the data onto lower dimensional subspaces that best represent the variability of the data. This requires first extending three important concepts of PCA into the manifold setting, namely (i) Variance, (ii) Geodesic subspaces, (iii) Projection

- **Variance**: the sample variance of the data is defined as the expected value of the squared Riemannian distance from the mean.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} d(\mu, x_i)^2 = \frac{1}{N}\sum_{i=1}^{N} \left\| \mathrm{Log}_u(x_i) \right\|^2 \tag{3}$$

  where μ is the intrinsic mean of the x_i.

- **Geodesic subspaces**: The lower dimensional subspaces in PCA are linear subspaces. For general manifolds the concept of a linear subspace is extended to that of a geodesic submanifold. A geodesic is a curve that is locally the shortest path between points (transformations in group) and is the generalization of a straight line. Thus, it is natural to use a geodesic curve as the one-dimensional subspace, i.e., the analog of the first principal

component in PCA. The comparison of PCA and PGA is illustrated in Figure 5. Figure 5(a) shows the principle component in 2D Euclidean feature space while Figure 5b shows the principle Geodesics in 2D manifold of feature space. Submanifolds geodesic at *x* preserve distances to *x*. This is an essential property for PGA because the variance is defined as the average squared distance to the mean. Thus, submanifolds geodesic at the mean will be the generalization of the linear subspaces of PCA.

- **Projection**: In PCA, the data is projected onto linear subspaces. For a projection operator defined for geodesic submanifolds, the projection of a point $x \in M$ onto a geodesic submanifold H of M is defined as the point on that is nearest to x in Riemannian distance. Thus, the projection operator $\pi_H: M \to H$ as

$$\pi_H(x) = \arg\min_{y \in H} d(x, y)^2 \tag{4}$$

It can be assured that projection is unique for any submanifold geodesic at the mean by restricting to a small enough neighborhood.

| (a) | (b) |
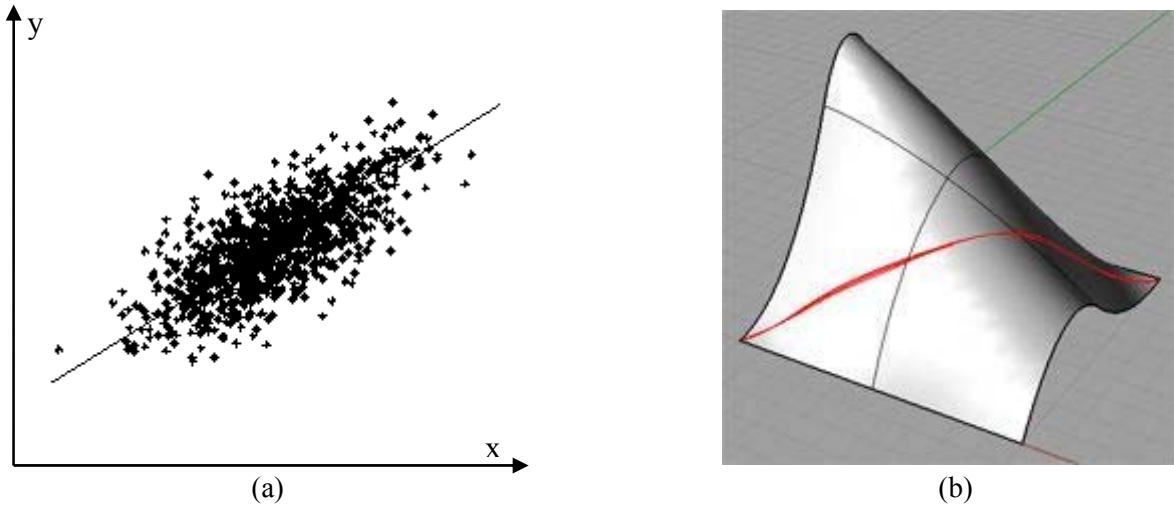
Figure 5: Comparison of PCA and PGA (only the first principle component is shown here), (a) PCA in 2D feature space, (b) PGA in 2D manifold of feature space, red line indicates the principal geodesics

The computation of PGA is to find a sequence of nested geodesic submanifolds that maximize the projected variance of the data. These submanifolds are called the *principal geodesic submanifolds*. The principal geodesic submanifolds are defined by first constructing an orthonormal basis of tangent vectors $v_1, \ldots, v_d \in T_u M$ that span the tangent space of M at the intrinsic mean u of x_i. These vectors are then used to form a sequence of nested subspaces $V_k = \mathrm{span}(\{v_1, \ldots v_k\}) \cap U$. The first principal direction is chosen to maximize the projected variance along the corresponding geodesic

$$v_1 = \arg\max_{\|v\|=1} \sum_{i=1}^{N} \left\| \mathrm{Log}_\mu(\pi_H(x_i)) \right\|^2 \tag{5}$$

where $H = Exp_\mu(\mathrm{span}(\{v\}) \cap U)$, is the exponential map and let $U \subset T_u M$ be a neighborhood of 0. The remaining principal directions are then defined recursively by spanning different basis vectors.

3 Methodology

This chapter is to present and describe the whole strategy for monitoring the urban traffic using ALS data in detail. At first, an overview of the complete strategy is given without going into algorithmic details. Second, the individual steps of the implementation of the vehicle extraction will be explained. The three *Routes* of the vehicle extraction strategy outlined in Section 3.1 are reflected in the outline of this chapter. The analysis of the global context is described at the beginning. Subsequently, it deals with the extraction of vehicle objects from data areas of the global context without elevated roads. Section 3.2.3 explains the object based point cloud analysis (OBPA) method used to deal with data areas of the global context with elevated roads, and if necessary, introduces generated intermediate results from the OBPA method into the local context guided method and finally fuses their extraction results for areas with dense placement of vehicles. The third part of this chapter is to first introduce and investigate the moving object model (motion artifacts effect) in single-path airborne LiDAR data, which lays down the theoretical foundation for LiDAR to monitor the traffic from airborne platforms. Furthermore, a scheme is proposed to distinguish the vehicle motion state from ALS data based on the extracted points of single vehicles using a binary shape classifier constrained by Lie group metrics. The shape parameters describing the vehicle geometric property in view of motion artifacts have been estimated in the shape parameterization process applied a-prior, where the vehicle points of ambiguous shapes are directly removed from the classification step of motion state and classified as uncertain moving class. Finally, the velocity of all moving vehicles can be quantitatively derived with knowledge about their shape deformations. Additionally, the velocity estimation accuracy will be extensively studied and evaluated for different estimation methods under error prorogation principle.

3.1 Overview of the complete strategy

The complete strategy lays down the essential fundament for the control of the extraction and analysis process. Through it the framework that when and what the knowledge about object is used for extraction and how the results from intermediate operations will be involved into the further procedure is given. Therefore, the strategy includes knowledge about the extractability of features or sub-objects as well as the efficiency of algorithms. The whole strategy consists of two main stages described in more detail in the next subsections: (1) vehicle extraction (Section 3.2), (2) Analysis and estimation of vehicle motion (Section 3.3),

Because of the task and data complexity to be treated here, a hybrid strategy for vehicle extraction is to be applied - namely a model-driven approach alternating with a date-driven approach, controlled by global context analysis in the initial phase. At the start of the model-driven approach "Region of Interest" (RoI) for vehicle extraction, i.e. ground surface is extracted. Based on the generated hypotheses the extraction can now focus on the places where vehicles are able to be detected in a more reliable and easy way. Subsequently, series of operations including morphological reconstruction and watershed transformation lead to the accurate delineation of most vehicles. Since the vehicle and context model incorporates 3D information as well as spatial mutual relationships among objects in the scene, it is more beneficial to perform the vehicle extraction algorithm directly

at 3D data level. To this end, a 3D segmentation-based classification approach is proposed to accomplish the vehicle extraction task too. In contrast to other approaches, it neither uses pulse intensity information for extraction nor extracts geometric features in 3D locality as, e.g. Gruen and Li (1997) or Zhang et al. (2001) did, by estimating and selecting local neighborhoods using TIN model. These procedures are conceptually elegant, but they imply feature matching algorithms and geometrical/topological analysis in 3D during the extraction process, which becomes cumbersome for scenes of higher complexity. On the other side, it is desired to avoid feature extraction in point cloud of urban areas because inaccuracies of the estimated local geometry due to erroneous neighborhood selection could disturb shape properties of object structures. The strategy in this work is to stay in 3D raw data exclusive of pulse intensity as long as possible and to extract vehicle objects in large extent and other high semantics. It is assumed that geometric component of ALS data is more essential, since it involves most informative hypotheses and encodes significant information for semantic modeling. The concrete workflow of the complete strategy is illustrated in Figure 6.

Stage 1: Vehicle extraction

This step comprises a hybrid approach intended to detect all the points of vehicle instances which should serve as the data foundation for motion analysis. The approach combines a local context-guided method with the OBPA method to accomplish the task, the switch from one method to another within the framework is controlled depending on which kind of the global context exist in the data area to be analyzed. Therefore, this stage starts with exploiting global contextual knowledge to make this information available for following processing stages. After distinguishing between urban areas with elevated roads and urban areas without elevated roads, using global context as guide for automatic vehicle extraction in urban areas becomes a trivial thing and the judgment is to be made upon that which method is selected to proceed subsequently. The output of this operation gives a rough idea about the position of most potential vehicles that could appear as well as about method which is most likely to yield promising results. In this work there are totally three *Routes* to perform vehicle extraction from airborne LiDAR data depending on global context analysis, as depicted in Figure 6. The first *Route* dedicated to the dataset without elevated roads will apply the method of vehicle extraction based on exploiting the (local) context while the second *Route* is dedicated to the dataset with elevated roads using the OBPA method to extracted vehicles. For third *Route*, it is dedicated to a refinement of the OBPA method for vehicle extraction in areas with elevated road. Following an initial 3D segmentation it enables the context-guided method to extract vehicles from the large context objects such as ground (road) surface again, which is made available by introducing the grouping operation using normalized cuts to the results generated by OBPA method. The extracted vehicles obtained in the refinement process (*Route* 3b) are to be merged with extraction results obtained by the OBPA method performed previously (*Route* 3a) to form the final result. The combination of the OBPA and local context-guided method can make the extraction more complete for the vehicles which are located very close to each other such as in parking lots.
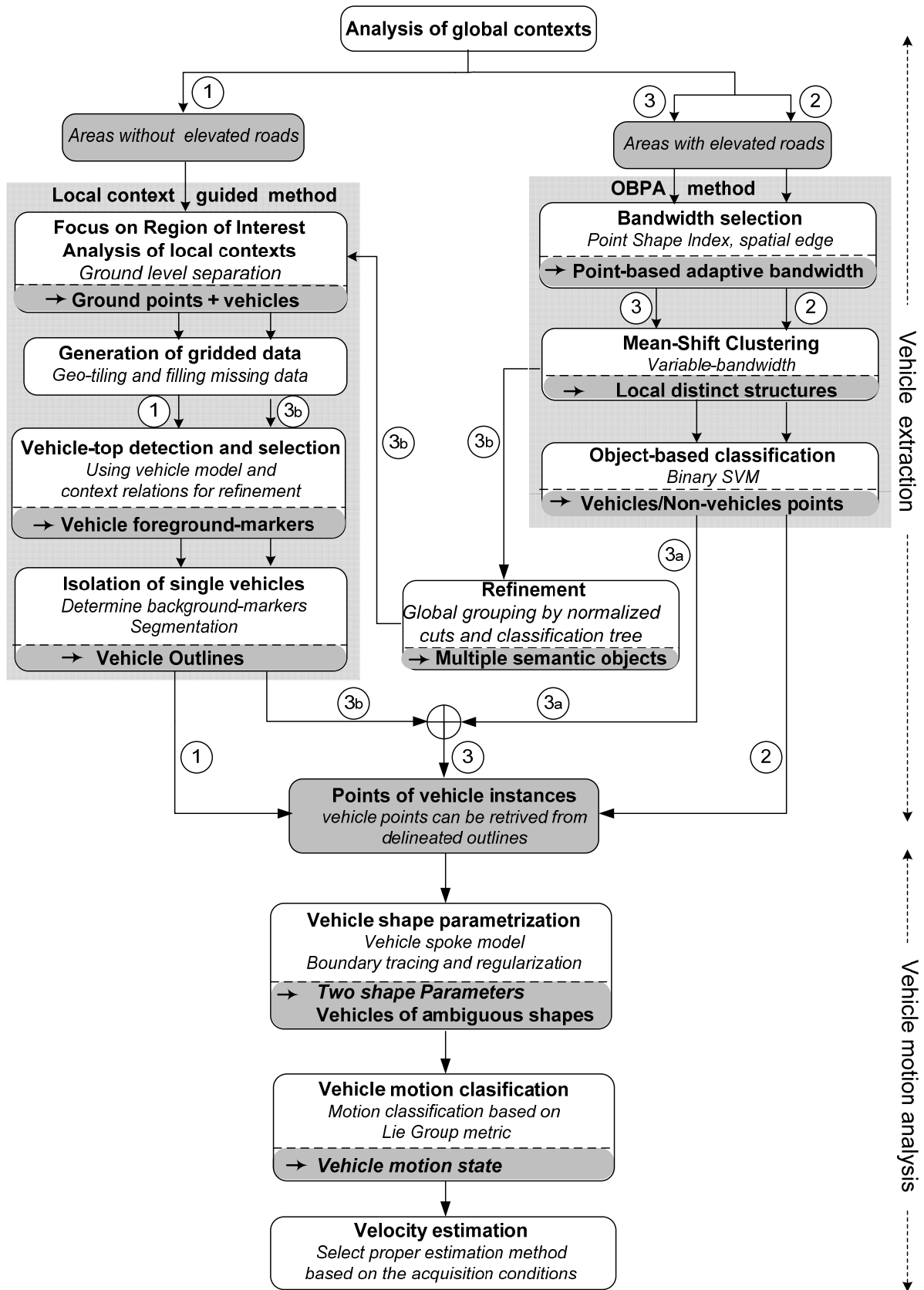
Figure 6: Workflow of the strategy

The (local) context-guided method utilizes the fundamental contextual relation between vehicle and ground surface to reduce the problem domain and provide the RoI. Then, the vehicle height model generated by last step is to be utilized to delineate vehicle instances by advanced geodesic morphology. The OBPA method is directly performed on the ALS raw point cloud in genuine way. Local geometric modes are detected using mean shift clustering by varying the local bandwidth adaptively; the obtained point segments can give a sensible partition of scene to represent large-scale object (parts) while keeping small-scale objects remain. Finally, the binary SVM classifier is adopted to label the point segments as vehicle or non-vehicle objects based on the various geometric and shape features derived at object-level. Both methods utilize the geometric components of ALS data only. Especially for data areas with both elevated roads and parking lots delivered to the second Route, the combination of two methods by a refinement operation can not only guarantee a considerable completeness of vehicle extraction, but also for each vehicle provide higher geometric accuracy in terms of shape fidelity.

Stage 2: Vehicle motion analysis

After vehicles are extracted from ALS data as set of points each, the next step focuses on motion analysis. It consists of two sub-tasks – determination of vehicle motion state and velocity estimation.

The motion state can be determined based on 3D shape classification of vehicle points subject to motion artifacts. To this end, each of extracted vehicle instances is firstly modeled by spoke model and then delivered to shape parameterization based on boundary tracing and regularization. The two controlling parameters describe the shape properties of vehicle spoke model in the point cloud – the aspect ratio and the shearing angle are obtained by the shape parameterization process and further delivered to motion classification process based on Lie group metric. The Lie group based classifier can essentially encode the nonlinear variability of vehicle shapes represented in ALS data and perform the binary classification task on distinguishing between moving and stationary vehicles more robustly and effectively compared to other common classification methods working on the linear Euclidean space only. Then, vehicle points can be classified into three categories concerning the motion state – moving, stationary and uncertain which is already detected in the boundary regularization process based on imposing the constraint on quadrilateral that every two sides should be mutually parallel. The vehicle instances of ambiguous shapes are excluded from the motion classification and directly labeled as uncertain class.

From obtained vehicles of moving class, the velocity can be estimated based on the measuring the shape deformations by inverting the motion artifacts model of ALS data. In general, the velocity estimation concept can be divided into two categories depending on whether the moving directions of vehicles are known or not. Usually, the velocity of sensor flight can easily be known from on-line navigation system in advance while the intersection angle between the sensor flying path and vehicle moving direction has to be explicitly derived in addition. The moving direction can be indicated either by the principal axis for each set of vehicle points or by the road direction given as the a-prior knowledge. The correct moving direction is chosen from the two alternatives in accordance with of the parameterized shape of vehicle points or the shape deformation direction (elongated or shortened). Accordingly, the velocity of moving vehicle can be estimated independently based on either the along-track stretching or the across-track shearing effect. Moreover, the velocity can also be derived by combining along-track with across-track velocity

components by the sum of squares operation which are estimated, respectively. However, if the moving direction of vehicles cannot accurately be estimated in advance, the velocity can be derived by uniting and solving the two bivariate equations in the moving object model directly while the intersection angle between the velocity vectors of flying sensor and moving vehicle is determined as a solution simultaneously. Finally, theoretic accuracy analysis of velocity estimation methods is to be performed in detail. The empirical relationships of error propagation between the estimated velocity and several observation parameters are to be specified in an analytic form.

3.2 Vehicle extraction

In this section, the individual steps of the implementation of the vehicle extraction will be explained. The three *Routes* in the strategy outlined in Section 3.1 are reflected in the outline of this section. Section 3.2.1 describes the analysis of the global context at the beginning of the vehicle extraction. Subsequently, Section 3.2.2 deals with the extraction of vehicle objects that correspond to data areas without elevated roads (*Route* 1). Section 3.2.3 explains the OBPA method used to deal with data areas with elevated roads (*Route* 2), and if necessary, introduces generated intermediate results from the OBPA method into the local context guided method and finally fuses their extraction results for areas with presence of densely placed vehicles (*Route* 3).

3.2.1 Global context analysis

After dividing the context model in Section 2.1.2 into a global and a local component, the global context is described here. Figure 7 summarizes the part of global context analysis. They serve primarily to focus the object extraction process to such areas that are likely to attract a good agreement between the object model and the actual object characteristics and thus make sure that the extraction is most successful. The extraction strategy presented in Section 3.1 regards the use of global contextual knowledge as a first step, since background information is obtained through this way, which is able to support the selection of vehicle extraction methods hereafter. The motivation for employing global context stems from the observation that it is possible to distinguish semantically meaningful data regions - so-called context regions - where vehicles show typical prominent features and certain relations to background objects have a similar importance in vehicle extraction. Consequently, the relevance of different components of the vehicle model and the importance of different context relations must be adapted to the respective context region. In urban areas, for instance, relations between vehicles and (elevated) roads are more important and elevated roads such as overpass and bridge are of a great significance due to their more frequent occurrence in city areas to make large quantity of traffic loads be dredged in an effective way, and henceforth, make vehicle extraction dependent on distinguishing different context regions. Actually, all the vehicles on elevated roads move at a considerable speed and usually are located not very close to each other but with a certain inter-object distance. Unlike other vehicles that move or park on ordinary roads, the position of these vehicles are no longer adhered to the context object – "ground surface". Based on these facts, the method proposed to extract vehicles should be also adapted to tackling different situations in ALS data. Urban areas have been distinguished between contexts with elevated road and without elevated road. To this end, two vehicle extraction methods are proposed to solve the problems emerging in two global contexts, respectively. The distinction of an ALS dataset of dense urban areas from the region with elevated roads to the region without elevated roads can be

performed by manual judgment. Since the global context analysis performed here is only limited to distinguishing between two hypotheses, the automation of this process is not indispensable which can be easily replaced by a human intervention. Otherwise, the distinction of the global context regions can be automatically done by analyzing the statistical property of the height distribution of ALS point cloud based on histogram analysis.
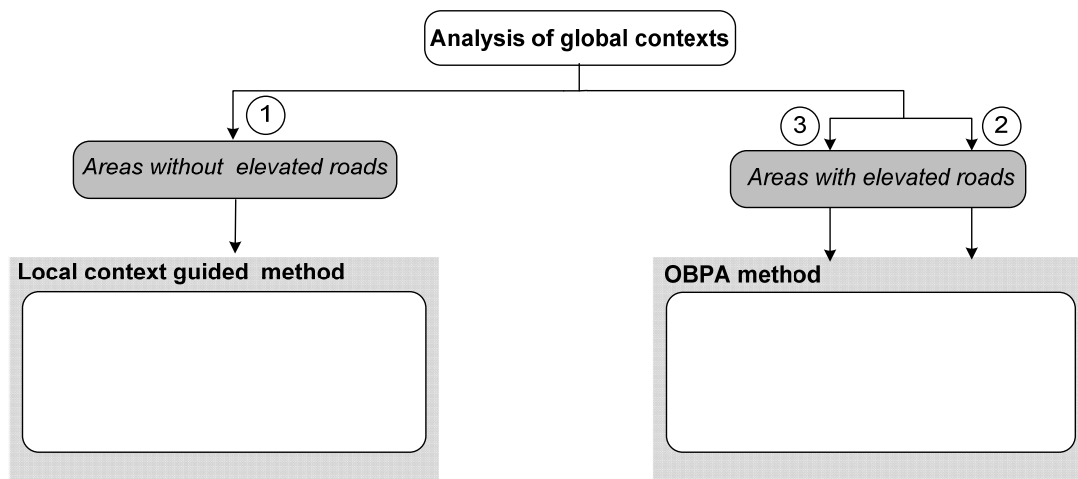
Figure 7: Extraction strategy: part "global context analysis" (cf. Figure 6)

3.2.2 Local context guided method

3.2.2.1 Ground surface separation

The local context model from Section 2.1.2.1 describes relations between vehicle objects and context objects "road", "ground"," building footprint "and" tree" and derives certain conclusions on derivation of a extraction strategy. They serve primarily to focus the vehicle extraction to such areas, that a good accordance between the vehicle model and the actual vehicles is expected and thus can guarantee a successful extraction.

The local contexts for vehicle that concern the buildings, ground surface and vegetation include the relations "is close to", "lies above" and "hangs over and occludes". For the analysis of these relations, the accurate reconstruction of the building or detailed extraction of vegetation needs to be carried out. However, they have established themselves as independent research topics for a long time. The automatic extraction of such objects in urban areas from airborne LiDAR data is itself not trivial task which has not become a solution mature enough and needs to be further investigated thoroughly. In contrary, the strong relation between vehicle and ground surface including road and other terrain surfaces is believed to be fundamental to the whole procedure of vehicle extraction. Meanwhile, the automatic extraction of ground point from ALS data has been well studied in the context of data filtering for terrain model generation and becomes an efficient and rapid solution to topographic information extraction for urban areas. Consequently, it is decided to start by exploiting the essential knowledge in the local context model to benefit all following processing stages. Therefore, the automatic separation of ground points from other object points in LiDAR data is used in this work to support the vehicle extraction. The output of these operations gives a rough idea about the position of potential vehicles as well as about regions where no vehicle is expected to

appear, namely providing the RoI towards vehicle extraction. Section 4.2.1 outlines this algorithm. Moreover, information about the vehicle model itself such as area and eccentricity are to be drawn in the vehicle-tops selection process (cf. Section 4.2.3) to remove the false alarms within the detected vehicle-tops which are formulated as local maximum too.

The ground points refer to the bare earth defined as by Sithole et al. (2003) plus vehicle. Object points comprise detached (buildings and trees) and attached objects (bridges). Hence, ground level separation is used to mask out background objects (e.g. buildings) to provide a compact dataset consisting only of vehicles, real ground and few disturbing objects above, e.g. pole or traffic signs, wall and tree stem etc. This step essentially facilitates the search process of vehicles by reducing the problem domain.

An adaptive thresholding method based on height distribution proposed by Bartels et al. (2006) is selected to perform the task, which can directly implement the key contextual constraint on vehicle extraction. It is stated that naturally measured samples (ground terrain points) will lead to a normal distribution according to the central limit theorem. Thus, the assumption is that object points may disturb the normal distribution and ground points can be obtained by removing those from the raw point cloud. Vehicle points may bias the normal distribution of ground points, but meanwhile are compensated by many topographical features in cities, such as ramps, connection paths to tunnel or garage, therefore making the height of ground surface points including vehicles still obey the rule of normal distribution. Two meaningful measures of skewness sk kurtosis ku to describe degrees of asymmetry and peakedness of the sample distribution are applied, which are defined by

$$
\begin{aligned}
sk &= \frac{1}{N \cdot \sigma^3} \cdot \sum_{i=1}^{N} \left(s_i - \mu_a \right)^3 \\
ku &= \frac{1}{N \cdot \sigma^4} \cdot \sum_{i=1}^{N} \left(s_i - \mu_a \right)^4
\end{aligned}
\tag{6}
$$

where N is the total number of the LiDAR points s_i, σ is the standard deviation and μ_a is the arithmetic mean.

For a normal distribution, sk is zero and ku is three. The algorithm balances the point's distribution by iteratively removing the highest (maximal) value of the point cloud until the skewness and kurtosis of the rest points converges on zero and three simultaneously. The remaining points are supposed to be the ground surface, whose height obeys the normal distribution (c.f. Figure 9). For more hilly areas, an iterative implementation of the algorithm is used to overcome the undersegmentation by treating the imperfect set of object points (nDSM) again as a DSM and rerunning height distribution balancing. Finally, a refinement of ground points could be achieved by combining DTM filtering with layer-cutting operation (a layer of 3m thickness above the ground surface is to be cut out) in order to exclude small-size non-topographic effects and meanwhile preserve vehicle points (Yao et al., 2010). Figure 8 shows one example of separation results distinguishing between ground (dark blue) and object (green) points. As depicted in Figure 8, detached objects were clearly detected, however, a few attached objects not, due to the complex scene. The first echo DSM of parts of Toronto in Canada in Figure 8 is constructed from 108, 790 LIDAR points. It shows a highly dense urban area with mostly detached objects, both with buildings and vegetation of various heights. Nearly all detached object points are correctly classified and

removed from the ground. The histogram of the LIDAR data tile in Figure 9 is also noteworthy showing both classified ground (dark blue) and object (green) points.
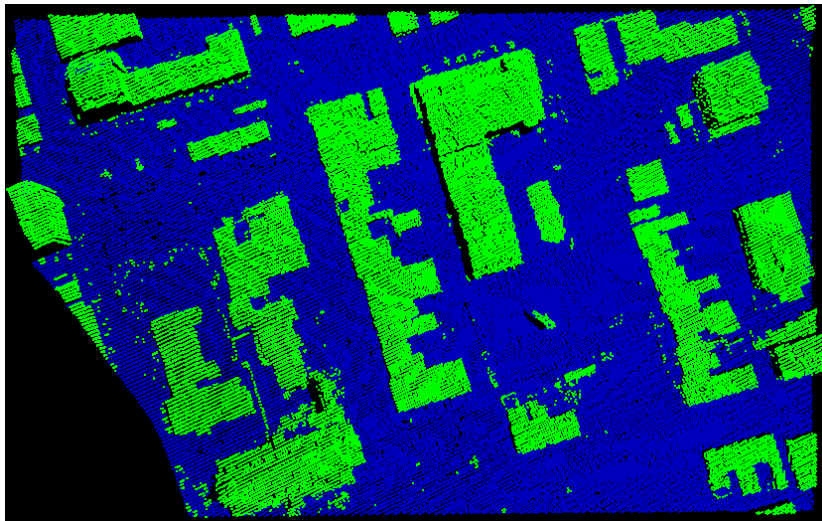
Figure 8: Separation of ground (green) and object (blue) points from ALS data of a city center. Black depicts non-point regions which are not acquired by scanning

The algorithm to separate the ground points from objects has several advantages. First, it is a non-parametric unsupervised classification algorithm that requires neither a specified threshold nor a fudge factor prior to its execution. Furthermore, the separation of object and ground points preserves the original LIDAR data and hence avoid the loss of information due to manipulation. Moreover, the algorithm is applicable to every LIDAR echo, where the processing of last pulses of LIDAR data is the fastest. This is due to the laser's ability to penetrate vegetation which results in less object points to be classified. The algorithm is therefore very flexible and can additionally be used for applications without restriction to the data format.
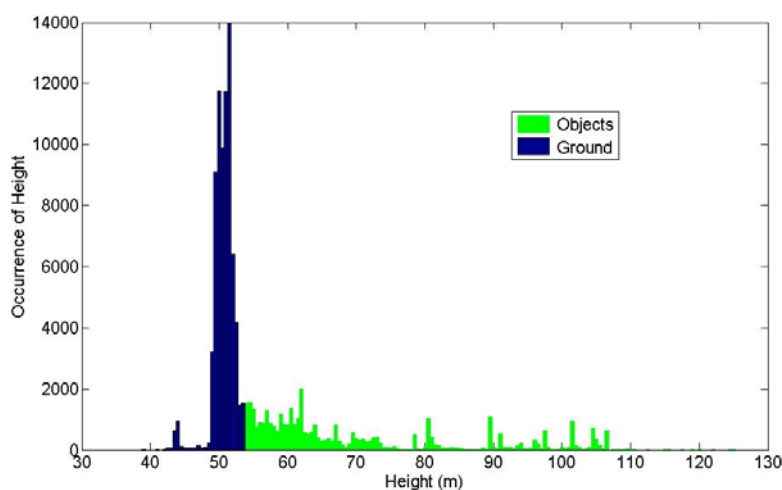
Figure 9: Height histogram of classified points in Figure 8

3.2.2.2 Geo-tiling and filling missing data

To make the vehicle extraction from ALS data greatly benefit from advanced operations of image processing, the raw ALS data has to be transformed into grid data. It is done here by geo-tiling

process, where rather than interpolation only the correspondence relation between single points and grid cell is established. The geo-tiling process is originally developed to overcome the spatial complexity of ALS raw data and accelerate the data-access. A geo-tile index is generated by the grid subdivision method (Laszlo, 1996), by which the data space is formatted into $n \times n$ grid cells with size $d \times d$ and the spatial index information about single points is stored (Figure 10). Then, each cell records the lowest last return of all pulses falling in the cell. To simultaneously preserve the local details and make data storage acceptable (Chen et al., 2006), the cell size d is set according to $d = \sqrt{1/\lambda}$; λ is average point density (points/ m^2). In practice, the pulse density varies owing to side overlaps between swaths. Therefore, a small cell size should be used when there are large pulse densities locally, which can be chosen by calculating the pth quantile of λ. Using geo-tiling for LiDAR data gridding has several advantages. First, it speeds up the data searching mechanism; second, it enables a flexible and efficient interaction between gridded data and raw airborne LiDAR points, e.g. retrieving 3D vehicle points in ALS points can be achieved by outlining their boundaries in gridded data followed by hole-filling.
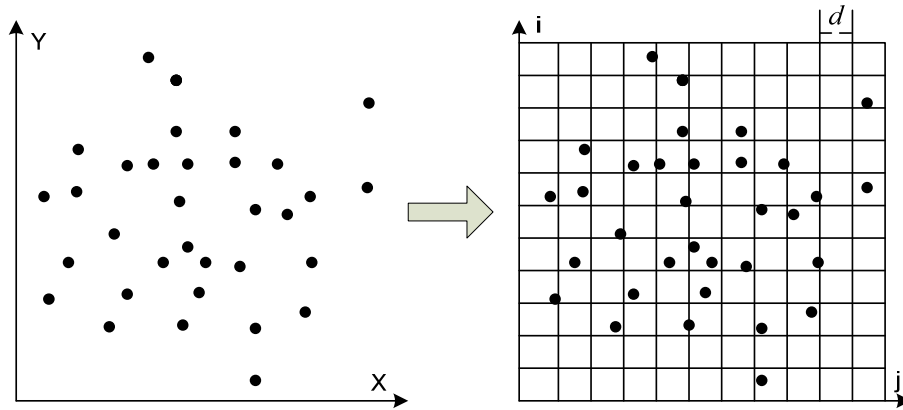
Figure 10: Schema of geo-tiling for LiDAR points indexing

When the pulse spacing is greater than d , there are no values for some cells by sparsely sampling. The same situation also happens caused by absorptive materials, such as waters. A general surface fitting strategy is adopted to locate and fill the missing data simultaneously. It is not an interpolant and uses a modified ridge estimator (Marquardt, 1970) to generate the surface of the form z(x, y) that approximates scattered data $\{x_i, y_i, z_i\}$ by successively resolving two coupled equation systems under a given constraint:

$$\begin{cases} \min \left\{ (\|\mathbf{A} * x - y\|)^2 + (\lambda * \|\mathbf{B} * x\|)^2 \right\} \\ \mathbf{A} * x = y \wedge \mathbf{B} * x = 0 \end{cases} \tag{7}$$

where $\mathbf{A}$ accounts for the interpolation system with plate-like metaphor; $\mathbf{B}$ accounts for the regularization system added to solve the underdetermination; λ is a parameter for controlling the plate "stiffness".

A nice feature of this surface fitting method is that it is able to smoothly extrapolate beyond the convex hull of data, something that interpolator cannot do. Since the object points rejected by the ground separation step presented in Section 3.2.2.1 is quite numerous and intendedly represented as blank areas (Figure 11), it is required to distinguish between unintended and intended data gaps. By adjusting two parameters of the estimator - *tilesize* and *smoothness*, the two situations can be

handled very well simultaneously. The small data gaps are filled while the large data gaps are preserved.

(a)

(b)
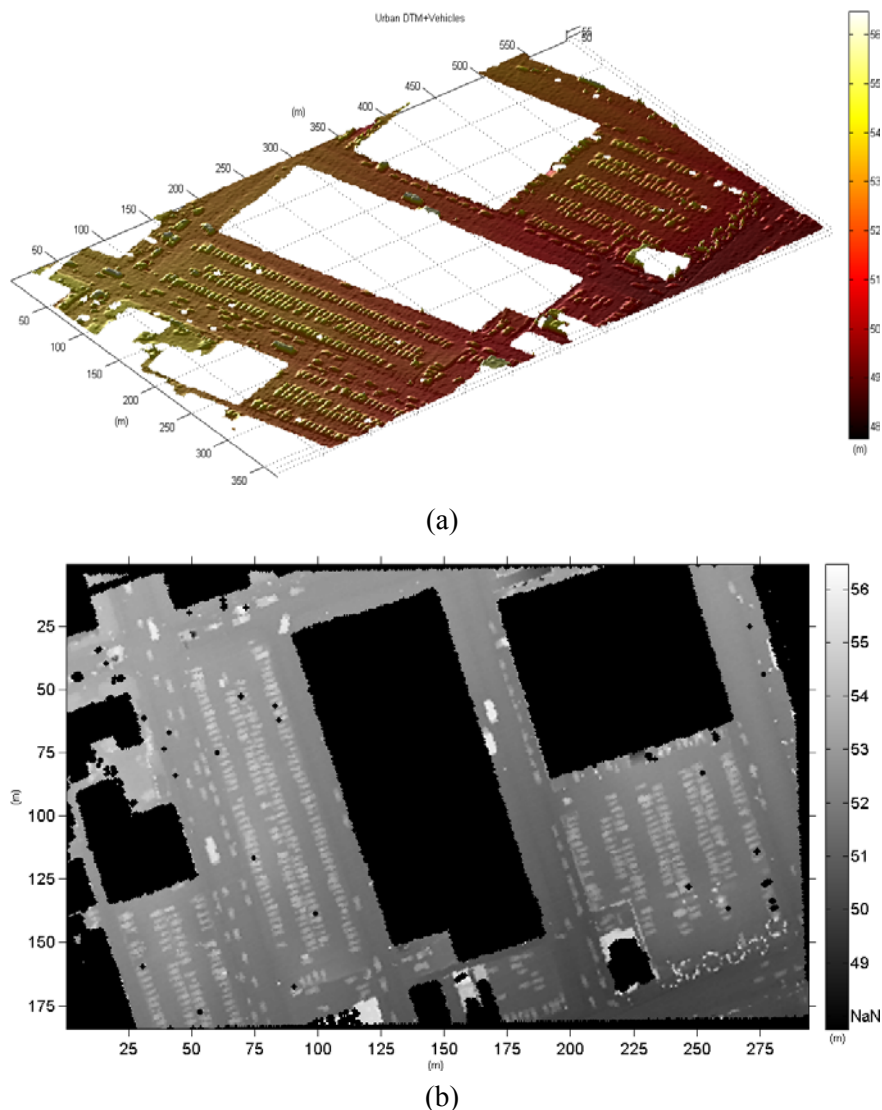
Figure 11: Filled height raster of ground points with surface fitting, black areas indicate object points masked out, (a) perspective-view; (b) top-view

3.2.2.3 Vehicle-top detection and selection

Based on the results generated in Section 3.2.2.2 – gridded and filled LiDAR ground data f, the next step is to find vehicle-tops as the foreground markers for vehicle segmentation, which refers to the roofs of vehicles. Local maximum filtering is often used to find the object-tops in a height model accompanied with large commission errors. As indicated in Section 2.1.1 and inspired by Soille and Pesaresi (2002), the common vehicles can be modeled as bulges in the height data and the vehicle-top consists of connected blobs of pixels inside each of the foreground objects, extended-maxima transform of the filled ground height grid f, denoted as $EMAX_h(f)$, is computed to provide initial candidates for the vehicle-tops. $EMAX_h(f)$ is the regional maxima ($RMAX$) of the h-maxima transformation of f, which is to perform the reconstruction by dilation of f from $f - h$, denoted as

$$f_{h\,\max} = R_f^{\delta}\left(f - h\right) \tag{8}$$

where $f_{h\,\max}$ is the *h*-maxima transformation of *f* (Soille, 2003).

(a)

(b)
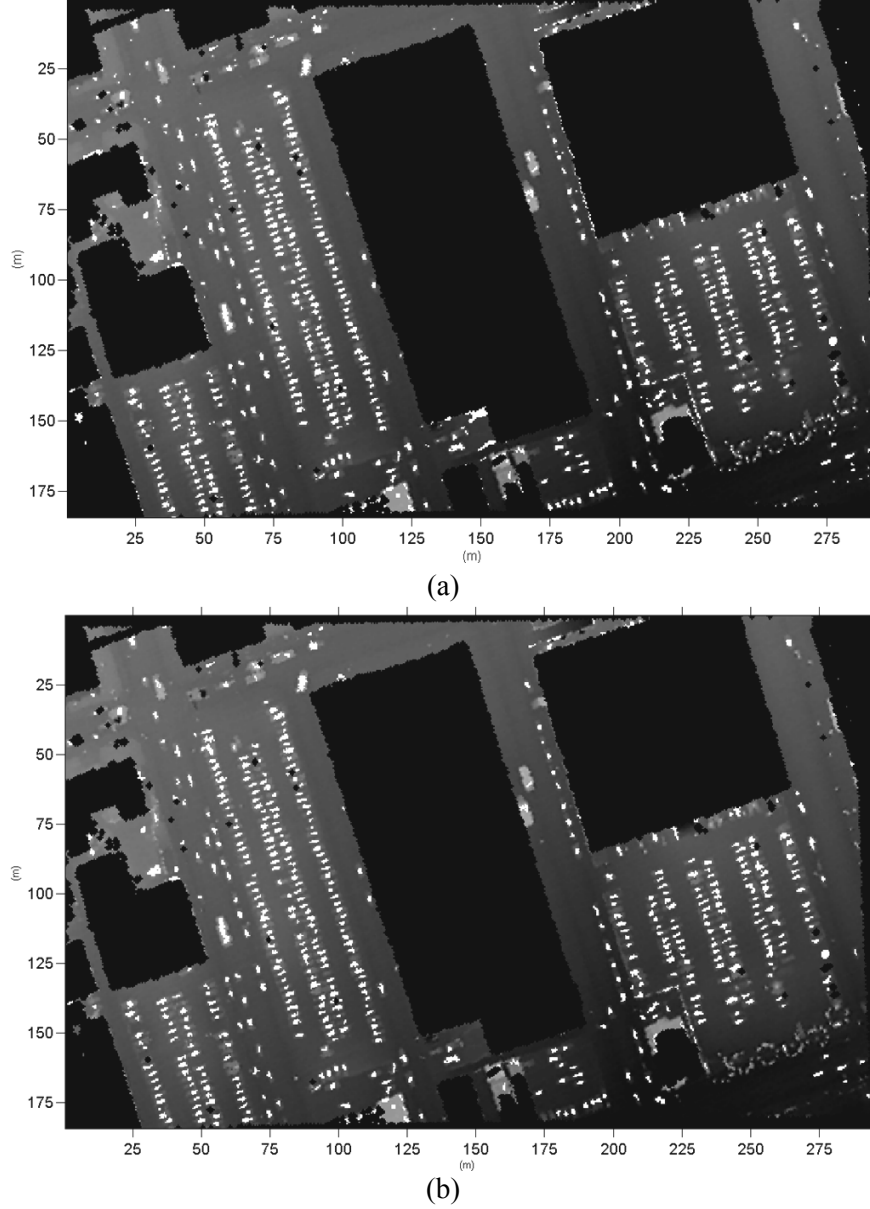
Figure 12: Detected vehicle-tops (white blobs) superimposed on Figure 11, (a) before refinement, (b) after refinement

The regional maximal considers the union of the maxima obtained at each level. Since the set union can be replaced by a summation and the image data type does not support negative values, it gives:

$$EMAX_h(f) = RMAX(f_{h\,\max}) = f_{h\,\max} + 1 - R_{f_{h\,\max}+1}^{\delta}(f_{h\,\max}) \tag{9}$$

Note that $EMAX_h(f)$ include the maxima not only caused by the vehicle-tops, but also ground irregularities and some anomaly objects which are not completely removed by the step presented in Section 3.2.2.1 (Figure 12 (a)). The next step is to generate refined results by selecting the

$EMAX_h(f)$ based on their shape measurements and local context relations. Since the vehicle-tops usually have a small blob-like shape in reality and nearby building foots might be detected incorrectly, the area α and eccentricity e of the detected vehicle-tops are derived and examined with respect to the thresholds ($\alpha_{T1}, \alpha_{T2}, e_T$) using the criterion: $\alpha_{T1} \leq \alpha \leq \alpha_{T2} \wedge 0 \leq e \leq e_T$, where α_{T1} =4 pixels α_{T2} = 25 pixels e_T =0.68. $\alpha_{T1}, \alpha_{T2}, e_T$ are derived by sample data given ahead. After refinement, spurious local maxima other than vehicle-tops were greatly reduced (Figure 12 (b)). The threshold h directly affects the performance of detecting vehicle-tops. However, the choice of h is less data-dependent and can be set by trial. Moreover, an optimal value of h can be obtained from training data given beforehand. According to experiments, the change of h would mostly affect the omission errors not the commission errors. It is due to the fact that vehicles usually have quasi-constant height of above the ground, whereas the area α and eccentricity e of the detected vehicle-tops are used to filter out spurious local maxima (i.e. false alarms) other than vehicle-tops with the quasi-same height. Applying such a strategy, both commission and omission errors for vehicle-top detection can be alleviated.

3.2.2.4 Segmentation

The marker-controlled watershed transformation is applied to segment vehicles and delineate their boundaries which are delivered to retrieving the points belonging to single vehicles. In marker-controlled watershed transformation, the segmentation function is first filtered by minima imposition so as to remove all irrelevant minima. Thus, basins are flooded from selected sources and over-segmentation is to be suppressed (Soille, 2003).

Suppose the segmentation function is gradient magnitude of f, denoted as $G(f)$, and a marker image f_m has been specified at each pixel p:

$$f_m(p) = \begin{cases} 0, & \text{if } p \text{ belongs to a marker,} \\ t_{max} + 1, & \text{otherwise.} \end{cases} \tag{10}$$

where t_{max} is the maximum value of the input image. Minima imposition is to first calculate a pixel-wise minimum between $G(f) + 1$ and the marker image f_m as $(G(f) + 1) \wedge f_m$, and then perform a morphological reconstruction by erosion of $(G(f) + 1) \wedge f_m$ from the marker image f_m.

$$f_{mp} = R^{\varepsilon}_{(G(f)+1) \wedge f_m} (f_m) \tag{11}$$

where f_{mp} is the resulted image by imposing minima and will be undergone the ordinary watershed transformation.

To extract the vehicles as accurate and complete as possible, the marker image should contain the markers containing both foreground markers ─ vehicle-tops and background markers─ non-vehicle objects. The two markers enhance each other by simulating the flooding from two corresponding basin catchments to build a more balanced dam line. Therefore, they should not lie too close to each
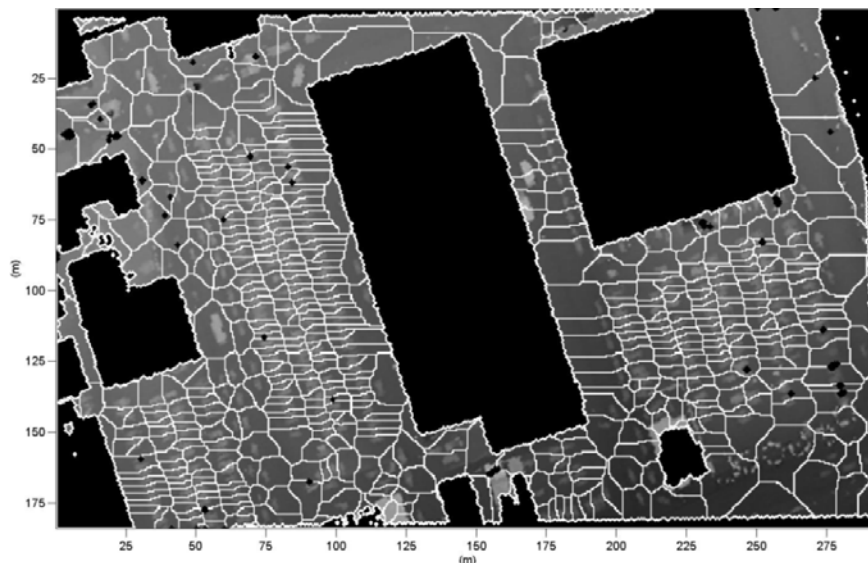
Figure 13: Thinned background markers

other and the background areas usually need to be thinned to the ridge lines of their distance-transformed version (Figure 13). Especially, it makes sense for places such as heavily traveled roads and parking lots. It has been verified by tests that the background markers should be considered here especially when we want to extract vehicles which are placed closely to each other. This situation usually appears in parking lots or heavily traveled road sections. The background markers can prevent the flooding simulation process in the watershed transformation from over-flooding effect. That means that the boundaries of adjacent objects to be segmented (each object has one boundary) can be flooded (approached) from two directions (basins), so that they can lead to building the watershed (dam) line by enforcing mutual constraints, corresponding to the vehicle outlines in real cases more accurately. Figure 13 has been created anew to avoid the computations occur in the no-data (NaN) regions. Minima imposition reconstructs the segmentation function $G(f)$ in such way that there are only minima (flooding sources) corresponding to markers. This illustration highlights the importance of finding the correct vehicle-top/non-vehicle marker function when applying such a strategy to extract vehicles. Figure 14 shows two examples of vehicle extraction results when applying the proposed method to one dataset, the difference between which is only limited to whether the background marker for watershed transformation is used or not. This comparative illustration between Figure 14 (a) and (b) can obviously indicate the positive impact of the background marker on the vehicle delineation accuracy.
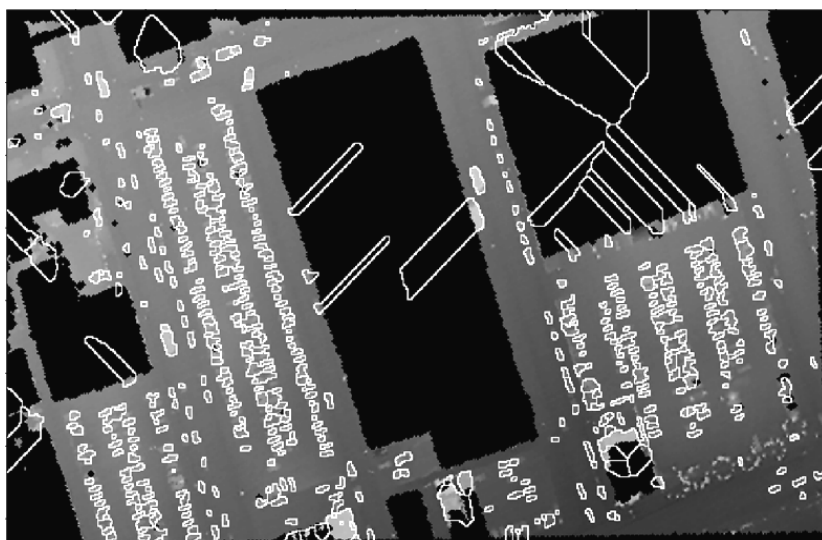
(a)
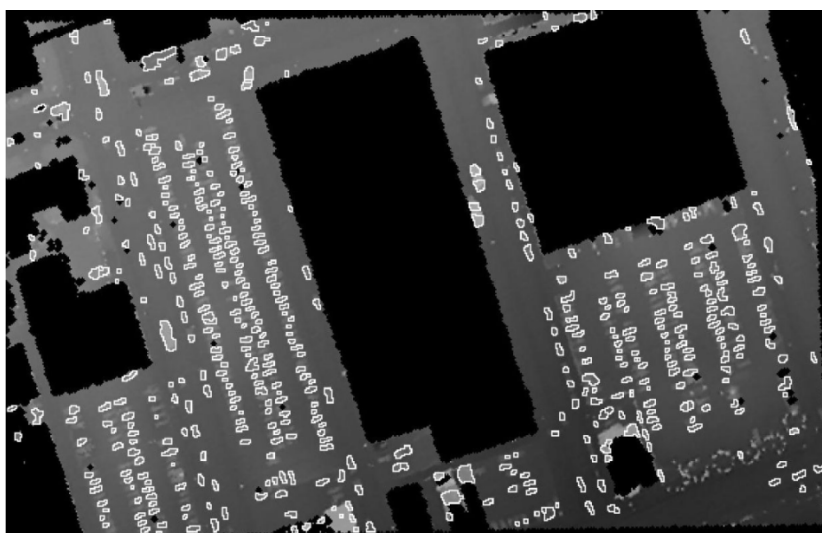
(b)

Figure 14: Vehicle delineation results by marker–controlled watershed transformation (a) without background markers, (b) with background markers

3.2.3 Object-based point cloud analysis method

3.2.3.1 Framework

The adaption of the concept of object-based image analysis framework (OBIA) to the LiDAR data analysis is firstly restricted to approaches using rasterized 2.5D ALS data, which greatly facilitate the application of image analysis algorithms. However, ALS systems deliver (X, Y, Z) coordinates of the unevenly sampled points over the scene and their reflected intensities as measurement. To maintain the full resolution information in ALS data and to avoid the conversion to a 2.5D model, the object based point cloud analysis (OBPA) concept is introduced for object extraction from the raw product of LiDAR sensors – unorganized point cloud (Rutzinger et al., 2008). Based on former studies the OBPA method was further developed and adapted to our task in this work. It is composed of

segmentation (Section 3.2.3.2), object feature derivation and classification (Section 3.2.3.3) and refinement step (Section 3.2.3.4).

Object extraction based on OBPA is substantially a strategy for object recognition by jointly employing segmentation and classification strategy, where the point cloud is firstly partitioned into significative segments to approximate the spatial extension of various object instances and then each segment is assigned a semantic label based on feature values derived on the segment-level. Generally, the point-based classification discriminates the dataset at class-level rather than instance-level due to the strong semantic orientation and the lack of the spatial awareness of the feature space used. That means that different object-instances within the same object-class need to be further separated for the purpose of object extraction. Consequently, the OBPA concept allows extracting various urban targets from ALS data simultaneously, when the 3D segmentation process is capable of giving a significative representation of the scanned scene. The step of segmentation plays a key role and becomes the core part of the whole strategy for extracting objects from ALS data. In addition, the partition of urban scene into various semantic object categories provide the opportunity to make use of local context relations and apply to ground surface segment for anew detecting vehicles located on it. It can be expected that the completeness rate of vehicle extraction could be improved.

Vehicle extraction using the OBPA method is based on an intelligent 3D segmentation approach working directly on point cloud. It deals with semantic inference in point cloud data, which has proved to be a very complicated task (Melzer, 2007). The segmentation process is performed in the sense of local mode detection by adaptive mean shift (MS). It allows a genuine clustering on raw ALS data and does a scale–aware mode-seeking process in the 3D geographic spatial feature space. The obtained segments can give a significative partition of scene to represent vehicle and non-vehicle objects, despite that they look like somehow fragmented. After achieving coherent point segments potentially corresponding to local entities, a binary classification based on Support Vector Machine (SVM) is performed to evaluate shape and positional features of point segments to extract vehicles. Moreover, a modified normalized cut (NCuts) can be used to group point segments to build large-scale objects (such as building, ground surface and road) in a sense of global optimization. For objects classified as ground (road) surface, the context-guided method is to be applied again to extract such vehicles that were missed by the OPBA method in areas of dense placement. Especially for data areas with global contexts in the *Route* 3 (Figure 6) where vehicles are placed very close to each other such as in parking lots, the combination of the OBPA and local context-guided method can make the extraction more successful for the vehicles. The extracted vehicles can be further delivered to motion analysis step based on the shape deformation features to infer the overall traffic situation by determining the motion state of each vehicle.

3.2.3.2   3D segmentation by adaptive mean shift clustering

Mean shift is an extremely versatile tool for feature-space clustering. So far, MS has been successfully applied to image segmentation tasks by exploiting the spectral feature space. The choice of feature-based analysis lies in the recognition that the integration of different cues into the analysis offers advantages of unveiling potential structures of data. As the feature-based analysis depends on the quality and reliability of selected features, the feature computation and selection should play a fundamental role in the design of a segmentation algorithm. Since the derivation of

geometric features such as height textures, planarity and curvature is biased by neighborhood definition and we want to avoid this, the 3D geographic space spanned by (X ,Y ,Z) coordinates of point cloud is chosen here to explicitly represent the feature space which is dedicated to clustering.

Thanks to its non-parametric and mode-seeking mechanism, MS is believed to have great potential with respect to flexibility and reliability in dealing with ALS data of complex urban areas. A remarkable feature is its ability to move the data points towards their respective modes of the empirical distribution. Given n laser data points x_i (i= 1, 2,…,n) in 3-D space, the kernel density estimator at point x can be written as

$$\hat{f}_{h,K}(x) = \frac{c_{k,d}}{nh^3} \sum_{i=1}^{n} K\left(\left\| \frac{x-x_i}{h} \right\|^2 \right)$$

(12)

where $c_{k,d}$ is a normalization constant, h is the bandwidth, and $K(\cdot)$ is the kernel, that models how strongly the data points are taken into account for the estimation. The objective of the feature-space analysis is to find the local maxima of the density $f(x)$, i.e., the modes of the density, which are located among the zeros of the gradient $\nabla f(x) = 0$. The MS procedure is an efficient way to locate these zeros without having to estimate the underlying probability density model before (Comaniciu et al., 2002). The density-gradient estimator can be derived by differentiating Equation(12),

$$\hat{\nabla}f_{h,K}(x) = \frac{2}{ch^2} \hat{f}_{h,G}(x) \cdot m_{h,G}(x)$$

(13)

where the profile of kernel G is defined as $g(x) = -k'(x)$, with c as its normalization parameter. In Equation (13), the first term is the density estimate at x with the kernel G and the second term is the MS

$$m_{h,G}(x) = \frac{\sum_{i=1}^{n} x_i \cdot g\left(\left\| \frac{x-x_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n} g\left(\left\| \frac{x-x_i}{h} \right\|^2 \right)} - x$$

(14)

From Equation (14), it can be found that the MS is the difference between the weighted mean, using the kernel G for weights, and x, the center of the kernel. According to Equation (13), the MS can be written as

$$m_{h,G}(x) = \frac{1}{2} h^2 c \frac{\hat{\nabla}f_{h,K}(x)}{\hat{f}_{h,G}(x)}$$

(15)

Equation (15) indicates that the MS vector at point x with kernel G is proportional to the normalized density-gradient estimate obtained with kernel K, and it thus always points toward the direction of maximum increase in the density. In other words, the local mean is shifted toward the region in which the majority of the points reside (Comaniciu et al., 2002). Once gets sufficiently close to a mode of the estimated density, it converges to it. The set of all locations that converge to the same mode defines the cluster of that model.
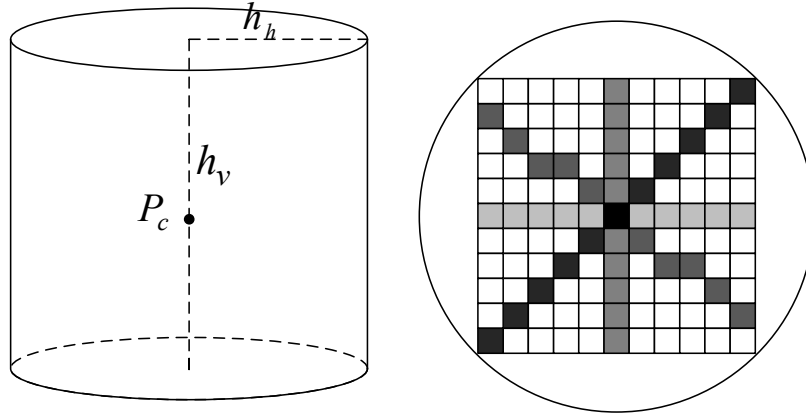
Figure 15: Left: cylindrical kernel for density estimation; Right: direction lines on the tiles of 2D projection of surrounding points around the centric point P_c

Based on the above analysis, MS can easily be adapted to the clustering task for laser point clouds intended for finding various geometric modes. It is intuitive for MS to obtain a genuine 3D clustering on unstructured ALS data without any pre-processing. It can be assumed that the density modes found in laser data tend to emerge around local distinct structures to form weak primitives for scene description. Airborne LiDAR data is represented as a spatial feature space denoting the coordinates and locations for different laser points. The multivariate cylinder-shaped kernel (Figure 15) is defined for density estimation

$$K(x) = \frac{C}{h_v h_h^2} \prod_{u \in \{h,v\}} k\left(\left\| \frac{x_u}{h_u} \right\|^2 \right) \qquad (16)$$

where C is a normalization parameter. h_v and h_h are the kernel bandwidths for vertical and horizontal subdomains. Since the MS feature-space analysis is task-dependent and urban objects stretch across totally different scales, the major challenge for applying the MS algorithm is to adaptively determine the bandwidths in the spatial domains based on different applications. For our application h_h is assumed to vary and need to be determined at each point, whereas h_v is kept as a constant and set to 2m.

There are two schemes of adaptive MS clustering – fixed bandwidth and variable-bandwidth (Comaniciu et al., 2003). One fixed bandwidth across the whole dataset can be determined according to the stability analysis of the decomposition (Fukunaga 1990). It is taken as the center of the largest operating range over which nearly the same number of clusters are obtained for the given data. However, the object-specific knowledge should be considered to control the kernel bandwidth. Therefore, a locally-adaptive MS scheme using variable bandwidth is proposed for airborne LiDAR data segmentation. It is based on exploiting local geometric homogeneity to search for the latent modes in a homogeneous area. Inspired by Zhang et al., (2006), a point shape index (**PSI**) is designed to measure the spatial structures and estimate the spatial bandwidth h_h for each point.

The point cloud is firstly spatially indexed by Geo-tiling process which can establish the corresponding relation between single points and the grid cell (Yao et al., 2010). PSI is computed point by point by extending a number of direction lines radiating from the central point, which are used to detect the object's overall contour. Firstly, the point homogeneity is defined as

$$PH_d(k,x) = \left| p_h^{cen} - p_h^{sur} \right| \tag{17}$$

where PH_i represents the spatial homogeneity of the dth direction between the centric point k and its surrounding point x, $p(\cdot)$ is the height value of the point within the kernel (Figure 15).

Here, we use a hybrid rule that fuses the spatial edge and shape information when the direction lines are being extended to adaptively determine the bandwidth. Edge detection for the resulting height grid data is performed using the Canny filter; it is an optimal edge detector having a low probability of false or missing edges and a high accuracy of edge positioning. After applying the Canny detector, it results in binary edge map where each pixel $edge[x]$ is represented by a value of either 0 (non-edge) or 1 (edge). After introducing the spatial constraint, the condition for extension of direction lines can be formulated as:

$$edge[x] \neq 1 \text{ and } PH_d(k,x) \leq T_1 \text{ and } L_d(k) \leq T_2 \tag{18}$$

where $L_d(k)$ is the length of the dth direction-line, which is measured as $L_d(i) = \max\left\{ \left| y^{\varepsilon 1} - y^{\varepsilon 2} \right|, \left| x^{\varepsilon 1} - x^{\varepsilon 2} \right| \right\}$. ($x^{\varepsilon 1}$, $y^{\varepsilon 1}$) denotes the planimetric coordinate of the point in one end of the direction line, and ($x^{\varepsilon 2}$, $y^{\varepsilon 2}$) denotes the planimetric coordinates of the point in the other end. T_1 is a pre-defined threshold for $PH_d(k,x)$ and T_2 is maximal number of points allowed in this direction line. If $edge[x] = 1$, it may show an obvious height discontinuity in the local area and hence the extension of direction-lines should be terminated there.

The PSI-based bandwidth estimation can be obtained by

$$2h_h(x) + 1 = PSI(x) \tag{19}$$

where $h_h(x)$ denotes the horizontal bandwidth for each point x, $PSI(x) = \frac{1}{D}\sum_{i=1}^{D} Ld_i$ which aims to examine the shape index of the centric point and is the averaged length of D-direction lines, and its value represents the spatial dimensions of groups of spatially related points. In such way, the horizontal spatial scale is determined according to the averaged diameter of the homogeneous region around the current point.

(a)

(b)

(c)

(d)
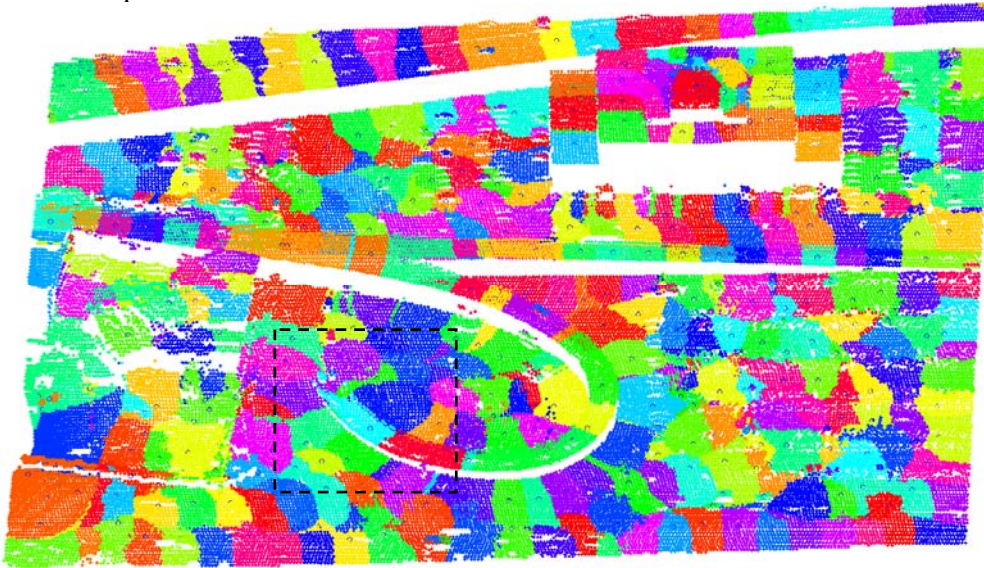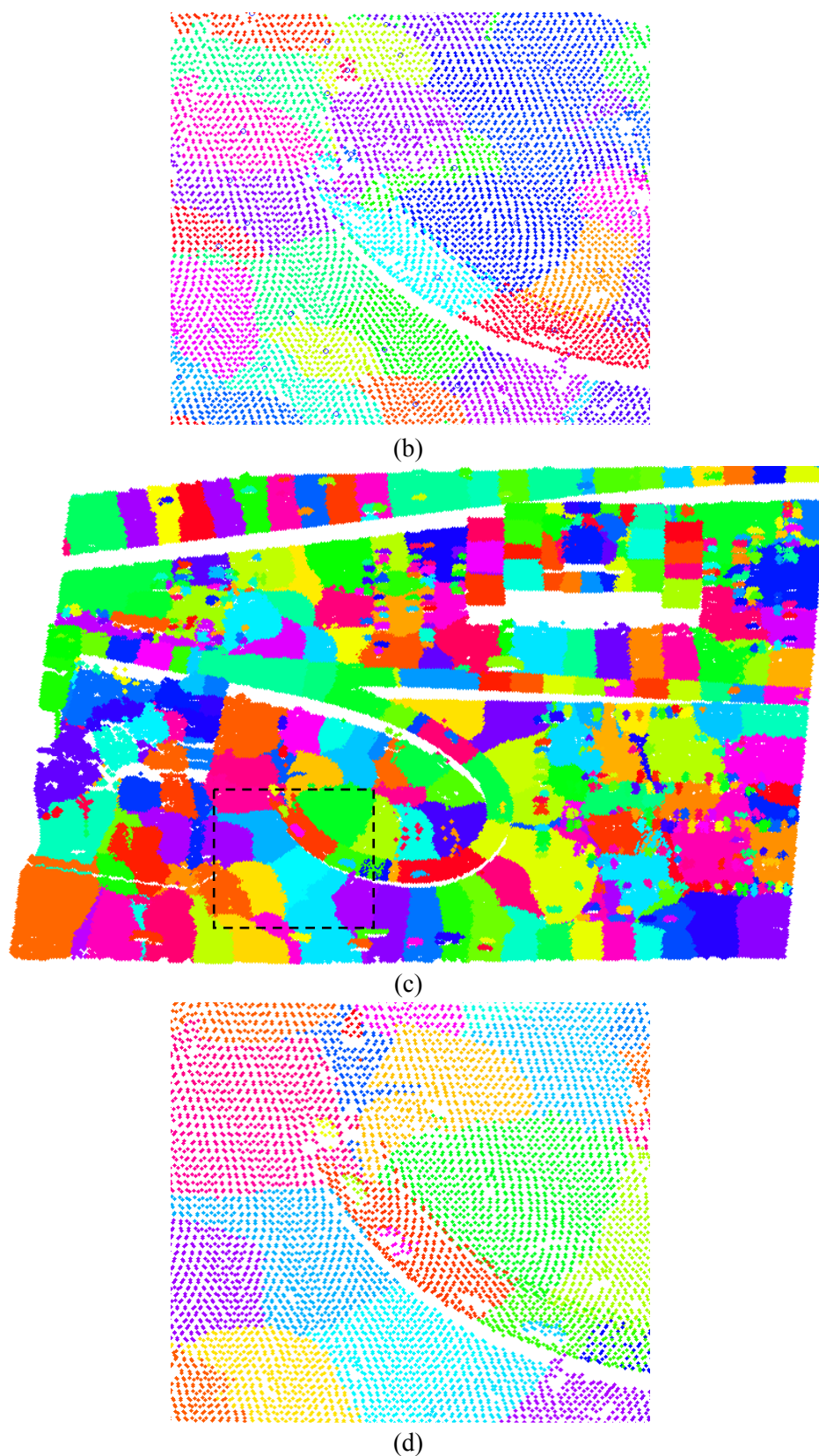
Figure 16: 3D segmentation of an ALS urban data by adaptive MS. (a) fixed-bandwidth (11m) MS, the number of objects = 592; the silhouette S =0.343; (b) zoom into the box areas of (a); (c) variable-bandwidth MS. The number of object = 478; the silhouette S =0.745; (d) zoom into the box areas of (c.)

The results of applying two schemes of MS clustering (*T1*=1.5 *T2*=25) to segmenting an example dataset are shown in Figure 16. First impression by visual analysis has inferred that the variable-bandwidth MS clustering using PSI gave a more natural and compact way concerning partition of the point cloud. The result obtained from the fixed MS analysis seems to be more fragmented in terms of representing significative objects, while the local details such as vehicle objects cannot be detected very well too (Figure 16(a) and (b)). It can be further seen in the enlargement of local section (Figure 16(b) and (d)) that several vehicles above the exit of the overpass can only be detected by variable-bandwidth MS. Furthermore, a clustering quality criterion named as silhouette *S* is used to quantitatively assess the intermediate segmentation results generated and thus enable us to gain an insight into the segmentation performance of AMS schemes from the viewpoint of clustering. The silhouette *S* is defined as

$$S = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n_j}\sum_{i:x_i \in C_j} S(i), \quad \text{where } S(i) = \frac{(b(i)-a(i))}{\max\{a(i),b(i)\}} \tag{20}$$

where $a(i)$ is average dissimilarity of *i*-point to all other points in the same cluster; $b(i)$ is minimum of average dissimilarity of *i*-point to all points in other cluster. *m* is the number of clusters while n_j is the number of points in the cluster C_j. The value of *S* ranges from -1 to 1. If silhouette value is close to 1, it means that dataset is well-clustered.
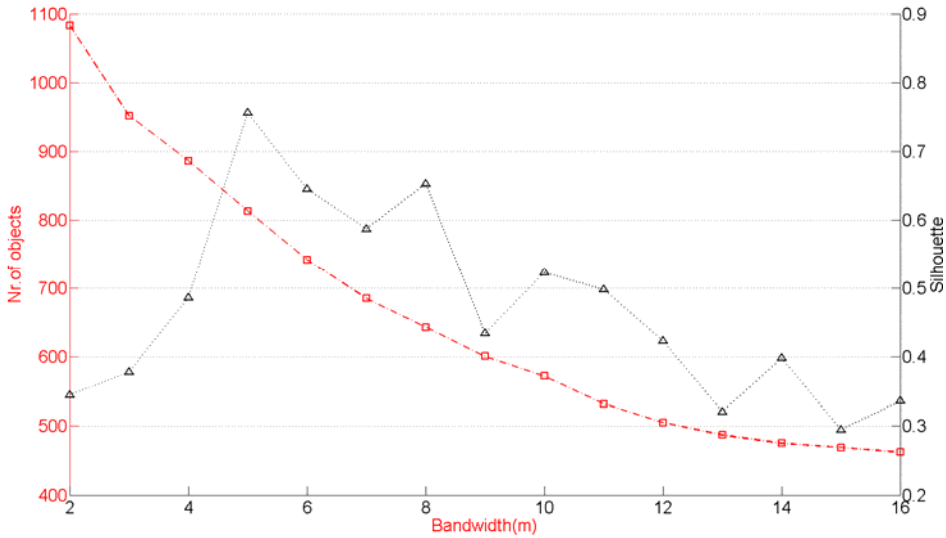
Figure 17: Plot of object number & validity as a function of the bandwidth for the fixed-bandwidth MS analysis on the same data as displayed in Figure 16

For the fixed–bandwidth MS analysis, the obtained number of objects and the silhouette *S* respectively as a function of different bandwidths are plotted together in Figure 17. It can be seen that the compromise between the clustering quality and the number of clusters cannot be optimally achieved by the fixed–bandwidth MS analysis. The number of objects is far beyond the reasonable number of existent objects in the scene. On the contrary, the variable-bandwidth MS analysis based on PSI can give a more sensible partition of the scene with both rational number of objects and acceptable clustering accuracy. Adaptive MS segmentation enables the points to cluster and form large semantic objects while keeping small ones left alone. Consequently, large segments can be used to describe significative and distinct parts of large-size semantic objects such as ground,

flyover and building, while small-size segments towards representing local details like vehicles are also preserved.

3.2.3.3   Classification of point segments

Classification is performed on the point segments resulting from the segmentation step. The objective is to assign semantic labels to data for extracting desirable urban objects. By adopting the point segments as the unit of classification, a large set of features can be calculated providing comprehensive information about their spatial, topological and contextual properties and the overall number of observations is dramatically reduced for a given dataset. Since urban areas usually feature heterogeneous objects and clutters, it is required to select the optimal object features for the discrimination of urban categories. Five object features were considered as potential sources of information for the classification (Table 1).

Feature	Abbreviation	Description	Formula
Area	*A(i)*	Given by the number of points of which the object *(i)* consists. It is represented as polygon of *n* vertices	$A(i) = \frac{1}{2} \lvert \sum_{k=0}^{K=n} (x_k y_{k+1} - x_{k+1} y_k) \rvert$
Elongatedness	*E(i)*	Ratio of the object *(i)* area and the square of its thickness *d*	$E(i) = \frac{A(i)}{2d^2}$
Planarity	*P(i)*	Eigenvalues analysis on 3x3 covariance matrix of the object *(i)*	$P(i) = \frac{\lambda_2 - \lambda_3}{\lambda_1}$
Vertical position	*Vp(i)*	Height of centroid of the object *(i)*	$Vp(i) = Z(centroid(i))$
Vertical range	*Vr(i)*	Extension of the object *(i)* in the Z-axis	$Vr(i) = \max(Z(i)) - \min(Z(i))$

Table 1: Features defined at object level for classification

In this work SVM is used to interpret and classify the ALS data with MS-based object-oriented feature space. This work utilizes SVM owing to its computational simplicity and superior accuracy when compared to conventional classifiers such as maximal likelihood classifier. It is not constrained to prior assumptions on the distribution of input data and is, hence, well suited for complex feature space. The SVM has advantages in nonlinear seperation problems. Moreover, because of the hybrid feature space and the spatially similar objects in the ALS data, the decision boundary should be nonlinear. The SVM was originally designed for binary classification; therefore, in this work, the binary SVM is directly implemented for our task in a genuine way (vehicle and non-vehicle) without handling the multiclass problem.

A specific application using SVM needs to handle several issues. It is required to normalize all the features into a uniform interval (here is [0, 1]) in such a way that they can then be input into the classifier. The normalization method of spectral inputs differs from that of spatial ones because of the dissimilar range distribution of spatial feature values. The normalization methods can be formulated by the linear stretching:

$$d'_k = \frac{d_k - d_{\min}}{d_{\max} - d_{\min}} \qquad (21)$$

where d_k denotes the original feature value of the point segment k, and accordingly $d_{\min}$, $d_{\max}$ represent the minimum and maximum value in that feature, and d'_k is the feature value after the normalization.

The commonly used kernel functions are the radial basis function (RBF) and the polynomial function (POLY). For classification of VHR satellite images, the POLY kernel was found to be better than RBF, because the POLY kernel is a type of function with overall influence, whereas RBF mainly responds to local structures around the central value (Zhang et al., 2006). In this study, the RBF kernel is used. In addition, the kernel based implementation of SVM involves problems related to the selection of multiple parameters, including the kernel parameters (p) and the regularization parameter C. Some standard methods exist that can facilitate the selection of parameters in the SVM classifier design. In our case, these parameters were selected automatically based on the LOO (leave-one-out) algorithm (Chapelle et al., 2002). It is based on the idea that the expected generalization error is to be minimized where the optimization of the parameters is carried out by a gradient descent search over the parameter space. Thanks to the 3D segmentation process applied previously, where the substantial objects (-parts) are already segmented and delineated very well towards object recognition, the classification step can be realized in an efficient way to achieve promising results of vehicle extraction from complex backgrounds.

3.2.3.4 Refinement

The refinement step presented here allows to combine the OBPA method with the local-context guided method to form final vehicle results (*Route* 3), which are especially desirable for the data areas showing a dense placement of vehicles. The only usage of the OBPA method to such data areas could results in less promising results due to the lack of bandwidth for many vehicles such as those in parking lots.

3.2.3.4.1 Global grouping of initial segments

MS segmentation results in over-segmentation, which are merely meant to approximating the geometric (sub) primitives but can hardly represent perceivable object (-parts) of large-scale in the scene. Consequently, an enhanced segmentation step is introduced in the sense of perceptual organization, attempting to group the point segments generated by MS operation to form large semantic objects while keeping small ones towards local modes left alone. The NCuts method from image segmentation (Shi and Malik, 2000) has been chosen to accomplish the task in the context of global grouping.

Instead of the single points, the point segments produced by MS segmentation become the grouping

unit and can be represented by a region adjacency graph (**RAG**) **G={V, E, W}** with *V* as the point segments representing the nodes and *E* as edges formed between every pair of nodes. The similarity between two nodes $\{i, j\} \in V$ is described by the weights w_{ij} which are derived by features associated with the point segments and construct an edge affinity matrix *W*. The graph can be partitioned into two disjoint sets *A* and *B = V−A* by removing the edges connecting the two parts. This problem can be mathematically formulated as graph-cuts by minimizing the *cut* value.

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \tag{22}$$

NCuts is a new graph-cuts method and can jointly consider the intragroup and the intergroup similarity within the total measure defined as

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \tag{23}$$

where $assoc(A,V)$ denotes the sum of the weights of all edges from nodes in A to all nodes in the graph. The minimization of $Ncut(A,B)$ is solved by the corresponding generalized eigenvalue problem.

One key factor for achieving a good partition performance lies in how to construct **RAG** for nodes by determining the connectivity between every 3D point segments. A key novelty of the proposed approach using the NCuts to group 3D point segments is to introduce a non-planar RAG to describe 3D topological relations and spatial connectivity between every point segments (Figure 18), where the graph edges are divided into two types:

Given the premise: $\min\left\{dist(\|R_i - R_j\|^2)\right\} < DistThresh$

● Horizontal edges (**HE**), which connect two point segments mutually horizontally aligned.

if $N\{R_i, R_j\} \perp \mathbf{C}_{Rij} \wedge acos\left(\frac{N_{Ri} \cdot N_{Rj}}{|N_{Ri}| \cdot |N_{Rj}|}\right) \le 5°$

● Spatial edges (**SE**), which connect two point segments mutually spatially aligned and contain all other edges that cannot fulfill the condition of **HE.**

It is assumed here that the laser data was by MS segmented into *n* regions $R_i, (i = 1, 2, 3..., n)$; $\mathbf{C}_{Rij}$ is the difference vector between the centroids of two point segments; N_{Ri} is the normal vector of point segments; $DistThresh$ is the predefined maximal distance allowed for connecting adjacent nodes and is uniformly set to 1.5σ in this work, where σ is the nominal point spacing of ALS data used.

Usually, NCuts algorithms try to find a "balanced partition" of a weighted graph via recursive cuts, which does not have a bias in favor of cutting small sets of isolated nodes in the graph (Tao et al., 2007). Unfortunately, small objects such as vehicle and tree crown belong to this class, which cannot be easily disconnected from other nodes to maintain as independent objects after graph cuts. To overcome this difficulty, a multi-child nodes strategy is applied to such nodes that each of which

is connected to adjacent nodes by **SE** only to split each of them into multiple child nodes. Most of the isolated nodes corresponding to small objects are spatially aligned with neighboring ones and so connected by **SEs**. The edges connecting all the child nodes to adjacent ones are defined as child edge (**CE**) (Figure 18(b)). The weights between the child nodes within one segment are all one, whereas the weights of edges connecting child nodes to adjacent segments are all the same and equal to the weight between these two regions. This yields a new weight matrix $\mathbf{W}_c = \mathbf{W} \otimes \mathbf{E}_c$, where $\otimes$ denotes the Kronecker product operator, and $\mathbf{E}_c$ is the $c \times c$ matrix with all unit entries. For our case, the number of child nodes for each region is set to $c = 3$.
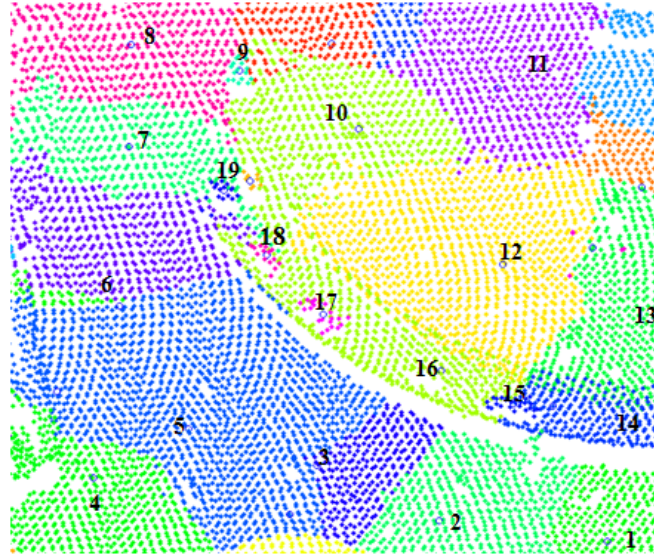
(a)
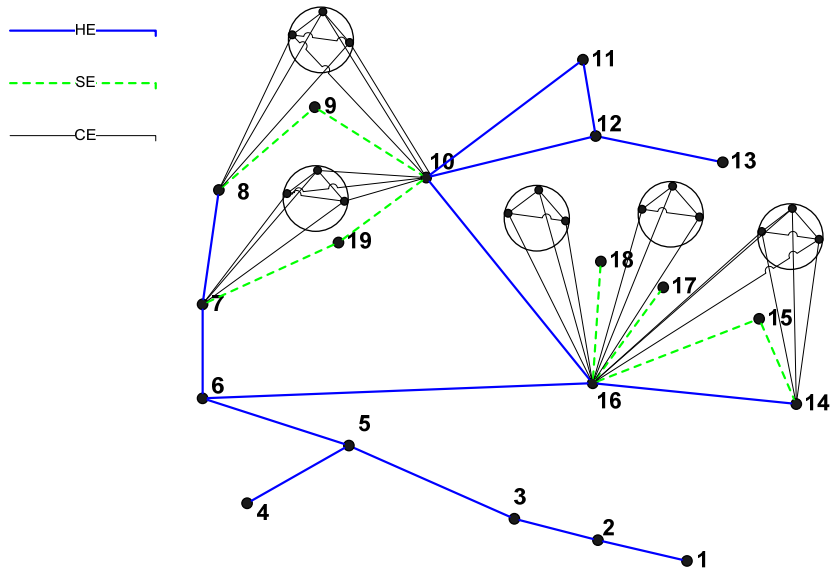
(b)

Figure 18: **RAG** for point segments (a): Numbered point segments for the section (black dotted box) of Figure 16(b), (b): the corresponding non-planar RAG with multi child-nodes (within each black circle, for point segments 9, 15, 17, 18, 19, respectively), each point segment is represented as a node in the graph

The subdivision of the *G* into several segments is realized in the hierarchical procedure:

Step 1: Create the non-planar RAG **G** by constructing two types of edge - **SE** and **HE**

Step 2: Generate multi-child nodes for **G** and compute $\mathbf{W}_c$ for all nodes.

Step 3: Find the solution of the eigenvalue problem and cut the graph *G* into two subgraphs G1 and G2 by binarizing the solution vector.

Step 4: Apply step 3 to the graphs **G1** and **G2**, respectively. Stop if the value for *NCut* exceeds the threshold $NCut_{thres}$.

Another key point is to define similarity measures for adjacent segments. The similarity measures should reflect the likelihood that two segments belong together. The point segments of laser data allow us to use several geometric, shape and physical properties to construct a joint spatial-physical feature space, which can provide the significant and distinguishing features for the segmentation of urban areas. The weight matrix **W** of all point segments can be computed by introducing the weight function $w(m, n)$ to compute the similarities $w_{m,n}$ between two connected segments *m* and *n*,

$$w(m, n) = \begin{cases} e^{-I(m,n)} \cdot e^{-Z(m,n)} \cdot e^{-P(m,n)} \cdot e^{-N(m,n)} & \text{if } m \text{ and } n \text{ are connected} \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

with $I(m, n) = \left(\frac{|i_m - i_n|}{\sigma_i}\right)^2$, $Z(m, n) = \left(\frac{|z_m - z_n|}{\sigma_z}\right)^2$, $P(m, n) = \left(\frac{|p_m - p_n|}{\sigma_p}\right)^2$, $N(m, n) = \left(\frac{\theta(n_m, n_n)}{\sigma_n}\right)^2$

where $I(m, n)$ is the mean intensity difference between the point segments, $Z(m, n)$ is the quadratic vertical distance between the centroids of two point segments, $P(m, n)$ is the planarity difference between two point segments, $N(m, n)$ is the intersection angle between the normal vectors of two point segments. In addition, $\sigma_i, \sigma_z, \sigma_p$ and σ_n are values controlling the sensitivity of four impact terms in Equation (24). The idea is to combine various impact factors into the similarity measure by exponential multiplication.
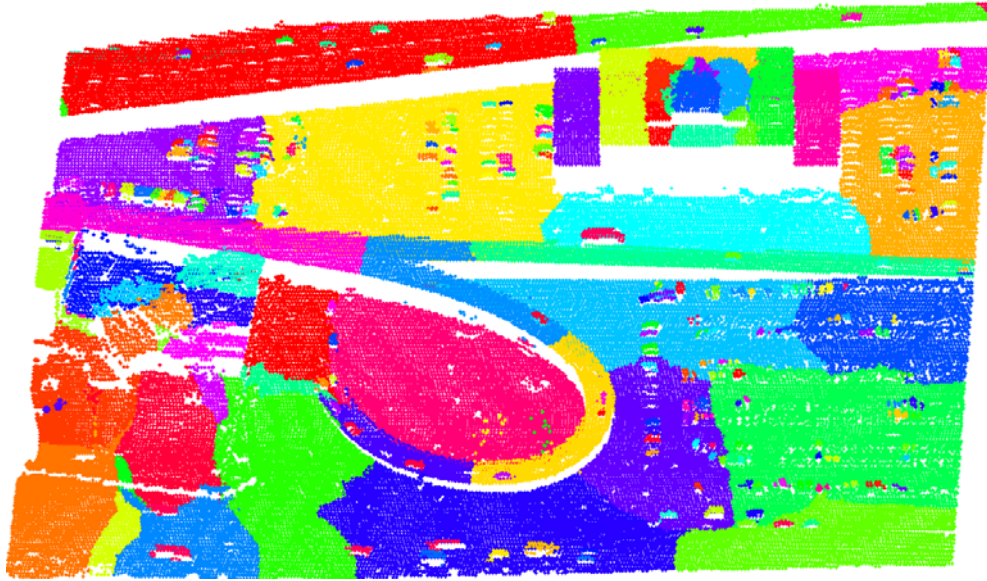
Figure 19: Segmentation result after using the modified Ncuts grouping when $NCut_{thres} = 0.37$

The segmentation result by performing the modified NCuts grouping based on the point segments is depicted in the Figure 19 as an example. Single large segments can be used to describe significative and distinct parts of large-size semantic objects such as ground, flyover and building, while many small-size segments towards local details like vehicles are also preserved.

3.2.3.4.2 Classification of multiple urban objects

A hierarchical rule-based classifier following a course to fine strategy is to be realized in the sense of classification tree. Development of classification trees is based on binary recursive splitting, which recursively partitions the dataset into homogeneous subsets by object attributes. The tree initially has one node, which is called the root node and contains all the observations in the data. The observation sets are divided into two groups whose nodes are each split into two child nodes alternately. The process of binary recursive splitting proceeds until either a maximal tree is built, where each terminal node is pure, or the node contains a number of observations equal to a predefined minimum. For each split, the best predictor variable for assigning the observations to one of the child nodes can be determined using the *Gini* index of impurity defined as

$$i(t) = 1 - p^2(j \mid t) \tag{25}$$

where $p(j \mid t)$ is the proportion of class j at node t.
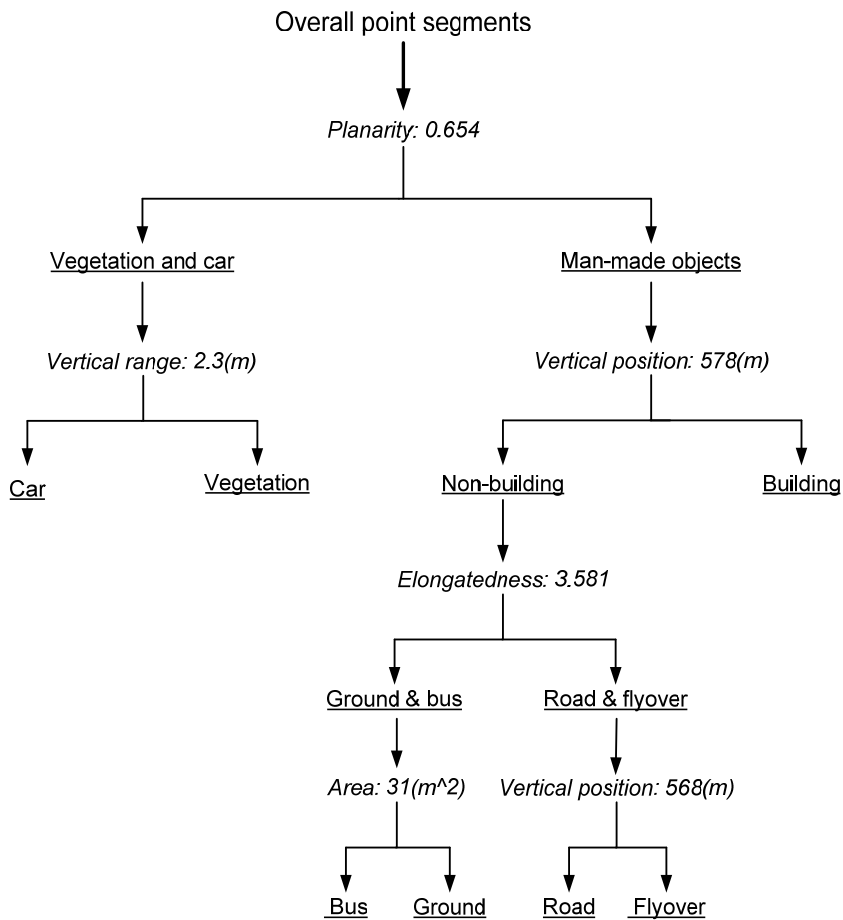
Figure 20: Classification tree for assigning the urban categories to segments in Figure 19, the threshold values were learned by training data extracted from the same dataset

To avoid over-fitting of the tree, a pruning operation is followed. Branches of the generated maximal tree are cut off successively considering the minimum increase in the misclassification. A 5-fold cross-validation procedure is used to assess these errors. The ALS data to be assessed are divided into five sub-samples from which four are used to yield the tree models and the last subsample is used to assess the prediction errors. This process is repeated five times and chooses the optimal tree (Figure 20) as the simplest one among those with average cross-validation errors within one standard deviation of the minimal cross-validation errors (Mallinis et al., 2008). Thanks to the 3D segmentation process proposed above, where the substantial objects (-parts) in the urban scene are delineated very well, the classification step can be realized in a easy way to achieve promising results of object classification and extraction in view of urban scene description. The relevant objects selected from obtained classification results, such as road, flyover and ground, can be further delivered to the local-context guided method to extract vehicles again. Finally, the vehicles extracted by respective methods are fused to yield the

3.3 Vehicle motion analysis

This section is to introduce and investigate the moving object model (motion artifacts effect) in single-path airborne LiDAR data, which lays down the theoretical foundation for monitoring traffic from airborne LiDAR platforms. Furthermore, a scheme is proposed to distinguish the vehicle motion state based on the extracted points of single vehicle instances using a binary shape classifier constrained by Lie group metric. The parameters describing the vehicle geometric property in view of motion artifacts have been estimated in the shape parameterization process applied a-prior, where vehicles of ambiguous shapes are directly removed from the classification step of motion state and labeled as uncertain class. Finally, the velocity of all vehicles of moving class can be quantitatively derived with knowledge about their shape deformations. Additional, the velocity estimation accuracy will be extensively studied and evaluated for different estimation methods under error prorogation principle.

3.3.1 Effects of moving objects in ALS data

3.3.1.1   Model of motion artifacts

In order to assess the feasibility of extracting dynamic information from modern LiDAR sensors installed in airborne platforms, the main characteristics of the sensors, including the data formation and composition method should be considered first. In most of airborne LiDAR scanning processes exclusive of flash LiDAR which are predominantly based on mechanical scanning, a rotating laser pointer rapidly scans the earth's surface with continuous scan angles during its flight. While the sensor is moving it transmits laser pulses at constant intervals given by the Pulse Repetition Frequency (PRF) and receives the echoes. With respect to the objective of extracting moving objects, the fundamental difference between scanning and the frame camera model is the presence of motion artifacts in scanner data. Due to the relatively short sampling time (camera exposure), the frame imagery preserves the shape of moving objects; if the relative speed between the sensor and the object is significant then increased motion blur may occur. In contrast, scanning will always produce motion artifacts, since the distance between sensor and target is usually calculated based on the

stationary-world assumption; significantly moving and even accelerating objects violate this assumption and therefore make the target be imaged "incorrectly" depending on the relative motion of the sensor and the object. The dependencies can be seen by adding the temporal component into the range equation of LiDAR sensor. Here, it is assumed that the sampling rate is consistent among all the vehicles independent of the scan angle. That is to say that all the vehicles are scanned with sufficient sampled points to represent their shape and artifacts.

In Figure 21 the geometry of data acquisition is shown. The sensor is flying in a certain altitude along the dotted arrow called along-track direction. The direction which is oriented perpendicular to the flight path is referred to as across-track direction. An example of generated shape artifacts by moving objects is also depicted in Figure 21, where the black dotted box indicates the vehicle shape obtained in the scanning process of airborne LiDAR while the original vehicle is depicted as rectangle nearby. It can be perceived that the moving vehicle is imaged as stretched parallelogram in the ALS data. Let θ_v be the intersection angle between the moving directions of sensor and vehicle where $\theta_v \in [0°, 360°]$, v_L and v the velocity of aircraft and vehicle respectively, l_s and l_v the sensed and original lengths of vehicle, respectively; and θ_{SA} the shearing angle that accounts the deformation of a rectangle–shaped vehicle as parallelogram. The analytic relations between shape artifacts and object movement parameters can be derived as

$$l_s = \frac{l_v \cdot v_L}{v_L - v \cdot \cos(\theta_v)} = \frac{l_v}{1 - \frac{v}{v_L} \cdot \cos(\theta_v)} \tag{26}$$

$$\theta_{SA} = \arctan\left(\frac{v \cdot \sin(\theta_v)}{v_l - v \cdot \cos(\theta_v)}\right) + 90° \tag{27}$$

where $\theta_{SA} \in (0°\ 180°)$ and is found as the left-bottom angle of observed vehicle. For the sake of full understanding of the appearance of moving objects in airborne LiDAR data, object motions are to be divided into two components and investigated for their respective influences on the data artifacts generated.
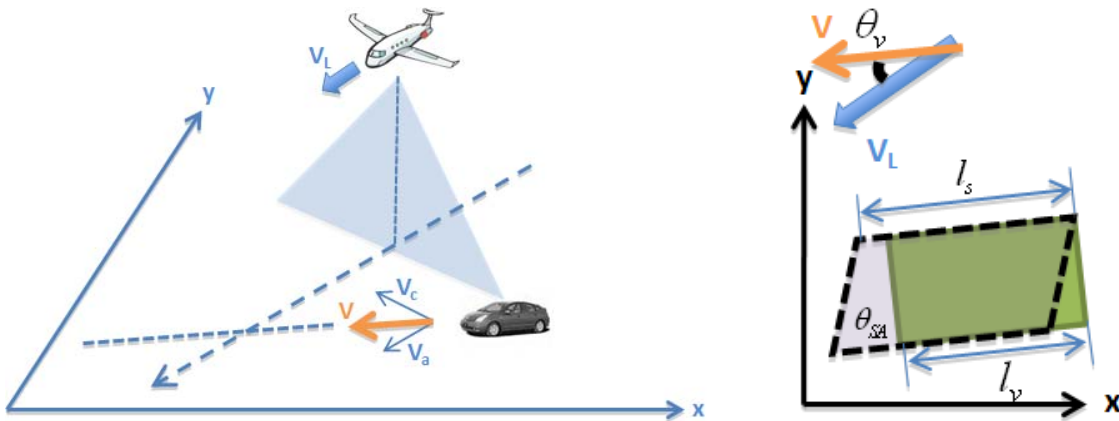
Figure 21: Moving objects undergo the scanning-over of airborne LiDAR

- **Along - track motion**

The target is now assumed to move with constant velocity v_a following the along-track direction.

The along-track motion changes the relative velocity between sensor and object compared to the surrounding stationary world which makes the laser footprints that hit upon target displaced consistently along the target moving direction. Therefore, along-track motion leads to the stretching or squeezing effect of the object shape depending on the relative velocity between target and sensor as illustrated in Figure 22.

The analytic relation between the object velocity in along-track direction v_a and the observed stretched length l_s can be summarized in Equation (28). The fact that v_a is the component of the vehicle velocity along the sensor track direction, not the modulus of $\mathbf{V}_a$, should be noted. Therefore, if $\theta_v = 180°$, it is $l_s < l_v$. Rather than the vehicle length that usually varies a lot and is not able to be exactly known in advance, the aspect ratio of vehicle is assumed to be easily determined due to its relative stability over one vehicle category. For this reason, the relation in Equation (28) is further modified to Equation (29) which explicitly connects v_a with the variation in the aspect ratio of vehicle shape in a mathematical way, thereby making motion state distinction and velocity estimation more feasible and reliable in view of applications in real-life scenes.

$$l_s = \frac{l_v}{1 - \dfrac{v_a}{v_L}} \tag{28}$$

$$Ar_s = \frac{l_s}{w_v} = \frac{Ar}{1 - \dfrac{v_a}{v_L}} \tag{29}$$

where Ar_s is the sensed aspect ratio of vehicle in ALS data while Ar is the original aspect ratio of vehicle.
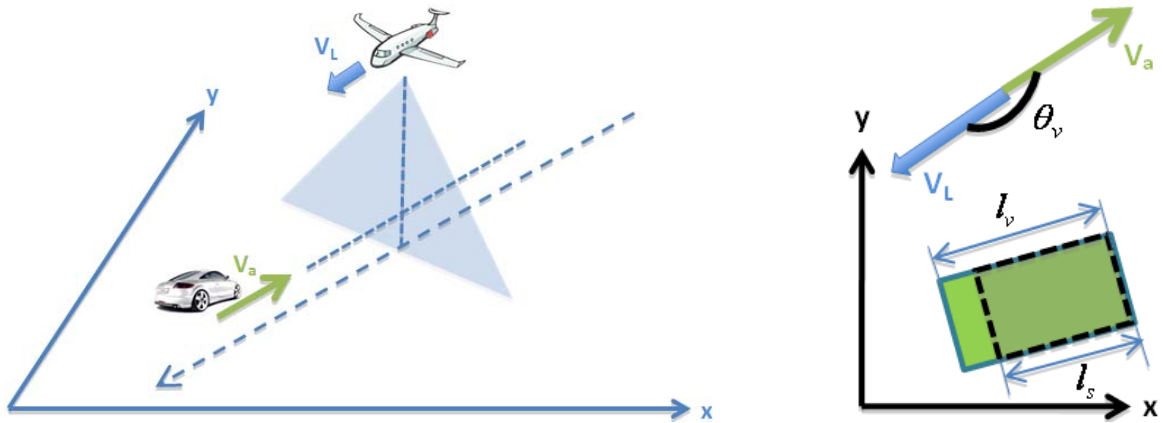
Figure 22: Along - track object motion, left: scanning scene, right: induced motion artifact

- **Across - track motion**

The target is now assumed to move in across-track with a constant velocity v_c. It results in a scanline-wise linear shift of laser footprints that hit upon target towards the moving direction when sensor sweeping over, so that the observed vehicle shape in the ALS data is deformed (sheared) to certain extent depending on the relative velocity between target and sensor. In Figure 23 this is shown.

In general, let v_c be the across - track motion component of the object velocity. Since $v_c =$

$v \cdot \sin(\theta_v)$, Equation (27) can be rewritten as Equation (30) for describing the analytic relation connecting the object velocity v_c with the observed shearing angle θ_{SA} through the sensor velocity v_L and the intersection angle θ_v.

$$\theta_{SA} = \arctan\left(\cfrac{1}{\cfrac{v_L}{v_c} - \cot(\theta_v)} \right) + 90°$$

(30)

The shearing angle θ_{SA} is zero for stationary objects and those moving only in along-track. Moreover, if the vehicle moves only in across-track direction, it is simply $\theta_{SA} = \arctan(v_c / v_L) + 90°$
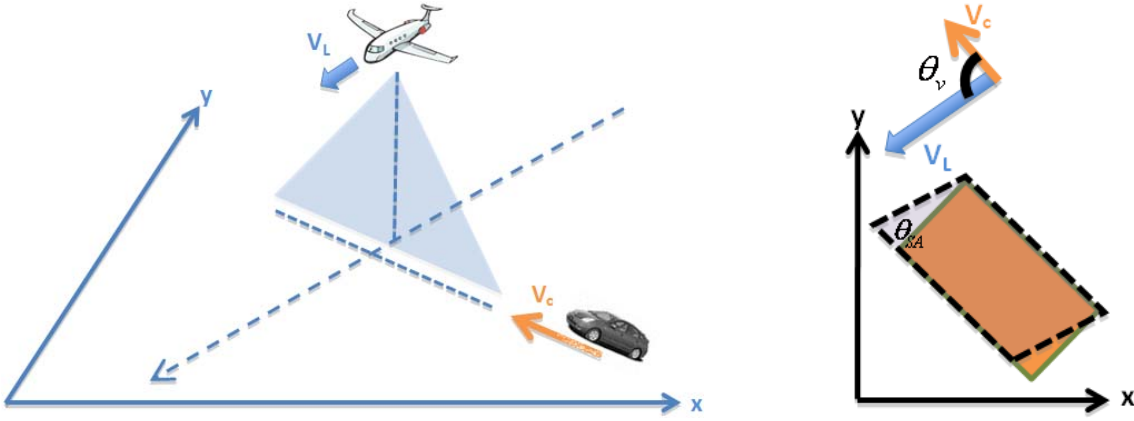
Figure 23: Across - track object motion, left: scanning scene, right: induced motion artifact

- **Acceleration**

In the majority of Ground Motion Target Indication (GMTI) applications except in highways, it is assumed that vehicles travel with constant velocity and along a straight path. However, within the scope of monitoring the traffic in urban areas, the effect of acceleration is weak and can be ignored by quantitative modeling; especially for airborne LiDAR data acquisition where sensor sweeps over vehicle bodies very fast during the scanning process. The produced deformation effects on vehicle point distribution could not play a distinct role. It can be easily seen and verified by substituting the acceleration into the terms in Equations (26) and (27). Theoretically, the acceleration could give rise to irregular shearing of the rectangle shape of vehicle in across-track direction if the object velocity is sufficiently high at the same time. As a consequence, actually occurring minor accelerations might cause an error in velocity estimates depending on the acceleration component. It can also be divided into two components as the object velocity. The along-track acceleration a_x appears in the quasi-linear term of the deformation equations depending on the quantity of a_x and could result in a nonlinearity of the vehicle shape stretching. For airborne LiDAR the deformation of the shape stretching effect is small even for strong and unrealistic a_x. The acceleration in across-track direction (a_y) causes a quadratic component in the range equation which results in a twist of the vehicle long/wide side (i.e. it happens when moving on curved roads and leads to the centripetal acceleration). Considering the common system parameters for city surveying tasks the amount of object shape distortion in airborne LiDAR data could be approximately calculated as a function of across-track acceleration a_y. It can be inferred that the irregular vehicle shape deformation due to the across-track accelerations is significant only for $a_y > 2 \, m/s^{-2}$. Actually, it could even make the deformation effects too complex to be modeled in analytic form in a consistent way and could

undermine the accuracy of velocity estimation. For typical accelerations in common traffic scenarios ($a < 2\,m/s^{-2}$) the effect is assumed to be almost negligible according to the empirical study.

3.3.1.2 Quantification of the effects of motion components

The above-mentioned effects can be exemplified for the case of a concrete airborne LiDAR configuration which is designed and applied to urban surveying task. Based on the formulas derived above, and by using the system parameters as specified in first row of Table 2, the quantitative impact of moving objects in ALS data can be estimated. These investigations help to derive boundary conditions for building up a strategy for detecting ground moving targets and estimating their velocities accurately using airborne LiDAR data which might be originally dedicated to a topographical mapping task.

Along-track motion: The effect of stretching of the sampled target points due to the along-track component of the target velocity is calculated based on Equation (28) by using the target velocity v instead of the along-track velocity component v_a. The sensed aspect ratio of a moving object in an ALS point cloud is shown in Figure 24 as a 3-D function of the intersection angle θ_v and the vehicle velocity v where the original aspect ratio of vehicle Ar is assumed to be 2 for this simulation case. As can be seen, moving vehicles are stretched (elongated or shortened) significantly from their real shape even for moderate along-track velocities (e.g. ca. 60% for 50 km/h). This effect dose not hamper the recognition of cars from ALS data, since their stretched shapes are still related to semantic information and do not change essential geometric characteristics of vehicle used to distinguish it from other semantic objects in urban areas . Figure 24 also shows that the sensed aspect ratio Ar_s is becoming an ambiguous value for indicating the vehicle motion when the intersection angle θ_v equals to or approaches $90°$ or $270°$ regardless of the target velocity. The reason is that the stretching effect on the vehicle shape caused by the along-track velocity is almost zero under this condition where vehicles are travelling in across-track direction.

| (a) | (b) |

Figure 24: Stretching effect of a moving object in ALS data, (a) as function of θ_v and v while v_L=100 km/h, (b) contour plot of (a)

The Ar_s f a moving target is symmetrically distributed around the axis of θ_v=180° with a moderate

basin in the middle being inclined towards the direction where the target velocity v increses. As observed in the right corners of Figure 24(b), Ar_s has increased acutely as the velocity ratio of sensor flight to target comes near to 1. In this case, it is clear that motion indication become easier and meanwhile the velocity estimation could be more sensitive to noise or errors accompanying the determination of shape parameters .Moreover; the along-track velocity component of object can raise the ambiguity limit with respect to motion estimation, since the original aspect ratio Ar of vehicle cannot accurately be determined for single vehicles beforehand. Figure 25 depicts the same function plot as in Figure 24 in the polar coordinate system but depending only on one varaible – intersection angle θ_v. The velocity ratio of sensor flight to moving target v_L/v must be specified in advance and has been selected as two different values. It even leads to giving more illustrative information which also confirmed the statement summarized above. In addition, it has further proved that the if the velocity of moving target is getting closer to that of sensor flight the stretching effect on the scanned vehicle shape in ALS data is becoming more distinct and hence make it beneficial to motion estimation tasks.
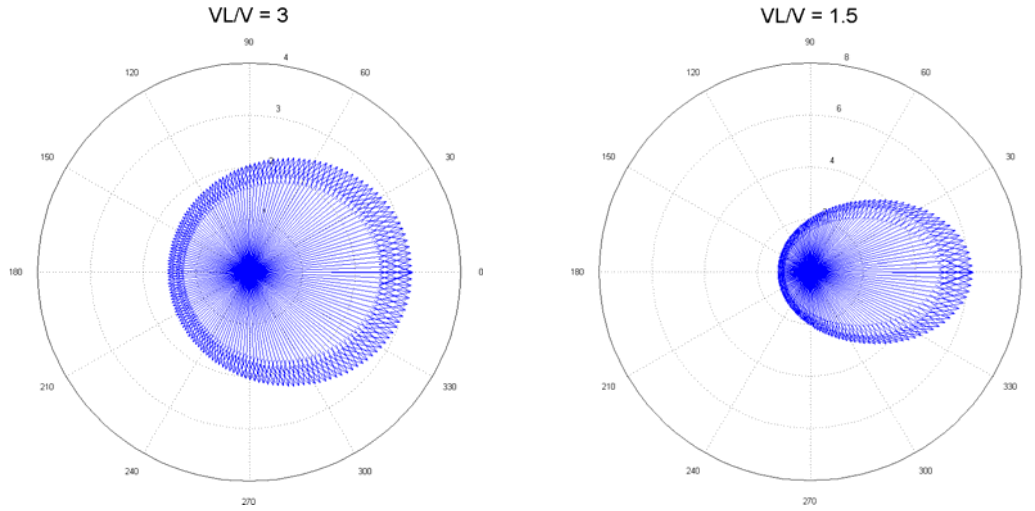
Figure 25: Visualization of the sensed aspect ratio Ar_s in polar coordinate system as the function of the intersection angle θ_v as the velocity ratio of sensor flight to moving target v_L/v changes from 3 to 1.5

Across-track motion: The shearing effect on the shape of vehicle points is triggered by the across-track component of the target velocity. Based on Equation (27) or (30) the shearing angle θ_{SA} of a target moving with velocity v can be calculated with respect to different intersection angles of θ_v. The results are presented in Figure 26 as a 3D function of intersection angle θ_v and vehicle velocity v.

As can be seen, the effect of across-track object motion has significant influence on the geometric appearance of vehicles within ALS data. Moving vehicles are deformed (sheared) to parallelogram from their real shape (given as rectangle-shaped) even for small across-track velocities (e.g. by ca. $25°$ for velocity of 30 km/h). The distribution of shearing measure is inversely (odd) symmetric around the axis of $\theta_v = 180°$. The shearing effect on vehicle shape facilitates motion indication and allows estimating the across-track velocity (also the velocity of vehicle if the intersection angle θ_v is known beforehand) without ambiguity, there by demonstrating the superiority to the method for along-track motion estimation. The minimal value of the measure for shearing effect happens in

such cases that vehicles travel purely in along-track directions. This can be easily comprehended and corresponds to the ideal obtained by intuitive study. Nevertheless, it is worth to note that the occurrence of maximal shearing effect shifts from 90° to ca. 40° for the intersection angle θ_v between the moving paths of sensor and target as the target's velocity increases from zero to approach that one of LiDAR sensor; this trend can be equivalently perceived from a plot of the 2-D function relating the shearing effect to the intersection angle which is deemed to explain the situation in a more illustrative way (Figure 27). It can also be found out clearly from Figure 27 that the maximum value of the shearing effect will intensify while shifting the moving direction from across-track to along-track (e.g. θ_v vary from 90° to 0°). Consequently, for detecting all ground moving targets which suffer from this deformation effect, the theoretic analysis on quantification results such as depicted in Figure 26 and Figure 24 could provide certain guidance on the planning of flight path and velocity of airborne sensor depending on the target velocities needed to be estimated.

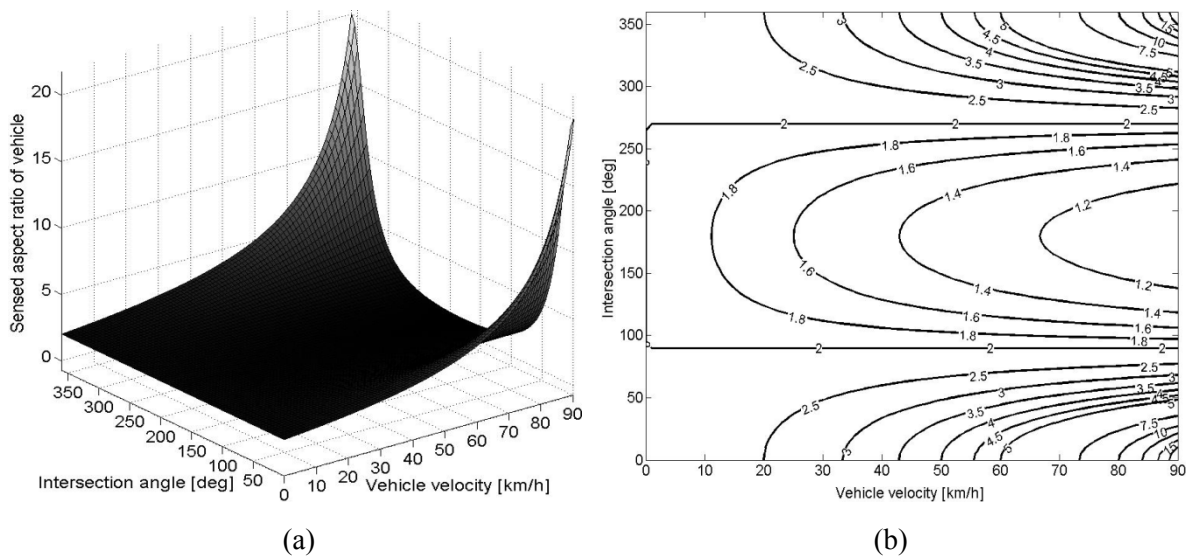| (a) | (b) |
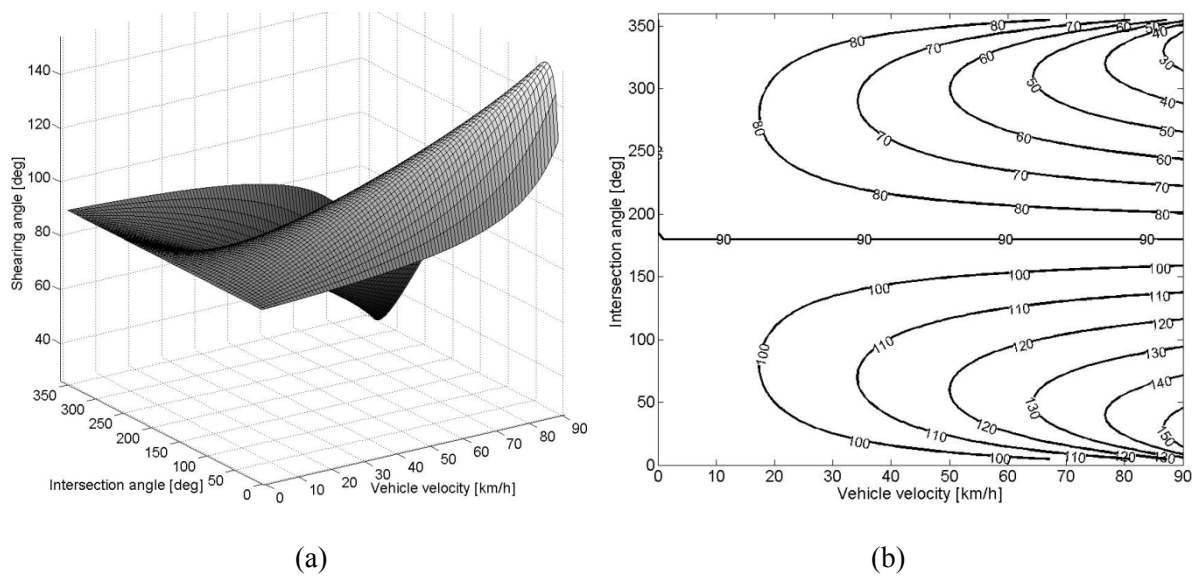
Figure 26: Shearing effect of a moving object in ALS data, (a) as function of θ_v and v while v_L=100km/h, (b) contour plot of (a)
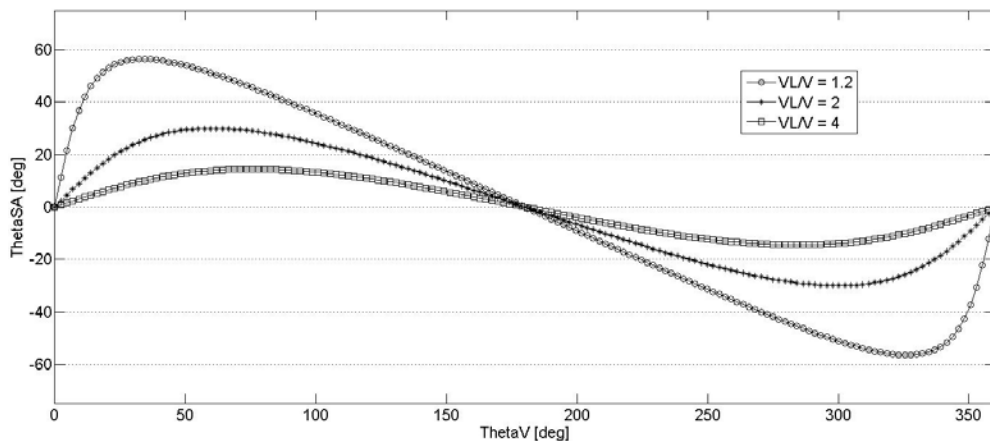
Figure 27: Shearing effect of a moving object in ALS data as function of θ_v, when the velocity ratio of sensor flight to moving target v_L/v is fixed. Lines of different styles depict different velocity ratios used for simulation

3.3.2 Vehicle motion classification

3.3.2.1 Vehicle shape parameterization

The whole strategy proposed in this thesis to detect the vehicles' motion state and estimate their velocities follows the strategy of detecting all the vehicles in advance which has been already achieved by the approach presented in Section 3.2. The separation of vehicle detection from vehicle motion analysis enables us to detect not only moving vehicles but also stationary vehicles thereby giving an overview on the traffic scenario of large areas scanned by airborne LiDAR.

It is desirable for the task of vehicle motion classification to represent the original shape of vehicles and their motion artifacts by a unified model. The spoke model for a set of vehicle points (Figure 31) is proposed to be used as the general framework for modeling the vehicle shape due to its flexibility and efficiency. For the spoke model, the point cloud of single vehicles is fitted with multiple connected planes which are used to describe the vehicle shape by two controlling parameters of each spoke - the orientation and radius. The spoke model could consistently encode geometric information for the classification of vehicle category (Yarlagadda et al., 2008). However, the laser point density acquired under common conditions towards the surveying task is usually insufficient for modeling the vertical profile of vehicle body. The situation is even degraded by the appearance of motion artifacts and the variation of scan angle of LiDAR sensor. Since the motion artifacts generated for moving objects focus on the horizontal deformation of object's shape, the spoke model of vehicles can be projected from original 3-dimension onto the horizontal plane to avoid the redundancy and unnecessary errors. Due to the limitation on the point density of ALS data, all vehicle categories might be modeled with one spoke plane in most real cases. Then, the shape parameterization is performed on the projected point sets of the vehicle spoke model by following two steps:

- **Boundary tracing**: A modified convex-hull algorithm (Jarvis 1977) is used to determine the boundary of the point set of vehicle. Based on the classical convex-hull algorithm, the modification is made upon constraining the searching space of a convex hull formation to a local neighborhood which is defined beforehand. The study showed that the approach can yield satisfactory results if the point distribution is consistent throughout the whole dataset. Such condition could be fulfilled for the single-path ALS data which are especially considered for analyzing the object motion. The boundary tracing method for a set of vehicle points using the modified convex-hull starts with the left - most point P. Then, we use the convex hull algorithm to find the next boundary point by only considering the points within the neighborhood of P, which is defined as a rectangle region with the size corresponding to the point spacing of ALS data, being slightly larger than twice of the point spacing in the along and across scan directions. In this way, only immediate adjacent points at about one ground-spacing are taken into account for the convex hull algorithm, such that a compact boundary is found. The approach proceeds to the newly found boundary point and repeat the step mentioned above until the point P is selected again, as depicted in Figure 28.
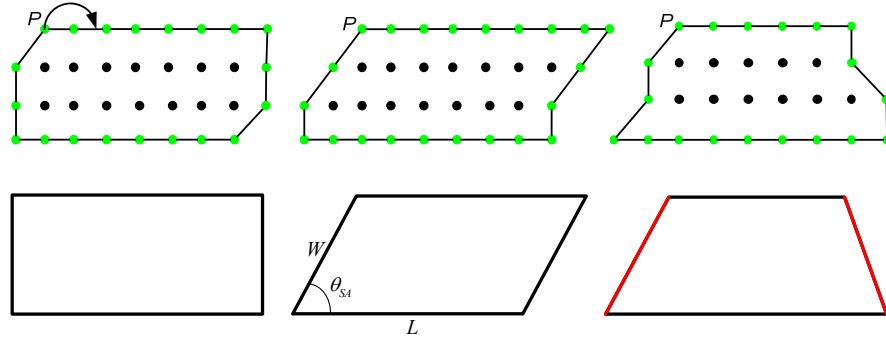
Figure 28: Vehicle shape parameterization. From left to right: stationary vehicle, moving vehicle, vehicle of ambiguous shape. Green points mark the boundary of extracted vehicle; red lines indicate the non-parallel sides of the fitted shape

- **Boundary regularization:** Since the uneven point sampling is always present in ALS data, the traced boundary cannot be directly used to parameterize the vehicle shape due to its irregular shape and possible artifacts introduced by the former step. The two step boundary regularization is introduced to tackle the problem based on the fact that the vehicle point sets should appear as parallelogram having two mutually perpendicular orientations of boundaries. The first step is to extract the points that lie on line segments with identical directions. This is done by sequentially following the boundary points and locating positions where the slopes of two consecutive edges significantly differ. Points on these edges are grouped to one line segment. Therefore, a set of line segments $\{l_1, l_2, ..., l_n, n \geq 4\}$ from which four longest line segments $\{L_1, L_2, L_3, L_4\}$ are selected. Each of the selected line segments is fitted by the straight line. Based on the slope $M_i = -A_i/B_i$ obtained for each line segment, they are sorted into different groups, each of which contains quasi-parallel line segments. In this way, the long line segments of a vehicle boundary are grouped into two "horizontal" and "vertical" groups based on their slopes. The next step is to determine the least squares parallelogram fitting to these line segments under the constraint that the lines segments are parallel to each other within one group. The solution consists of sets of parameters required to describe four line segments, which are formulated as following line equations:

$$A_i x_j + B_i y_j + 1 = 0 \qquad i = 1,2,3,4; j = j(i) = 1,2,3, \dots \ n_i \qquad (31)$$

with the condition: $M_1 = M_3$ and $M_2 = M_4 \ \Leftrightarrow \ L_1 \ (L_2)$ and $L_3 \ (L_4)$ are opposite sides, where n_i is the number of points of each line segment. Once the spoke model of a set of vehicle points is parameterized (Figure 28, bottom row), two parameters describing the vehicle's shape can be derived:

1. The angle of shear θ_{SA} of vehicle:

$$\theta_{SA} = \arctan\left(\left|\frac{M_2 - M_1}{1 + M_1 \cdot M_2}\right|\right) \qquad (32)$$

2. The aspect ratio Ar of the parameterized vehicle points:

$$Ar = L/W \qquad (33)$$

where M_1, M_2 are slopes of line segments of each group.
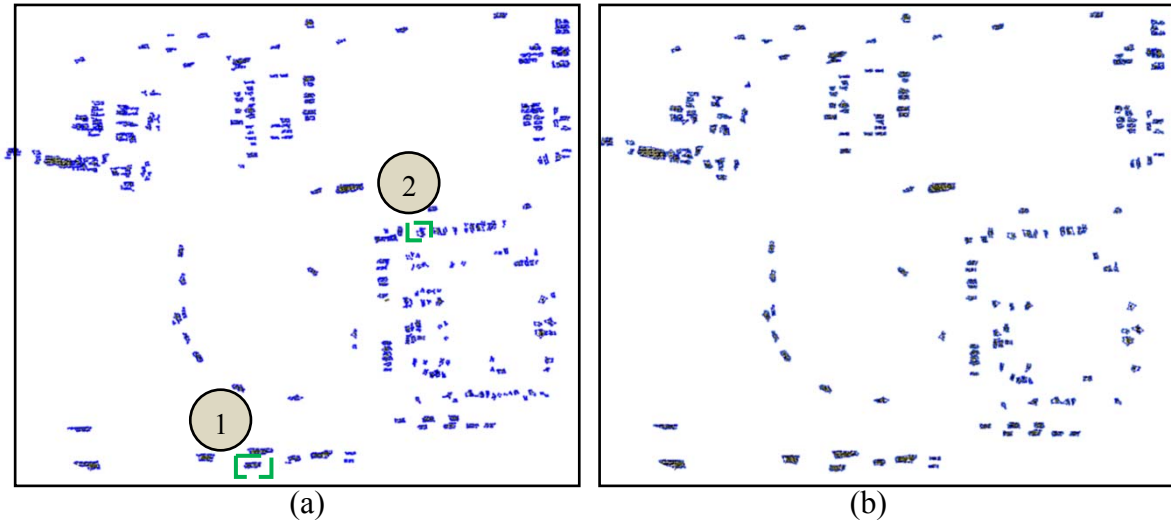
Figure 29: Example for vehicle shape parameterization: (a) boundary tracing, (b) boundary regularization
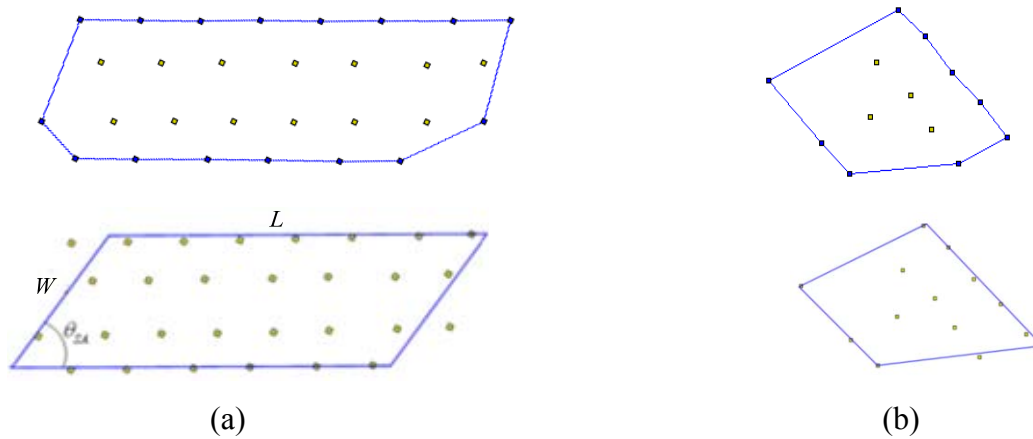
Figure 30: Zoom into the shape parameterization results of labeled single vehicles of Figure 29, (a) vehicle #1, (b) vehicle #2, top row is traced boundary while bottom row is regularized boundary

For certain extracted vehicle point sets the parallelogram fitting could fail in the irregular shapes of boundary due to unstable sampling of LiDAR points or vehicle extraction inaccuracies. These vehicle points may emerge as e.g. trapezoid (Figure 28, third column) other than parallelogram. Such vehicle points deliver ambiguous shape features and make the motion state determination troublesome within the framework of the proposed motion artifact model, because they could e.g. belong to a stationary vehicle with missing parts but were misclassified as moving vehicle. Therefore, this category of vehicle points has to be excluded and attributed to uncertain motion status prior to the motion classification step. An example of applying the vehicle shape parameterization to extracted vehicles from real ALS data is shown in Figure 29. Figure 30 further illustrates the enlargement of two labeled single vehicles in Figure 29, where the vehicle I is a moving one that is parameterized as parallelogram and vehicle II is parameterized as trapezoid and deemed to have uncertain motion state despite that it is actually motionless in a parking lot.

3.3.2.2 Distinction of motion state by shape classification

Once the vehicle shape parameterization is completed, a set of vehicle points can be geometrically

described as spoke model with two parameters, whose configuration can be formulated as

$$X = \begin{pmatrix} U_1 \\ . \\ . \\ . \\ U_k \end{pmatrix}, \qquad U_i = \begin{pmatrix} \theta_{SA}^i \\ Ar_i \end{pmatrix} \tag{34}$$

where k denotes the number of spokes in the model. The vehicle shape variability caused by the motion artifacts effect is of nonlinear nature and thereby can be represented as a transformation space. To this end, the geometric similarity between vehicle instances can be measured by group distance metric. The classification framework for distinguishing different vehicle categories (pick-up, off-road, passenger car or SUV) based on examining the geometric features can be easily adapted to motion classification schema equivalently based on shape features extracted for each set of vehicle points.
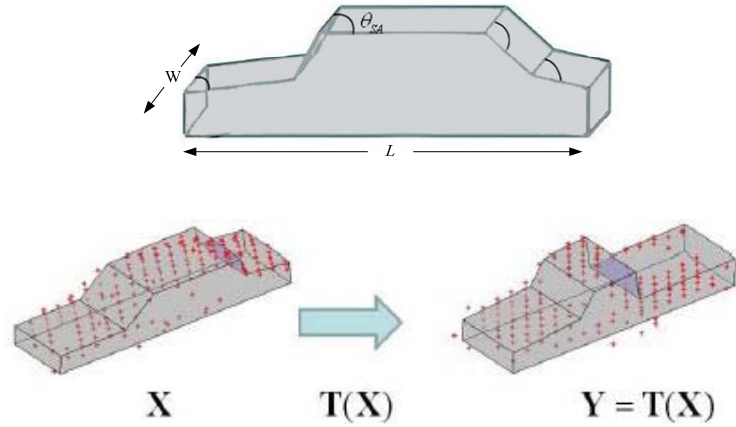
Figure 31: Vehicle spoke model and shape transformation between passenger car and pick-up (modified from Yarlagadda et al., (2008))

Consequently, a new vehicle shape configuration Y can be obtained by a transformation of a original vehicle configuration X written in matrix form: $Y = T(X)$, where $T = \begin{pmatrix} M_1 & . & 0 \\ . & . & . \\ 0 & . & M_k \end{pmatrix}$, $M_i = \begin{pmatrix} R_i & 0 \\ 0 & e^{ai} \end{pmatrix}$,

R_i denotes the 2-D rotation acting on the angle of shear θ_{SA}, e^{ai} denotes the scale acting on Ar. By varying T, any vehicle shape configuration X (motion state) may be represented in T as the transformation of an identity atom (Figure 31). Since the Cartesian product of Lie group elements is a Lie group and T is the Cartesian product of transformation matrices M_i acting on the individual spokes, T forms a Lie group G (Rossman 2002). The intrinsic mean μ of a set of transformation matrices $T_1, T_2, \ldots, T_n$ of vehicle shapes is defined as

$$\mu = \arg \min \sum_{k=1}^{n} d(T, T_k)^2 \tag{35}$$

where argmin(d) indicates the argument of the minimum of d, $d(\cdot, \cdot)$ denotes Riemannian distance and $d(T_1, T_2) = \left\| \log(T_1^{-1} T_2) \right\|$, where $\|\cdot\|$ is the Frobenius norm of the algebra elements, i.e. $\|T\| = \sqrt{Tr(TT^H)}$ with Tr trace of a matrix and T^H indicates the conjugate transpose of T. Analogous

to the principle components of a vector space, there exists 1-parameter sub-group $H_v(t)$ called the principle geodesic curve (Fletcher et al., 2003) which describes the essential variability of the data points lying on the manifold. The parameter t sweeps out a 1-parameter sub-group $H_v(t)$ of the Lie group G. For any $g \in G$, the distance between g and $H_v(t)$ is defined as

$$d(g, H_v) = \min\{d(g, \exp[A_v(t)]), \; t \in \Re\} \tag{36}$$

The first principle geodesic curve for elements of a Lie group G is defined as the 1-parameter subgroup $H_{v(1)}(t)$, where

$$v^{(1)} = \arg\min \sum_{i=1}^{n} d(\mu^{-1}g_i, H_v)^2 \tag{37}$$

Let $p_{i,1}$ be the projection of $\mu^{-1}g_i$ on and define $g_i^{(1)} = p_{i,1}^{-1}\mu^{-1}g_i$. The higher k-th principle geodesic curve can be determined recursively based on Equation (37). It has been confirmed that principle component analysis on the Lie group, i.e. principle geodesic analysis (PGA) on the manifold, can better describe the variability of data that is inherently nonlinear (Fletcher et al., 2003). Figure 32 provides an intuitive schematic depiction of the beneficial effect toward the classification that the PGA could bring. Considering the feature spaces (Figure 32) as visualized in the 2D linear Euclidean space, by means of the PGA all the feature points representing vehicle shapes of two motion states are mapped into a new feature space with such configuration that tend to demonstrates better separability with quasi-linear boundary between two classes instead of the mutually interlaced distribution without clear boundary.

(a)   (b)
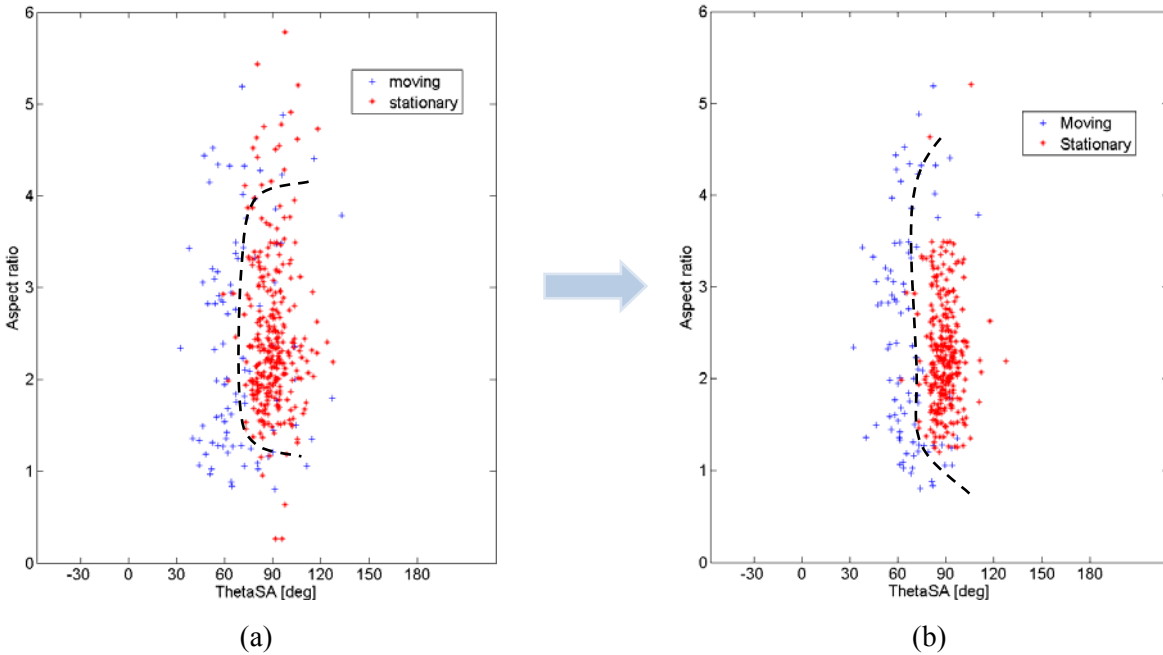
Figure 32: Transformation of a feature space of vehicle shapes of two motion states using PGA: (a) original feature space, (b) feature space after PGA. The shearing angle θ_{SA} for moving vehicle class is normalized into the acute angle range ($<=$ 90°) for simplicity. Black dotted lines indicate the possible lines of separation

The motion distinction task can then be formulated as a binary classification problem. The input to the Lie distance classifier comprises a set of labeled samples from two vehicle categories

$C_j (j=1,2)$ - moving and stationary. The intrinsic mean μ and the principal geodesics $H_{v^{(n)}}$ are computed for each C_j using the samples. Once the principal geodesics are available, the classification of an unlabeled vehicle sample x can be performed by finding the category with the closest distance on the first principal geodesics $H_{v^{(1)}}$ to x. The corresponding motion status of a extracted vehicle j is found by

$$j^* = \arg\min \left\| \log(H_{j,v^{(1)}}^{-1} \mu^{-1} x) \right\|, \quad j \in \{1,2\} \tag{38}$$

The classification of vehicle motion state can successfully run based solely on the first principal geodesics. Although there are significant variations in shape over one vehicle category, the first principal geodesics $H_{v^{(1)}}$ is assumed to summarize essential shape features of vehicle points sampled by airborne LiDAR in view of distinguishing their motion states.

3.3.3 Velocity estimation

3.3.3.1 Estimation concept

The estimation of the velocity of detected moving vehicles can be done based on all effects moving objects may cause in single-pass ALS data by inverting the analytic functions in the motion artifacts model to explicitly relate the velocity with other observed and known parameters, which were originally presented in the Section 3.3.1. Thus, the approaches may use different measurements and derivations to estimate the velocity, which can be initially divided into two main groups, depending on whether the moving direction of vehicles is known or not:

i)   Given the intersection angle θ_v which can be further separated into following three methods using respective observations to estimate the velocity

- the measure for shearing angle of detected moving vehicles from their original orthogonal shape of rectangle
- the measure for stretching effect of detected moving vehicles from their original size
- the combination of the along-track and across-track velocity components which are estimated based on abovementioned effects, respectively

ii)   Not given the intersection angle θ_v

- the solution to a system of bivariate equations constructed by uniting the two formulas

The three methods in the first group assume that the moving direction of vehicles is given beforehand whereas the last one from the second group not. For the last case, the moving direction of vehicles can be estimated together with the velocity by uniting the variable of velocity with the variable of the intersection angle to build a system of bivariate equations and solving it, thereby giving the motion estimation from ALS data a great flexibility to deal with many arduous cases encountered in real-life scenarios. That means that not only the quantity of but also the direction of vehicles' velocity can be derived from the motion artifacts. All possible methods have their advantages and disadvantages and differ in the accuracy of their results which are to be analyzed and evaluated in detail in following subsections, respectively.

3.3.3.2 Velocity estimation based on the across-track deformation effect

The shearing angle of moving vehicles caused by the across-track deformation effect allows for a direct access to the object's velocity without the need of auxiliary information only except that the moving direction has to be assumed as known a-prior and input as an observation. Still, information about the relative orientation of the road axis corresponding to the particular motion of vehicle is needed in order to derive the real heading velocity of vehicles. If we assume that a detected moving vehicle originally acts as a rectangle object, the velocity estimate v of the detected moving vehicle based on the shearing effect of its shape is defined by

$$v = \frac{v_L \cdot \tan(\theta_{SA} - 90°)}{\cos\theta_v \cdot \tan(\theta_{SA} - 90°) + \sin(\theta_v)} \tag{39}$$

The value of intersection angle θ_v can be determined based on principal axis measurements of vehicle points as the flight direction of airborne LiDAR sensor can be always assumed to be known beforehand thanks to sustained navigation systems. Given Equation (39) which shows that the accuracy of the velocity estimate based on the across-track deformation effect σ_v^c is a function of the quality of the moving vehicle's heading angle relative to the sensor flight track θ_v and the accuracy of the shearing angle measurement θ_{SA}, the standard deviation of the derived velocity estimate v is calculated using the error propagation law and derived as

$$\sigma_v^c = \sqrt{\left(\frac{\partial v}{\partial \theta_v}\right)^2 \sigma_{\theta_v}^2 + \left(\frac{\partial v}{\partial \theta_{SA}}\right)^2 \sigma_{\theta_{SA}}^2} \tag{40}$$

$$= \sqrt{\begin{aligned} &\left(-\frac{v_L \cdot \tan(\theta_{SA} - 90°) \cdot (\cos(\theta_v) - \tan(\theta_{SA} - 90°) \cdot \sin(\theta_v))}{(\sin(\theta_v) + \tan(\theta_{SA} - 90°) \cdot \cos(\theta_v))^2}\right)^2 \sigma_{\theta_v}^2 \\ &+ \left(\frac{2v_L \cdot \sin(\theta_v) \cdot (\tan(90° - \theta_{SA})^2 + 1)}{\cos(2\theta_v) \cdot \tan(90° - \theta_{SA})^2 - 2\sin(2\theta_v) \cdot \tan(90° - \theta_{SA}) - \cos(2\theta_v) + \tan(90° - \theta_{SA})^2 + 1}\right)^2 \sigma_{\theta_{SA}}^2 \end{aligned}}$$

with v_L being the instantaneous flying velocity of the airborne senor system.

From empirical evaluations based on the analysis of the NavTeq road database the accuracy of the road heading direction is assessed to be $\sigma_{rhead} = 1°$ (Meyer et al., 2006). According to the value range of the error matrix derived in the procedure of vehicle shape parameterization the standard deviation of the vehicle's heading direction and the accuracy of the determination of vehicle shearing angle can be here assessed to be $\sigma_{vhead} = 3.5°$ and $\sigma_{\theta_{SA}} = 2°$, respectively. The accuracy of sensor flying direction is obtained via the navigation system output and assumed here to be $\sigma_{shead} = 2°$. The accuracy of the intersection angle θ_v between the moving directions of vehicle and LiDAR sensor can then be derived based on either the road heading direction or the vehicle's moving direction, since both of them could indicate the vehicle moving direction. Therefore, the accuracy of the intersection angle θ_v is assessed here to be $\sigma_{\theta_v} = 2°$.

Given the system parameters of airborne LiDAR and assuming the flying velocity of sensor $v_L = 120$km/h, the accuracy of velocity estimate σ_v^c based on the across-track deformation (shearing)

effect is derived by inserting the empirical error measures for the observations into Equation (40). Although the accuracies of the derived estimates and v are only directly related to two observed parameters describing the vehicle shape deformations - Ar_s and θ_{SA} as shown in their error functions (Equation (40) and (42)), respectively, for motion estimation applications, it is more intuitive and reasonable to interpret their accuracy relations with the dependence on the generic variables that can be directly perceived and controlled such as vehicle velocity and moving direction relative to airborne sensor. Moreover, two observation parameters cannot be arbitrarily specified to build up the independent variables for the 3D accuracy function as their relationships are subject to the physical law governed by the moving model. Fortunately, by doing simple transformations of different functional relations one can relate the error measure of velocity estimate σ_v directly to the institutive variables θ_v and σ_v/v by viewing σ_v or v as dependent variable and θ_v and v (or θ_v and σ_v/v) as independent variables. These configurations are valid for all the estimation methods.

The resulting error σ_v^c is shown in Figure 33 as a function of real target velocity v and vehicle heading angle relative to the sensor flying path θ_v and normalized with v. The relative velocity error σ_v^c/v is desirable to be used here to describe the accuracy, since the estimation accuracy only makes sense when the derived error is compared to the corresponding estimated velocity. It can be seen from Figure 33 that the heading velocity v of vehicles can be estimated with a high accuracy of $\sigma_v^c/v \leq 8\%$ if they were moving on urban roads (e.g. speed limit $v_T = 50$ km/h) with an intersection angle relative to the sensor heading direction of $\theta_v \geqslant 18°$. For vehicles running nearly in along-track direction ($\theta_v < 10°$) this approach fails to provide reliable velocity estimates. Additionally, one can get a standard deviation σ_v^c of approximately 5-10 km/h for the center part of the airborne LiDAR swath in urban areas no matter in which heading direction the vehicles travel.

(a)   (b)

Figure 33: Accuracy of vehicle velocities estimated from the across-track shearing effect, (a) standard deviation σ_v^c as a function of target velocity v and the intersection angle θ_v, σ_v^c is given in km/h. (b) Relative velocity error σ_v^c/v

As explained in Section 3.3.1 both across and along-track acceleration a_{xy} may give rise to irregular shearing of the rectangle shape of vehicle in across-track. Usually, it is assumed that the acceleration of vehicles in urban areas is zero during the time of sweeping-over. As a consequence, actually occurring minor accelerations might cause an error in velocity estimates depending on the

acceleration component. According to empirical studies based on Inertial Navigation System (INS) measurements with cars driving on city streets and highways, accelerations up to $a_{xy}=2\text{m/s}^2$ are likely to happen, which lead to contributing to max. 10% of σ_v as a "worst case" error source for the accuracy calculations according to considerable theoretic and simulation studies. Therefore, it is stated here that the error component aroused by vehicle accelerations can be omitted without changing the result of accuracy analysis noticeably. Moreover, it is generally shown that the velocity of slow moving targets cannot be reliably estimated even for very large intersection angle θ_v, whereas the speed of fast moving targets (approaching to the velocity of the airborne sensor platform) can be estimated with better relative accuracy.

3.3.3.3 Velocity estimation based on the along-track stretching effect

Besides of the above mentioned approach, the heading velocity of a moving vehicle can be derived by measuring its along-track stretching effect from its original vehicle size. The functional relation is given by

$$v = \frac{(1 - Ar/Ar_s) \cdot v_L}{\cos(\theta_v)} \tag{41}$$

where $Ar_s = l_s/w$ is the sensed aspect ratio of the detected moving vehicle, while Ar is the original aspect ratio of the vehicle and assumed to be input as constant. The accuracy of the velocity estimate based on the along – track stretching effect σ_v^a is a function of the quality of the aspect ratio measurement for the detected moving vehicle and the accuracy of the vehicle's heading direction relative to the sensor flight track. σ_v^a can also be calculated by error propagation law as follows:

$$\sigma_v^a = \sqrt{\left(\frac{\partial v}{\partial \theta_v}\right)^2 \sigma_{\theta_v}^2 + \left(\frac{\partial v}{\partial Ar_s}\right)^2 \sigma_{Ar_s}^2} \tag{42}$$

$$= \sqrt{\left(-\frac{v_L \cdot \sin(\theta_v) \cdot (Ar/Ar_s - 1)}{\cos(\theta_v)^2}\right)^2 \sigma_{\theta_v}^2 + \left(\frac{Ar \cdot v_L}{Ar_s^2 \cdot \cos(\theta_v)}\right)^2 \sigma_{Ar_s}^2}$$
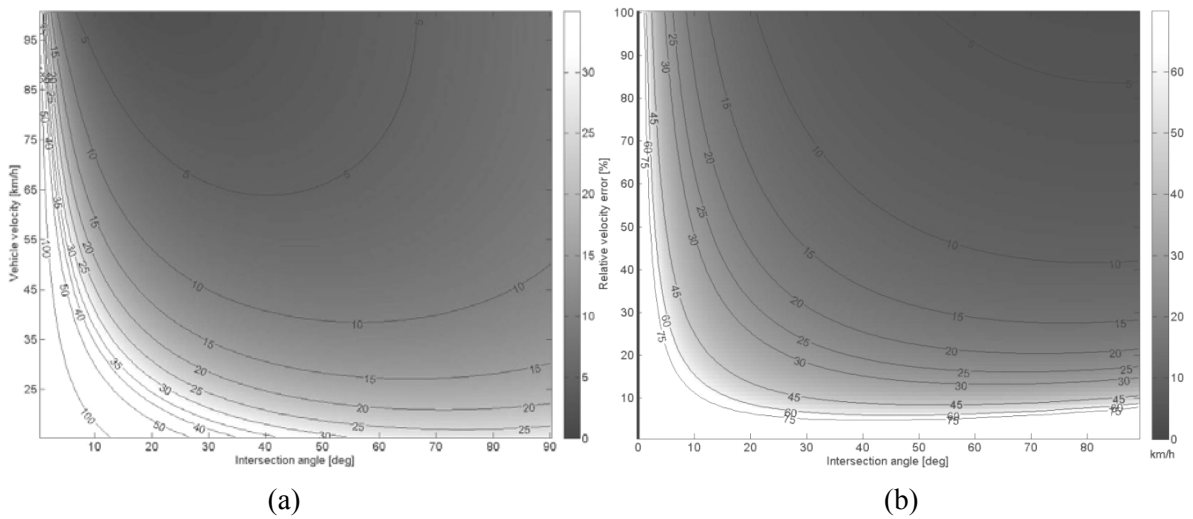

| (a) | (b) |
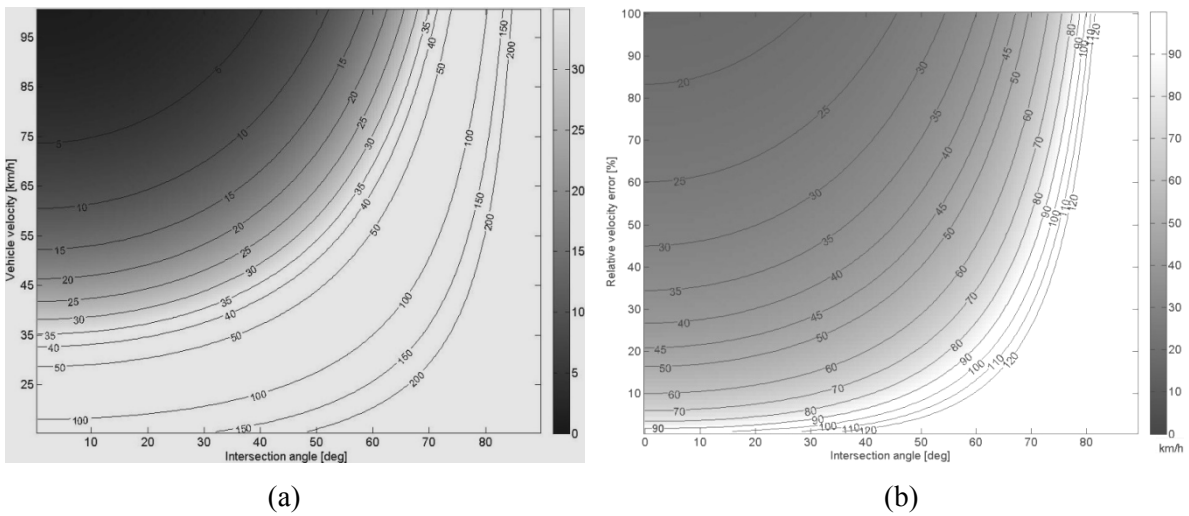
Figure 34: Accuracy of vehicle velocities estimated from the along-track stretching effect, (a) standard deviation σ_v^a as a function of target velocity v and the intersection angle θ_v, σ_v^a is given in km/h; (b) Relative velocity error σ_v^a/v

From empirical evaluations based on the analysis of the NavTeq road database the accuracy of the road heading direction is equivalently assessed to be σ$_{rhead}$ = 1°. From the vehicle shape parameterization procedure the standard deviation of the vehicle's heading direction and the accuracy of the determination of vehicle aspect ratio can be assessed to be σ$_{vhead}$ = 3.5° m and σ$_{Ars}$ = 2, respectively. The accuracy of sensor flying direction is obtained via the navigation system output and assumed here to be σ$_{shead}$ = 2° too. Therefore, the accuracy of the intersection angle θ_v is assessed here to be $\sigma_{\theta v}$ = 2°.

The accuracy of velocity estimates σ_v^a is derived by inserting the empirical error measures for the observations into Equation.(42). The resulting error σ_v^a is shown in Figure 34 as a function of real target velocity v and vehicle heading angle relative to the sensor flying path θ_v and normalized with v, respectively. It can be seen from that the heading velocity v of vehicles can be estimated with a high accuracy of $\sigma_v^a/v \leq 15\%$ too, if they were moving on roads with an intersection angle relative to the sensor heading direction of $\theta_v \leq 35°$. The estimation method based on the across-track shearing effect yields a wider and more robust range of the intersection angle for accurate velocity estimation than the method using along-track stretching shearing effect. The high accuracy of velocity estimation would happen at quasi along-track direction yet with a relative error value of ca. 10%, being a little bit worse than the highest accuracy of velocity estimation using the shearing effect. For vehicles running obviously in non-along-track direction ($\theta_v > 45°$) this approach fails to provide reliable velocity measures.

<div align="center">(a)   (b)</div>
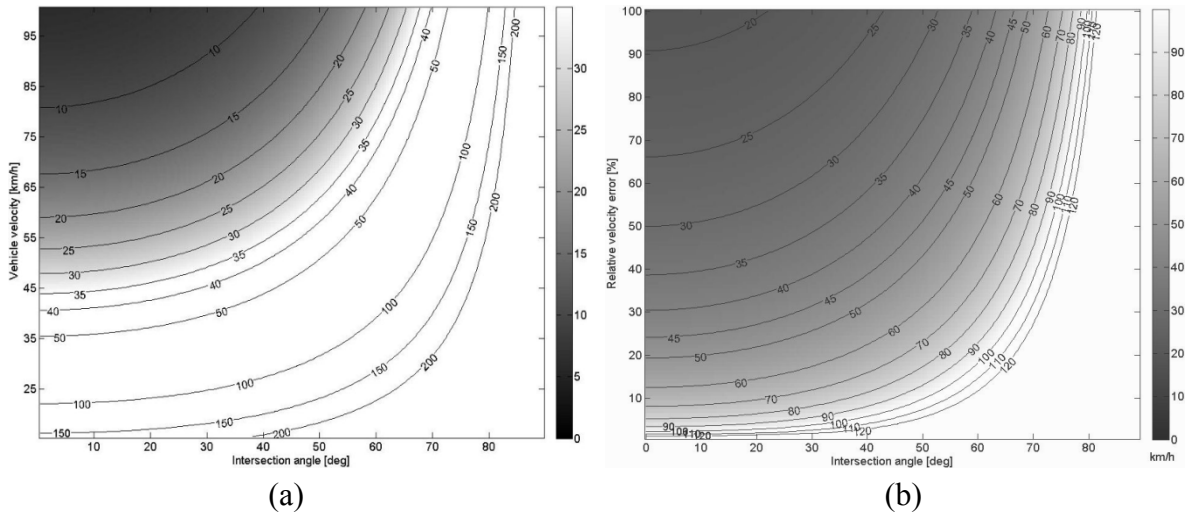
Figure 35: Accuracy of vehicle velocities estimated from the along-track stretching effect considering the error of the original aspect ratio *Ar*, (a) standard deviation $\sigma_v^{a'}$ as a function of target velocity v and the intersection angle θ_v, $\sigma_v^{a'}$ is given in km/h, (b) Relative velocity error $\sigma_v^{a'}/v$

As explained in Section 3.3.1 the proper estimation of the original aspect ratio *Ar* for each vehicle plays a key role in accurately determining the motion state and velocity of vehicles in ALS data. Usually, it is assumed that the aspect ratio of vehicles is two. As a consequence, actually inaccurately estimated aspect ratios cause an error in velocity estimates depending on the chosen original aspect ratio *Ar*. According to empirical studies based on random sampling measurements with real ALS datasets of various city areas, aspect ratios of vehicles *Ar* = 2 are mostly likely to happen for vehicles in city areas where passenger cars should represent the largest proportion of

traffic flow. Since vehicles of the aspect ratios up to Ar = 3-5 such as trucks and buses are rare to see in urban areas. Thus, σ_{Ar} = 0.5m is assumed as a "worst case" error source for the accuracy calculations of velocity estimates based on the vehicle aspect ratio. Considering this point, the standard deviation of this velocity estimate is also influenced by the accuracy of the original aspect ratio of vehicles which is specified independent on other measurements. Hence, for simplicity, the error formula can be newly formulated by inserting an additive error term σ_{Ar} characterizing the influence of aspect ratio Ar into Equation (42) to obtain the accuracy measure for velocity estimation by taking the uncertainty of original vehicle aspect ratio into account as

$$\sigma_v^{a'} = \sigma_v^a + k \cdot \sigma_{Ar} \tag{43}$$

where k is a multiplicative coefficient of the error term for the assumed aspect ratio, it is a sensitivity factor and controlled by the a-prior knowledge about vehicle size distribution in the test scene. The resulting error $\sigma_v^{a'}$ is shown in Figure 35, where it can be identified that the standard deviation $\sigma_v^{a'}$ of estimated velocities has increased by up to 10 km/h even for quasi along-track moving vehicles in urban areas, which correspond to an increase in the relative velocity error $\sigma_v^{a'}/v$ of 10% at least. Moreover, it is generally shown that the velocity of slow moving targets cannot be reliably estimated even for very small intersection angle θ_v, whereas the speed of fast moving targets (approaching to the velocity of the airborne sensor platform) can be estimated with better relative accuracy.

3.3.3.4 Velocity estimation based on combing two velocity components

Both of the above presented estimation methods might fail to give a reliable velocity estimate for vehicles, if they are moving in such direction that generated deformation effects for the vehicle shape are not dominated by either one of what two moving components account for(e.g. a moving vehicle with the intersection angle θ_v = 35° and the velocity v = 40km/h). To fill this gap and make the velocity estimate in the arbitrarily complex environment more robust and repeatable, it is proposed to use both of shape deformation effects for estimating velocities. The functional dependence of the velocity estimate can be given by sum of squares of two motion components which are derived based on two shape deformation parameters Ar_s and θ_{SA}, respectively:

$$v = \sqrt{\left(v_a\right)^2 + \left(v_c\right)^2} \tag{44}$$

$$\text{where} \begin{cases} v_a = v_L \cdot (1 - \dfrac{Ar}{Ar_s}) \\[3mm] v_c = \dfrac{v_L}{\cot(\theta_{SA} - 90°) + \cot(\theta_v)} \end{cases} \tag{45}$$

v_a and v_c are along-track and across-track motion components, respectively. The accuracy of the velocity estimate based on combing the two velocity components σ_v^{a+c} is a function of the quality of the along-track and across-track motion measurements for the detected moving vehicle and σ_v^{a+c} can be first calculated with respect to these two motion components by the error propagation law as:

$$\sigma_v^{a+c} = \sqrt{\left(\frac{\partial v}{\partial v_a}\right)^2 \sigma_{v_a}^2 + \left(\frac{\partial v}{\partial v_c}\right)^2 \sigma_{v_c}^2} \tag{46}$$

$$= \sqrt{\frac{v_a^2}{v_a^2 + v_c^2} \sigma_{v_a}^2 + \frac{v_c^2}{v_a^2 + v_c^2} \sigma_{v_c}^2}$$

where σ_{v_a} and σ_{v_c} are the standard deviations of along-track and across-track motion derivations. They can be further decomposed into accuracy with respect to three observation measures concerning the vehicle shape and motion parameters based on the function relations presented in Equation (45). Using the error propagation law, σ_{v_a} and σ_{v_c} are inferred as follows:

$$\sigma_{v_a} = \frac{\partial v_a}{\partial Ar_s} \sigma_{Ar_s} = \frac{Ar \cdot v_L}{Ar_s^2} \sigma_{Ar_s} \tag{47}$$

$$\sigma_{v_c} = \sqrt{\left(\frac{\partial v_c}{\partial \theta_v}\right)^2 \sigma_{\theta_v}^2 + \left(\frac{\partial v_c}{\partial \theta_{SA}}\right)^2 \sigma_{\theta_{SA}}^2}$$

$$= \sqrt{\left(\frac{v_L \cdot (\cot(\theta_v)^2 + 1)}{(\cot(90^\circ - \theta_{SA}) - \cot(\theta_v))^2}\right)^2 \sigma_{\theta_v}^2 + \left(\frac{v_L \cdot (\cot(90^\circ - \theta_{SA})^2 + 1)}{(\cot(90^\circ - \theta_{SA}) - \cot(\theta_v))^2}\right)^2 \sigma_{\theta_{SA}}^2} \tag{48}$$

Finally, after substituting Equations (47) and (48) into Equation (46), then the error propagation relation for the velocity estimate based on combing the two velocity components with respect to the three observed/derived variables Ar_s, θ_{SA} and θ_v is derived.

(a)   (b)
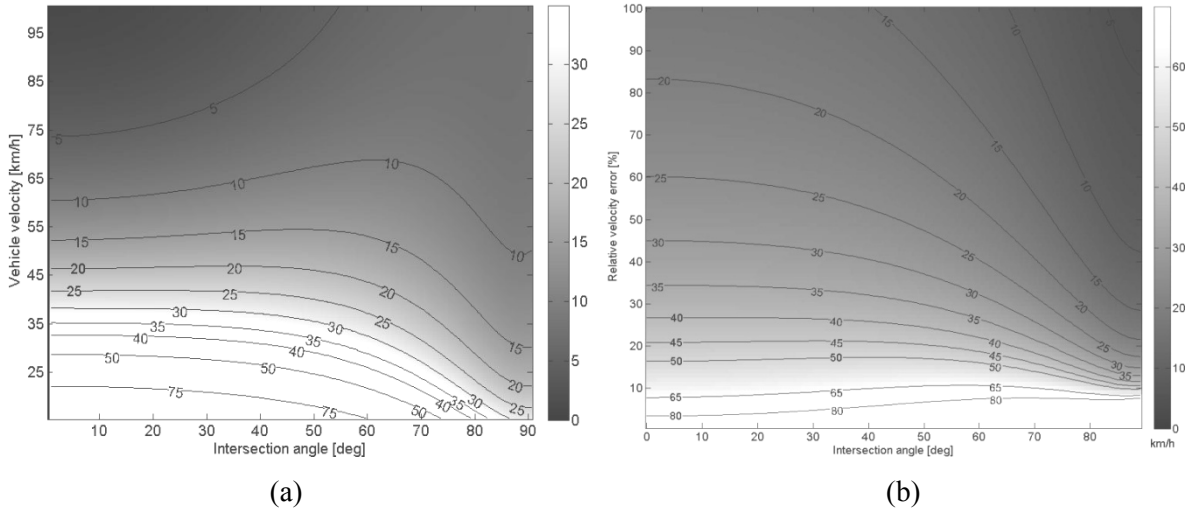
Figure 36: Accuracy of vehicle velocities estimated based on combing the two velocity components, (a) standard deviation σ_v^{a+c} as a function of target velocity v and the intersection angle θ_v, σ_v^{a+c} is given in km/h; (b) Relative velocity error σ_v^{a+c}/v

Based on the analysis of the NavTeq road database, the result of vehicle shape parameterization procedure and the navigation system output, the empirical error values for various observation measures including σ_{rhead}, σ_{Ars} $\sigma_{\theta v}$ and $\sigma_{\theta SA}$ can be assessed equivalently as in the preceding two methods presented above. The accuracy of velocity estimates σ_v^{a+c} based on combing the two velocity components is derived by inserting the empirical error measures for the observations into Equation (46). The resulting error σ_v^{a+c} is shown in Figure 36 as a function of real target velocity v and relative angle between vehicle heading direction and the sensor flying path θ_v and normalized with v - namely as relative velocity error. It can be seen from the error relation plot that the

heading velocity v of moving vehicles can be estimated with a high accuracy of $\sigma_v^{a+c}/v \leq 8\%$ whatever heading angle relative to the vehicle moving direction roads have on which they are moving. For vehicles running nearly in across-track direction with a relatively high velocity (e.g. $v > 60$ km/h) this approach tends to provide less reliable velocity measures, whereas those running with a relatively low velocity (e.g. $v < 40$ km/h) could only obtain reliable velocity measures by adjusting the moving direction to be in accordance with the sensor heading.

3.3.3.5 Joint estimation of moving velocity and direction

All of the already presented estimation methods are not able to give velocity estimates for vehicles if they are moving in an unknown direction or their moving detections cannot be accurately determined in advance. To solve this problem we propose to jointly consider estimated velocities and the intersection angle θ_v as the unknown parameters simultaneously while the variables describing the deformation effects caused by the motion components as observed. Actually, two analytic formulas describing the relations between motion artifacts and other parameters can be directly viewed as an equation system to which the velocity and the intersection angle are formulated as a set of solution. The system of bivariate equations for the velocity estimate relating unknown parameters to observations is given by:

$$\begin{cases} \theta_{SA} - 90° = \arctan\left(\dfrac{v \cdot \sin(\theta_v)}{v_L - v \cdot \cos(\theta_v)} \right) \\ 1 - \dfrac{v}{v_L} \cdot \cos(\theta_v) = \dfrac{Ar}{Ar_s} \end{cases} \tag{49}$$

The system of bivariate equations is to be solved using the substitution method. Firstly, transform second sub-equation of Equation (49) into

$$v = \frac{v_L}{\cos(\theta_v)} \cdot (1 - \frac{Ar}{Ar_s}) \tag{50}$$

and substitute it into first sub-equation of Equation (49), which has been converted into a more solution-friendly expression in advance:

$$\tan(\theta_{SA} - 90°) \cdot v_L = v \cdot (\tan(\theta_{SA} - 90°) \cdot \cos(\theta_v) + \sin(\theta_v)) \tag{51}$$

After substitution, the expression of Equation (51) can be rewritten as

$$\tan(\theta_{SA} - 90°) \cdot v_L = v_L \cdot (1 - \frac{Ar}{Ar_s}) \cdot \tan(\theta_{SA} - 90°) + \tan(\theta_v) \cdot v_L \cdot (1 - \frac{Ar}{Ar_s}) \tag{52}$$

Further do transformations to facilitate the solution and get

$$\tan(\theta_v) = \frac{\tan(\theta_{SA} - 90°) \cdot \left[1 - (1 - \frac{Ar}{Ar_s}) \right]}{1 - \frac{Ar}{Ar_s}} = \tan(\theta_{SA} - 90°) \cdot (\frac{Ar_s}{Ar_s - Ar} - 1) \tag{53}$$

$$\Rightarrow \theta_v = \arctan\left[\tan(\theta_{SA} - 90°) \cdot (\frac{Ar_s}{Ar_s - Ar} - 1) \right]$$

Finally, substitute the second sub-equation in Equation (53) into Equation (50) again and the velocity estimate of moving vehicle $\hat{v}$ can be derived as follows

$$v = v_L \cdot (1 - \frac{Ar}{Ar_s}) \cdot \sec \left\{ \arctan \left[\tan(\theta_{SA} - 90°) \cdot (\frac{Ar_s}{Ar_s - Ar} - 1) \right] \right\} \qquad (54)$$

It can be seen that the velocity of a moving vehicle can be directly estimated based on the shape deformation parameters without the need to assume the intersection angle θ_v as known a-prior. The intersection angle θ_v can be estimated as an intermediate variable solely based on two shape deformation parameters Ar_s, and θ_{SA} and is independent on the sensor flight velocity v_L. For accuracy analysis, two accuracy measures can be estimated for the intersection angle, namely the moving direction and the velocity values. The accuracies of the intersection angle $\sigma_{\theta v}$ and the velocity estimate σ_v can be derived as functions of the quality of the along-track stretching and across-track shearing measurements for the detected moving vehicle shape. Equivalently, $\sigma_{\theta v}$ and σ_v can be calculated with respect to these two deformation parameters by the error propagation law as:

$$\sigma_{\theta_v} = \sqrt{\left(\frac{\delta \theta_v}{\delta Ar_s} \right)^2 \sigma_{Ar_s}^2 + \left(\frac{\delta \theta_v}{\delta \theta_{SA}} \right)^2 \sigma_{\theta_{SA}}^2} \qquad (55)$$

$$= \sqrt{ \left(\frac{Ar \cdot \tan(90° - \theta_{SA})}{Ar^2 \cdot \tan(90° - \theta_{SA})^2 + (Ar - Ar_s)^2} \right)^2 \sigma_{Ar_s}^2 + \frac{Ar \cdot (\tan(90° - \theta_{SA})^2 + 1) \cdot (Ar - Ar_s)}{Ar^2 \cdot \tan(90° - \theta_{SA})^2 + (Ar - Ar_s)^2} \sigma_{\theta_{SA}}^2 }$$

$$\sigma_v = \sqrt{\left(\frac{\delta v}{\delta Ar_s} \right)^2 \sigma_{Ar_s}^2 + \left(\frac{\delta v}{\delta \theta_{SA}} \right)^2 \sigma_{\theta_{SA}}^2} \qquad (56)$$

$$= \sqrt{ \left(\frac{Ar \cdot v_L \cdot (Ar \cdot \tan(90° - \theta_{SA})^2 + Ar - Ar_s)}{Ar_s^2 (Ar - Ar_s) \cdot \sqrt{\frac{Ar^2 \tan(90° - \theta_{SA})^2 + (Ar - Ar_s)^2}{(Ar - Ar_s)^2}}} \right)^2 \sigma_{Ar_s}^2 + \left(\frac{Ar \cdot v_L \cdot \tan(90° - \theta_{SA}) \cdot (\tan(90° - \theta_{SA})^2 + 1)}{Ar_s (Ar - Ar_s) \cdot \sqrt{\frac{Ar^2 \tan(90° - \theta_{SA})^2 + (Ar - Ar_s)^2}{(Ar - Ar_s)^2}}} \right)^2 \sigma_{\theta_{SA}}^2 }$$

The empirical error values for two observation measures $\sigma_{Ar s}$ and $\sigma_{\theta_{SA}}$ were also assessed according to results of the vehicle shape parameterization procedure and can be assigned to the same values as used in the preceding estimation methods. The accuracy of derived intersection angle $\sigma_{\theta v}$ and velocity estimates σ_v based on the joint estimation of moving velocity and direction is derived by inserting the empirical error measures for the observations into Equation (55) and (56). The error of intersection angle $\sigma_{\theta v}$ is shown in Figure 37(a) as a function of real target velocity v and relative

angle between vehicle heading direction and the sensor flying path θ_v, the relative error is indicated in Figure 37(b). The resulting (relative) velocity error σ_v and σ_v/v are shown in Figure 38 as a function of real target velocity v and intersection angle θ_v. It can be seen from the plots that most of vehicles on road sections of urban areas could not allow for a high accuracy of moving direction estimation ($\sigma_{\theta v}/\theta_v > 25\%$) unless they move very fast (> 80 km/h). The high accuracy of velocity estimates could be only guaranteed for vehicles which obviously don't travel in across-track direction ($\theta_v < 75\%$). The overall accuracy of velocity estimates derived in this way is slightly degraded compared to other solutions where the moving direction is given.

(a)   (b)
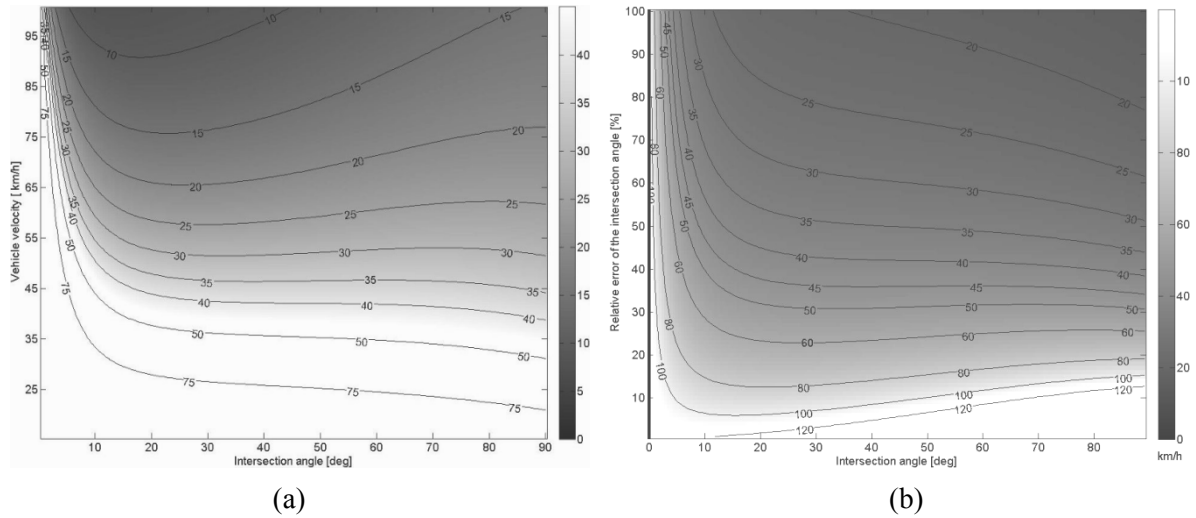
Figure 37: Accuracy of the intersection angle obtained based on the joint estimation of velocity and heading, (a) standard deviation $\sigma_{\theta v}$ as a function of target velocity v and the intersection angle θ_v, $\sigma_{\theta v}$ is given in degree; (b) Relative error of the intersection angle $\sigma_{\theta v}/\theta_v$

(a)   (b)
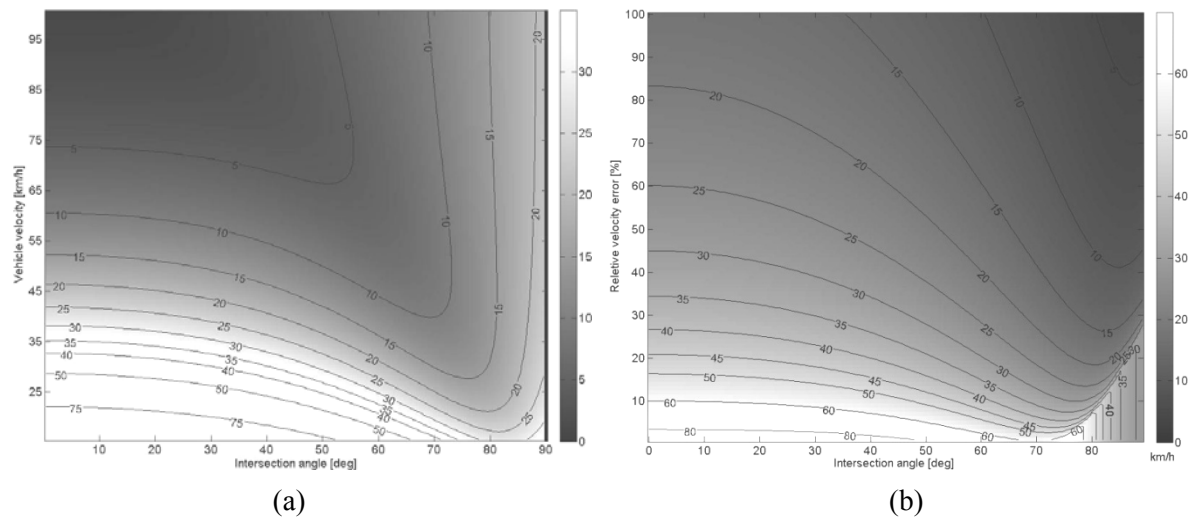
Figure 38: Accuracy of vehicle velocities obtained based on the joint estimation of velocity and heading, (a) standard deviation σ_v as a function of target velocity v and the intersection angle θ_v, σ_v is given in km/h; (b) Relative velocity error σ_v/v

Sections 3.3.3.2 – 3.3.3.5 show that there are several possibilities to estimate the velocity of moving vehicles from single-pass airborne LiDAR data. According to the quality analysis of the velocity estimates the combined usage of the along- and across track deformations is the most promising and reliable approach for a large variation range of heading angles. If vehicles move neither in

along-track nor in across-track direction (e.g. $35°<\theta_v < 50°$), the standard deviation of velocity estimates is almost fair for all estimators. However, the usage of joint estimation of motion direction and velocity establishes the unique solution to the generic case where the moving direction of target is unknown and can give fair results. Clearly, although the accuracy analysis is performed under the theoretic framework where the error sources are assumed to be independent to each other and the values of some errors are given empirically, for an analysis of traffic behavior and dynamics this accuracy analysis level is more than marginally sufficient and could give the valuable guidance on the selection of the proper method for vehicle estimation by using airborne LiDAR to monitoring specific traffic scenarios.

4 Experimental data

This chapter introduces the experimental data and evaluation scheme used to assess the results. All the used datasets of real-life scenarios were acquired under common conditions which are not optimized for detection of moving objects. The evaluation is conducted based on an established scheme for the external assessment of results of both vehicle extraction and motion detection.

4.1 Airborne LiDAR data

Five experimental datasets used in this thesis stem from three flight campaigns with airborne LiDAR systems which were conducted over city areas in North America and Europe. The main properties and parameters of these three ALS observations (flight velocity, point density, flight height, swath, and view mode) are summarized in Table 2. The experiments based on these datasets were conducted to validate the theory in real conditions. Please note that the conducted experiments are intended to yield comprehensive performance characterization of the proposed strategy to monitoring traffic when applied to real-life scenarios scanned by ALS under common conditions. That means that all the datasets are supposed to be acquired initially for the 3D city mapping tasks instead of the detection of moving objects. Since the nominal point density of the data is known, a vehicle is supposed to consist of ca. 20-64 laser points. Furthermore, the claimed vertical accuracy for the first flight campaign is 18 cm with 95 percent confidence and the horizontal accuracy is 1/3000 of the flying height. One example of test datasets – dataset Toronto I is displayed in Figure 39.

	Sensor	Pulse repetition rate [kHz]	Flying height [m]	Flying velocity [km/h]	Point density [points/m^2]	Swath [m]	View mode
Toronto I II III	Optech ALTM-200	90	520	120	4	650	nadir
TUM	Riegl LMS-Q560	110	480	140	5	480	forward
Enschede	FLI-MAP-400	75	275	100	10	420	nadir

Table 2: Acquisition configurations of airborne LiDAR campaigns

The first three experimental datasets -Toronto I, II and III were acquired covering the study area of Canadian metropolis - Toronto, located near the downtown area. The site is characterized by relatively high traffic density, with mutually intersected main roads and even city express road such as flyovers. ALS data recorded both first and last returns for each laser pulse. The scanning pattern was parallel line. To meet the requirements of traffic monitoring applications addressed in this thesis test datasets used are chosen from such sites that were flown only once.

Dataset Toronto I: The dataset Toronto I was acquired from first flight campaign and covers a flat urban area of ca. 400×600 m^2 with few large building blocks and a mixture of high and low vegetation. The terrain in this dataset is characterized by less varying topography with height

difference of about 10 m between the lowest and highest terrain points. The challenging elements here are the two express roads as flyover across over the area, some under passing roads, and several large parking lots around building and main roads.

Dataset Toronto II: The dataset Toronto II was acquired from first flight campaign over a flat urban area of ca. 400 $\times$ 600 m^2 with several large buildings and open yard with a mixture of high and low vegetations. The terrain in this dataset is characterized by less varying topography with height difference of about 20 m between the lowest and highest terrain points. The challenging elements here are the express road as a large flyover across over the test site and some under passing roads.

Dataset Toronto III: The dataset Toronto III was also acquired from first flight campaign in Toronto over a flat urban area of ca. 320 $\times$ 400 m^2 with several connected building blocks and open yard (parking lots) with a mixture of high and low vegetations. The terrain in this dataset is characterized by less varying topography with height difference of about 10 m between the lowest and highest terrain points. There are no elevated roads in the dataset. The challenging elements here are the complicated building structures and surrounding clutters.

The fourth and fifth datasets – TUM and Enschede were acquired over two European cities – Munich and Enschede, also located around the downtown area. The sites are characterized by relatively low traffic density, with fewer mutually intersected main roads of small width upon which few vehicles were observed to be traveling. ALS data recorded both first and last returns for each laser pulse, for second flight even full-waveform LiDAR data. The scanning pattern of both flights was parallel line.

Dataset TUM: The TUM dataset was acquired from the second flight campaign over a dense urban area of ca. 400 $\times$ 480 m^2 with densely connected buildings blocks with low vegetations on the road sides and in courtyards of building blocks. The terrain in this dataset is characterized by flat topography with height difference of about 5 m between the lowest and highest terrain points. The challenging elements here are densely constructed building blocks, a large amount of vegetations along road sides, and the oblique-view geometry of airborne LiDAR over the test site. The point density of this dataset is moderate compared to others.
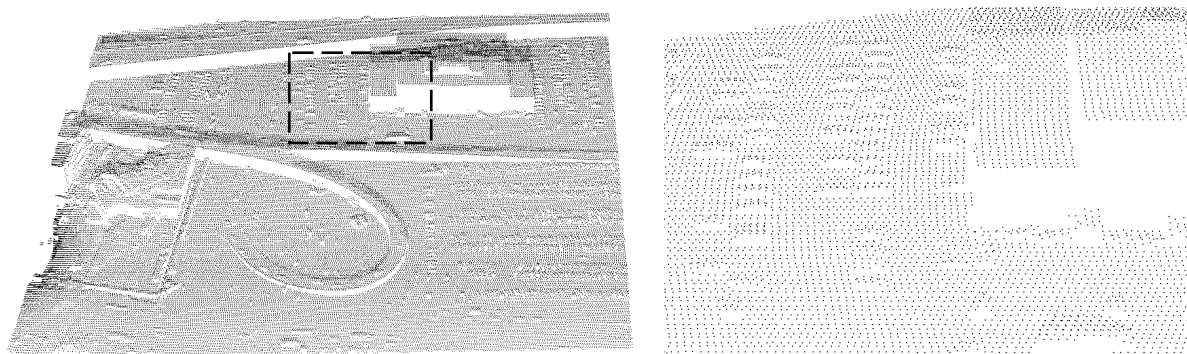
Figure 39: One example of test datasets: dataset Toronto I(left); Right: zoom-in of the data area marked by the dotted box

Dataset Enschede: The Enschede dataset was acquired from the third flight campaign over Enschede with an area of ca. 200 $\times$ 280 m^2. It consists of connected buildings blocks with low vegetations on the road sides and in courtyards of building blocks. The terrain in this dataset is characterized by

flat topography with height difference of about 8 m between the lowest and highest terrain points. The challenging elements here are densely constructed buildings, a large amount of vegetations along road sides and in the middle part. The point density of this dataset is the highest among the five test datasets and the data quality is assumed to be best as well.

4.2 External evaluation

4.2.1 Reference data

The external evaluation is used for the quantitative description of the quality of vehicle extraction and motion analysis. The quality is determined by comparing extraction/classification results with a reference dataset that is a record corresponding to the desired result. Even if the statement of quality is in the strict sense only valid for the subject of dataset used in the evaluation, it can nevertheless be estimated with them how an extraction/classification algorithm suitable for other datasets is, in which the objects to be extracted are marked with similar characteristics as in the evaluated dataset. Another advantage of an external evaluation is the possibility of systematic analysis of strengths and weaknesses of an approach (Hinz, 2003). Reference data are often created manually and supposed to be characterized by a high quality than the automatic extraction, so that the comparison of extraction with the reference can be deemed to be an objective assessment. However, they are usually not completely error-free. On the one side, errors can be concerned with geometry of the reference object with incomplete parts, e.g. the vehicle underneath trees, or its correct connection to the surroundings. On the other side, the semantics of objects with similar geometric properties in ALS data cannot be clarified in the recording of the references, e.g. ambiguous objects above ground like flowerbeds and low vegetations, traffic signs. The regions handled by these semantic uncertainties are excluded from the references for the evaluation.

If the references were created with the help of other (higher quality) data sources or instruments, they can be captured based on the information, which are not included in the data base of extraction and hence they are also available for the assessment of the extraction procedures. To avoid potential confusion it is distinguished between real and recordable reference (Klausmann et., 1999), where the recordable reference is created by a skilled human operator who is trained before. This implies that the human capability represents an upper bound on the achievable quality of the automatic extraction/classification. All the reference data used in this work are obtained based on the human inspection on either the original ALS datasets or ancillary data acquired simultaneously such as video data, namely corresponding to the recordable reference. However, even human observation is quite difficult in the unstructured ALS point cloud, making the reference data particularly for vehicle extraction more challenging and less accurate. This kind of reference data correspond to the so-called *minimal reference* (Hinz, 2003). The minimal reference includes objects whose semantics or dynamical state could be definitely determined, and therefore should be extracted or determined in any case. The focus is thus put on the completeness of extraction.
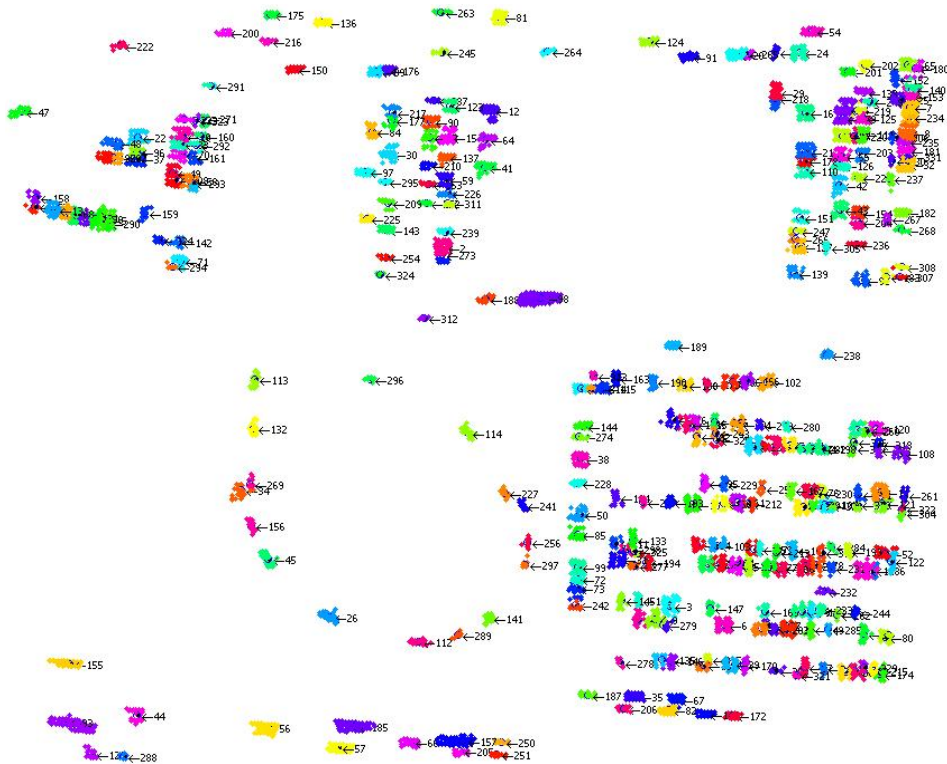
Figure 40: One example for the reference data for vehicle extraction – dataset Toronto I with 317 vehicles, every vehicle object is indicated by a color

In summary, for vehicle extraction there is a minimum reference available which is created by a human operator via visual analysis to mask out single vehicles in point cloud of ALS data. One example of the reference data for vehicle extraction is shown in Figure 40. During last two campaigns for data acquisition the video camera data were also made. For both of flights, an optical/infrared video camera mounted on the same airborne platform was used to provide concurrent observations. They could exhibit the reference data for dynamical traffic situation at the time of acquisition. For the first flight, there is unfortunately no concurrent additional sensor mounted on the platform, thereby making the direct comparison impossible. However, the vehicle movement can be empirically inferred by spatial contexts involved in data to enable the assessment of motion classification results. For both the distinction of vehicle motion state and velocity estimation, analogous to the evaluation of vehicle extraction, the evaluation is based on the reference data which are also minimum reference. The velocities of extracted moving vehicles can be accurately estimated by vehicle tracking in concurrently acquired video data und viewed as ground truth for velocity estimation. One example of such reference data for motion analysis is shown in Figure 41. The following section describes a well-established scheme for the external evaluation of vehicle extraction and motion analysis in short.
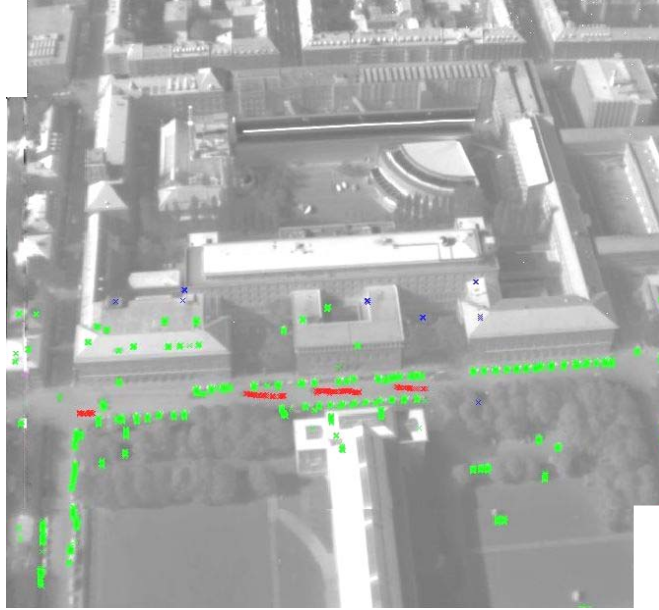
Figure 41: Video reference data for motion analysis displayed on a composite of two video frames, green: stationary vehicles, red: trajectories of moving vehicles

4.2.2 Evaluation schema

The method used for the external assessment evaluates extracted vehicles by the criteria of completeness, reliability and geometric accuracy. These can be calculated by the local allocation of object-parts of the extraction to the reference. On the other hand, by the measures of commission and omission errors, the properties of the distinction of vehicle motion state within the context of the traffic monitoring are evaluated. The determination of evaluation results for the whole strategy is divided into two steps: (1) The evaluation of vehicle extraction with reference, and (2) the calculation of quality measurements for motion analysis. Apart from algorithmic differences, the assignment of the two vehicle objects is a conceptually identical approach, as to be described in the section below. The calculation of the assessment is done based on the object-parts appropriately assigned together or with discrete points affiliated to the objects. Note, however, that there exists a relation between assignment and evaluation criteria. For instance, a tolerant assignment leads to less geometric precision of the extraction result. This threshold for data assignment needs to be determined according to application purposes.

➢ **Vehicle extraction**

Detection quality measure - The object-based evaluation scheme is used to assess the vehicle extraction algorithm in the sense of detection quality, which measures the ability of the algorithm to resolve semantic information. For this reason, the Absolute/Relative Accuracy for Object Extraction - *AAOE* and *EAOE* are implemented for measuring the completeness and correctness at object level, respectively, which are derived by

$$AAOE = \frac{N_{1,1}}{N_r}, \quad EAOE = \frac{N_{1,1}}{N_s} \tag{57}$$

where $N_{1,1}$ is the number of vehicles which has one-to-one relationship with the reference data, N_r is the true number of vehicles in the field, N_s is the total number of automatically extracted vehicles. One-to-one relationship means that the area ratio of one point segment to a reference object is within the range of [80% 120%], where the object area is established by corresponding the point with the pixel by the geo-tilling processing.

Geometric preservation accuracy – The shape delineation accuracy, which measures the shape accuracy of extracted points of vehicle objects compared to their true shape.

To particularly guarantee that the extracted vehicle points can be delivered to the motion analysis and obtain the rational velocity based on the moving vehicle model presented in Section 3.3.1 (Yao et al., 2008), the assessment of extracted vehicle points with respect to the shape accordance is necessary. The Hausdorff distance $H(E,R)$ is devoted to expressing the similarity measure. The two sets of vehicle points are delivered to the comparison by calculating the $H(E,R)$ by

$$H(E,R) = \max(h(E,R), h(R,E)) \tag{58}$$

where

$$h(E,R) = \max_{a \in E} \min_{b \in R} \|a - b\|,$$

and E, R are the extracted and reference vehicles, respectively; A schematic description for deriving the delineation accuracy by shape comparison is denotes in Figure 42.
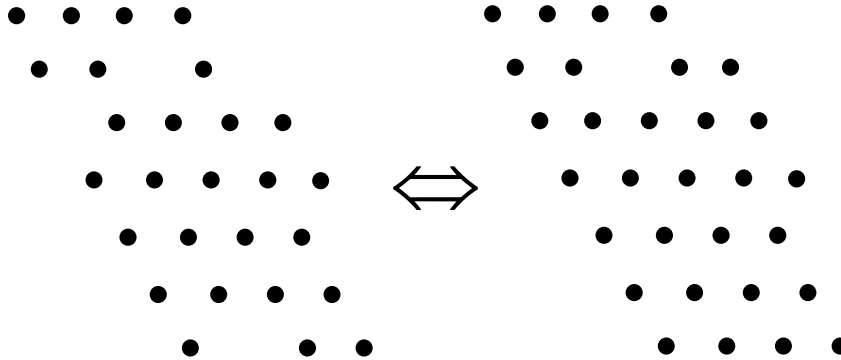
Figure 42: Shape comparison for a moving vehicle. Left: extracted points, right: corresponding reference vehicle points, i.e. $H(E,R) = 0.38m$.

➢ **Motion state distinction**

The moving objects represent the most important proportion of traffic flow information and play a key role in the traffic monitoring for urban areas. Therefore, it is decided to evaluate the motion analysis results based on examining the classified moving objects.

Since the motion state analysis is a binary classification process and an uncertain motion category is excluded in advance, the evaluation of results is primarily focused on the classification of data. Classification can be divided into two main categories: correct classification of moving and stationary vehicles, and incorrect classification. The incorrect classification can be further divided into two groups: classification of moving vehicle points as stationary vehicle points (termed Error type I), which is defined as omission errors

$$Error_I = \frac{N_{M,S}}{N_{M,M} + N_{M,S}} \cdot 100 \ [\%] \tag{59}$$

where $N_{M,S}$ is the number of vehicles which should be moving but are classified as stationary class. $N_{M,M}$ is the number of vehicles which should be moving and are also classified as moving class.

and the classification of stationary vehicle points as moving vehicle points (termed Error type II), which is defined as commission error:

$$Error_{II} = \frac{N_{S,M}}{N_{M,M} + N_{S,M}} \cdot 100 \ [\%] \tag{60}$$

where $N_{S,M}$ is the number of vehicles which should be stationary but are classified as moving class. $N_{M,M}$ is the number of vehicles which should be moving and are also classified as moving class.

Based on this terminology, type II errors have a greater effect than type I errors, since missing a few moving vehicle points is not equivalent to inserting non-moving elements into the moving vehicle representation. Most classification algorithms focus on reducing the type II errors, sometimes at the cost of a larger amount of type I errors. The results will show that the applied algorithm intends to minimize types II error, while keeping type I error negligible or at a low level.

5 Experimental results

This Chapter is to apply the proposed strategy to the datasets of Chapter 4 to examine the effectiveness of these methods for traffic analysis from airborne LiDAR platform. Five different datasets are selected to undergo the experiment. The content is organized according to the sequence of applied datasets. The parameters and thresholds for every intermediate step are adaptively determined by empirically examining training data chosen from different data areas but of the same flight campaign. The training and evaluation processes were based on objects where the interpretation was made unambiguously with the help of both geometric and context information. For the distinction of vehicle motion state, moving and stationary vehicles are collected to serve as positive and negative training samples, respectively. The velocity estimation method by combining two motion components presented in Section 3.3.3.4 is selected for the test datasets due to its reliability and generality, the original aspect ratio of vehicles Ar is assumed $= 2$.

5.1 Dataset Toronto I

Since the dataset Toronto I is characterized by the large parking lots and elevated road, it will undergo the third processing *Route* defined in the whole strategy (Figure 6) by combing the OBPA method with local context-guided method to extract vehicles. Parameter settings are chosen as follows: $T_1 = 1.5$, $T_2 = 25$, $NCut_{thres} = 0.35$. In Figure 43 the results of vehicle motion analysis are visualized. The result of vehicle extraction is not presented along, because it is believed to be able to be perceived and indicated within the motion analysis results too. The results are also reflected in the quantitative external evaluation (Table 3). Both travelling vehicles on roads and parking ones are successfully extracted prior to be delivered to motion state analysis, particularly including many vehicles that were placed in the large parking lots. The extraction achieved a vehicle extraction completeness of about 76% and a reliability of near 85%. A RMS$\{H(E,R)\}$ of 0.41m implies a mean

AAOE	EAOE	RMS$\{H(E,R)\}$
78.5%	85.1%	0.41m

(a)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	232	7	45	284		
	Moving	8	28	2	38	Type I	22%
					322	Type II	20%

(b)

Table 3: Evaluation for dataset Toronto I: (a) vehicle extraction, (b) motion state classification

geometric inaccuracy of the value corresponding to the data resolution. Nearly 80% of moving vehicles from extracted vehicles have been detected correctly while 20% of detected moving vehicles are false alarms and should be motionless. The velocities of moving vehicles derived based on the motion artifacts effect are shown as color-coded in Figure 43(b).
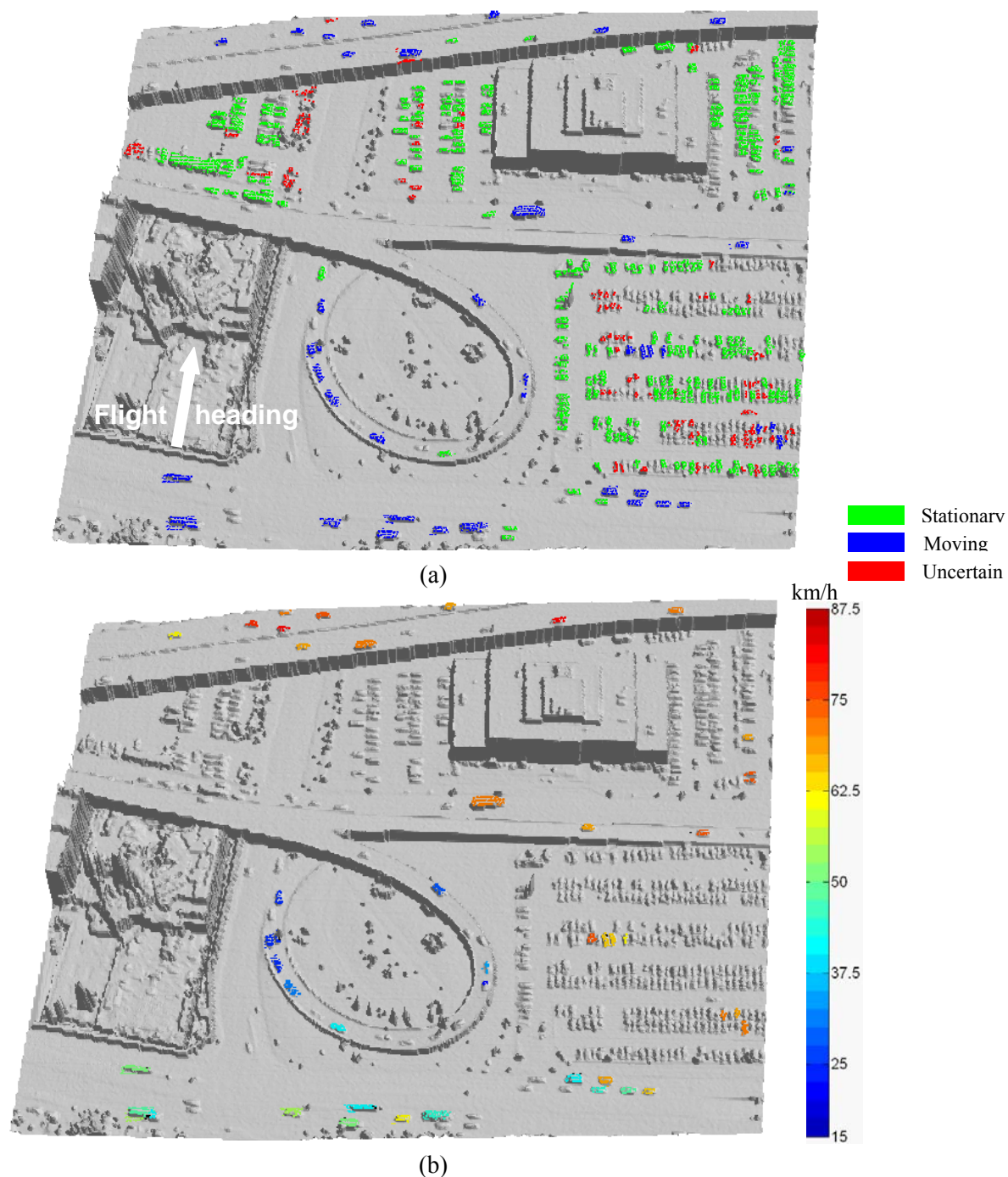
Figure 43: Vehicle motion analysis results for dataset Toronto I: (a) motion state classification, (b) velocity estimation

5.2 Dataset Toronto II

Since the Toronto dataset II is characterized by the elevated road, it will undergo the third processing

Route defined in the whole strategy by using the OBPA method only to extract vehicles. Parameter settings are chosen as follows: $T_1 = 1.5$, $T_2 = 25$. In Figure 44 the results of vehicle motion analysis are visualized. Both travelling vehicles on roads and parking ones are successfully extracted, including many vehicles that were placed along the road margins. The visual results are also reflected in the quantitative external evaluation (Table 4). The extraction achieved a vehicle extraction completeness of about 80% and a reliability of near 84%. The extraction shows its geometric precision with a RMS$\{H(E,R)\}$ of 0.36m implying a mean geometric inaccuracy of the value corresponding to less than the data resolution. Nearly 80% of moving vehicles from extracted vehicles have been detected correctly while 15% of detected moving vehicles are false alarms and should be motionless. The velocities of moving vehicles derived based on the motion artifacts effect are shown as color-coded in Figure 44(b).

AAOE	*EAOE*	RMS$\{H(E,R)\}$
80.3%	84.5%	0.36m

(a)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	57	5	12	74		
	Moving	6	23	4	33	Type I	21%
					107	Type II	18%

(b)

Table 4: Evaluation for dataset Toronto II: (a) vehicle extraction, (b) motion state classification
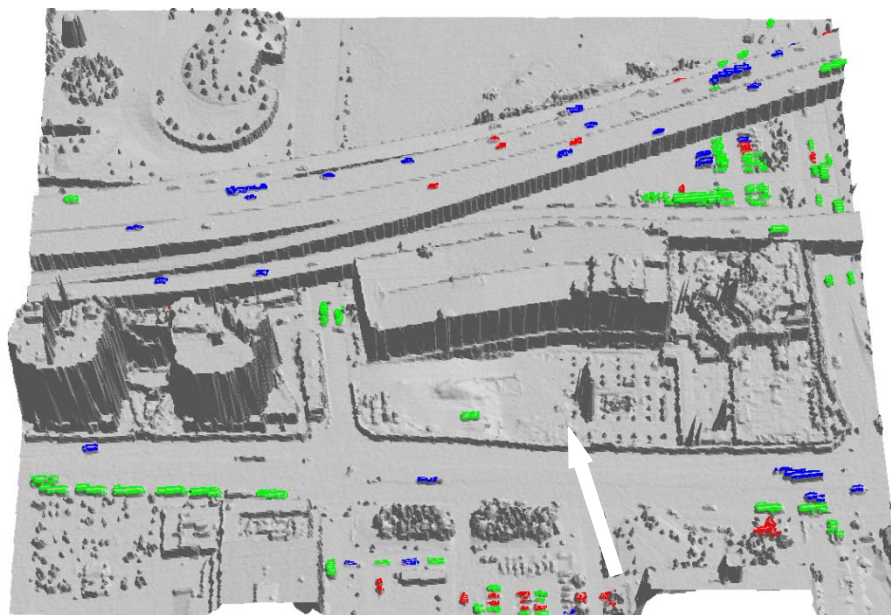
(a)
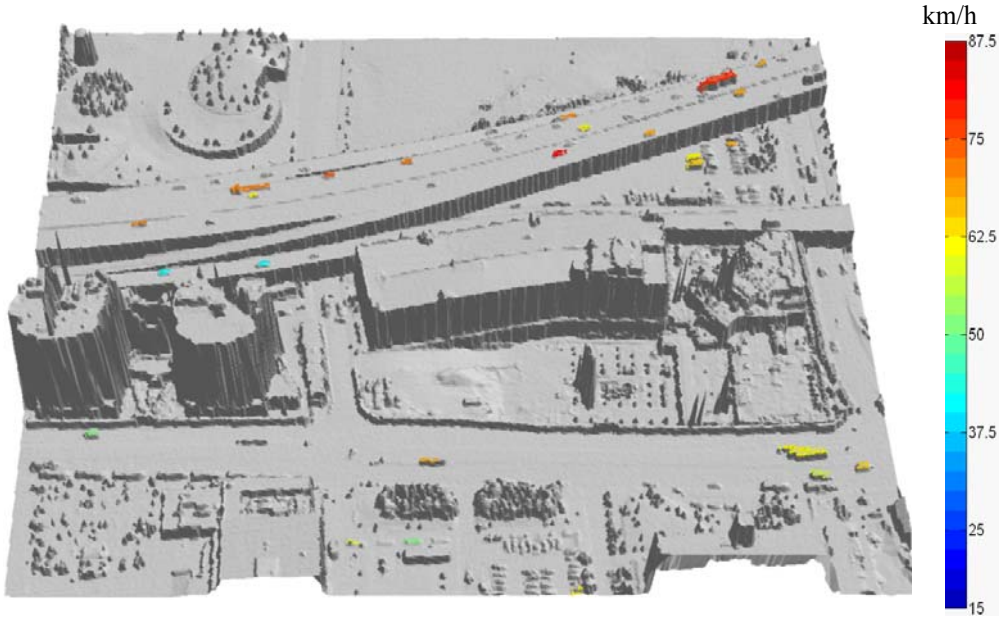
(b)

Figure 44: Vehicle motion analysis results for dataset Toronto II: (a) motion state classification, (b) velocity estimation

5.3 Dataset TUM

Since the TUM dataset is characterized by no elevated road, it will undergo the first processing *Route* defined in the whole strategy by using the local context guided method only to extract vehicles. On the other side, this dataset is used to assess the boundary conditions of the proposed strategy to extracting urban traffic information when applied to common airborne LiDAR data, as the oblique-view airborne LiDAR data has proven to be unsuitable for vehicle extraction task due to the unstable laser reflection properties against the mental surface of vehicle body under this view angle (Yao et al., 2008). Parameter settings are chosen as follows: $h = 0.5$.

AAOE	*EAOE*	*RMS{H(E,R)}*
68.3%	78.3%	0.43m

(a)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	57	2	8	67		
	Moving	3	4	1	8	Type I	38%
					75	Type II	34%

(b)

Table 5: Evaluation for TUM Dataset: (a) vehicle extraction, (b) vehicle motion classification

In Figure 45 the results of vehicle motion analysis are visualized. The quantitative external

evaluation is summarized in Table 5. The result achieved a vehicle extraction completeness of about 69% and a reliability of near to 80%. The extraction shows its geometric precision with a RMS{$H(E,R)$} of 0.43m implying a mean geometric inaccuracy of the value corresponding to less than the data resolution. Nearly 62% of moving vehicles from extracted vehicles have been detected correctly while 34% of detected moving vehicles are false alarms and should be motionless. The velocities of moving vehicles derived based on the motion artifacts effect are listed in Table 6 in comparison to the ground truth data obtained by the concurrently acquired video data.

Figure 45: Vehicle motion analysis result for dataset TUM (displayed as overlaid on the DSM exclusive of trees)

Target #	v_r km/h	v_e km/h	$\left|\Delta v\right|$ km/h
1	28.6	20.7	5.1
2	24.9	21.3	3.6
3	15.7	23.1	7.4
4	12.5	20.1	7.6
5	10.3	18.5	8.2

Table 6 Comparison of estimated velocities v_e with reference v_r for TUM dataset

5.4 Dataset Enschede

Since the Enschede dataset is also characterized by no elevated road, it will undergo the first processing Route defined in the whole strategy (Figure 6) by using the local context guided method only to extract vehicles. Parameter settings are chosen as follows: $h = 0.5$.

To enhance the visual impression on vehicle extraction in this dataset, the results of vehicle extraction and motion analysis are illustrated in Figure 46, respectively. The quantitative external

evaluation is summarized in Table 7. The result achieved a vehicle extraction completeness of about 87% and a reliability of nearly 83%. The extraction shows its geometric precision with a RMS$\{H(E,R)\}$ of 0.23m implying a mean geometric inaccuracy of the value corresponding to less than the point spacing. Nearly 80% of moving vehicles from extracted vehicles have been detected correctly while 30% of detected moving vehicles are false alarms and should be motionless. The velocities of moving vehicles derived based on the motion artifacts effect are listed in Table 8 which are compared to the reference data obtained by the concurrently acquired video data.

(a)

(b)

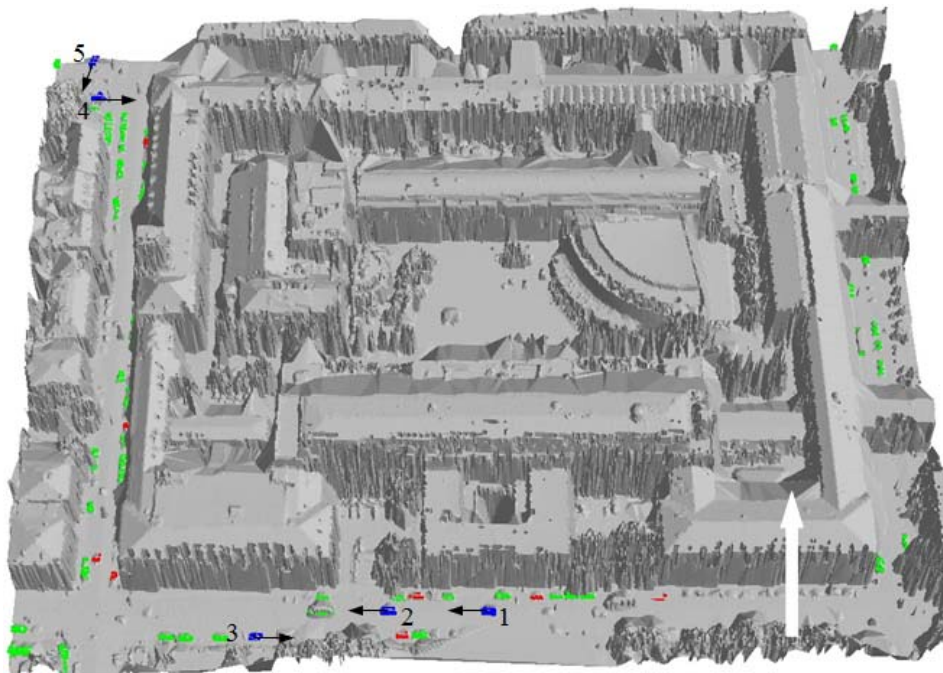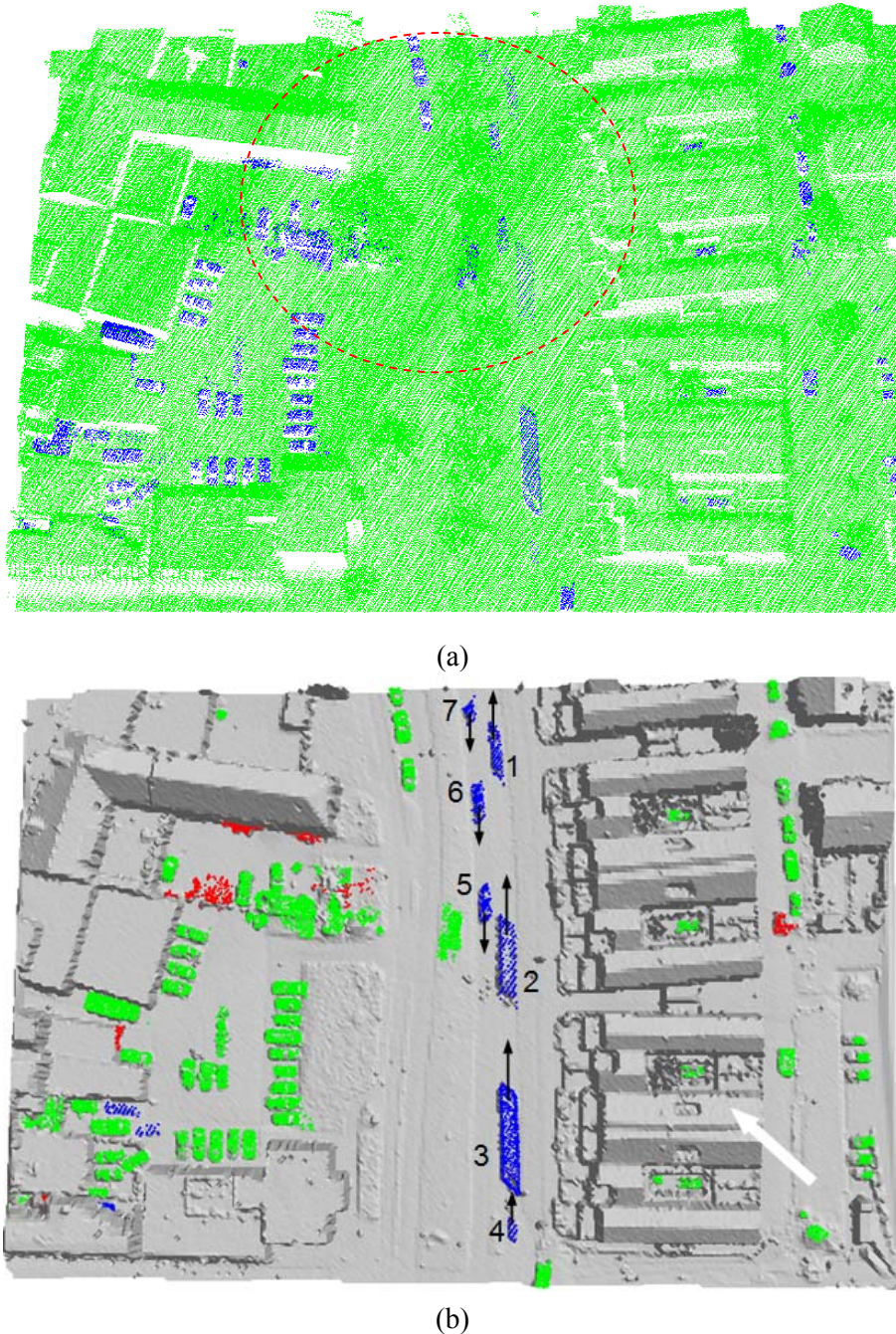Figure 46: Vehicle analysis results for dataset Enschede: (a) vehicle extraction, (b) vehicle motion analysis (displayed as overlaid on the DSM without trees)

AAOE	EAOE	RMS{H(E,R)}
87.2%	86.5%	0.23m

(a)

		Classified				Errors	
		Stationary	Moving	Uncertain			
Reference	Stationary	68	2	7	77		
	Moving	2	7	2	11	Type I	22%
					88	Type II	27%

(b)

Table 7: Evaluation for Enschede dataset: (a) vehicle extraction, (b) distinction of motion state.

Target #	v_r km/h	v_e km/h	$\left|\triangle v\right|$ km/h
1	61.2	64.3	3.1
2	62.2	58.6	3.6
3	54.6	57.3	2.7
4	46.5	51.6	5.1
5	58.7	54.5	4.6
6	57.7	55.6	2.1
7	62.9	59.7	3.2

Table 8: Comparison of estimated velocities with reference for Enschede dataset

5.5 Comparison of vehicle extraction methods towards motion analysis

Initially, different procedures were adopted to deal with traffic scenes of two global contexts via two methods of vehicle extraction, as shown in Chapter 3. However, from technical viewpoints, it could be argued that the OBPA method (vehicle extraction method II)   can also be independently applied to the test datasets of the global context regions without elevated roads, which are normally processed by the local context guided method (vehicle extraction method I) to extract vehicles. In this section, it is desirable to evaluate and compare the performances of these two vehicles extraction methods when applied to the same datasets, especially with respect to the extraction of vehicle dynamical information. Given same ALS datasets (where the quality of vehicle extraction is comparable between two methods, so it has to exclude elevated roads in the dataset) we would like to see the respective strongnesses and shortcomings of two vehicle extraction methods in the context of urban traffic monitoring by using airborne LiDAR.

To realize the comparative study between two methods raised above, two datasets with unbiased global context relations are selected here to undergo the experiment. It should make sure that the datasets to be used for comparison study are equivalent with respect to data characteristics and

context relations, thereby ensuring comparability of the results. The first test dataset used to be comparative study is dataset Toronto III while the second test dataset is dataset Enschede used in the experiments presented in Section 5.4, as it also shows a balanced data conditions for two vehicle extraction methods. Both datasets are to undergo the first and second processing *Routes* defined in the whole strategy by using either the local context guided or the OBPA methods to extract vehicles.

(a)

(b)
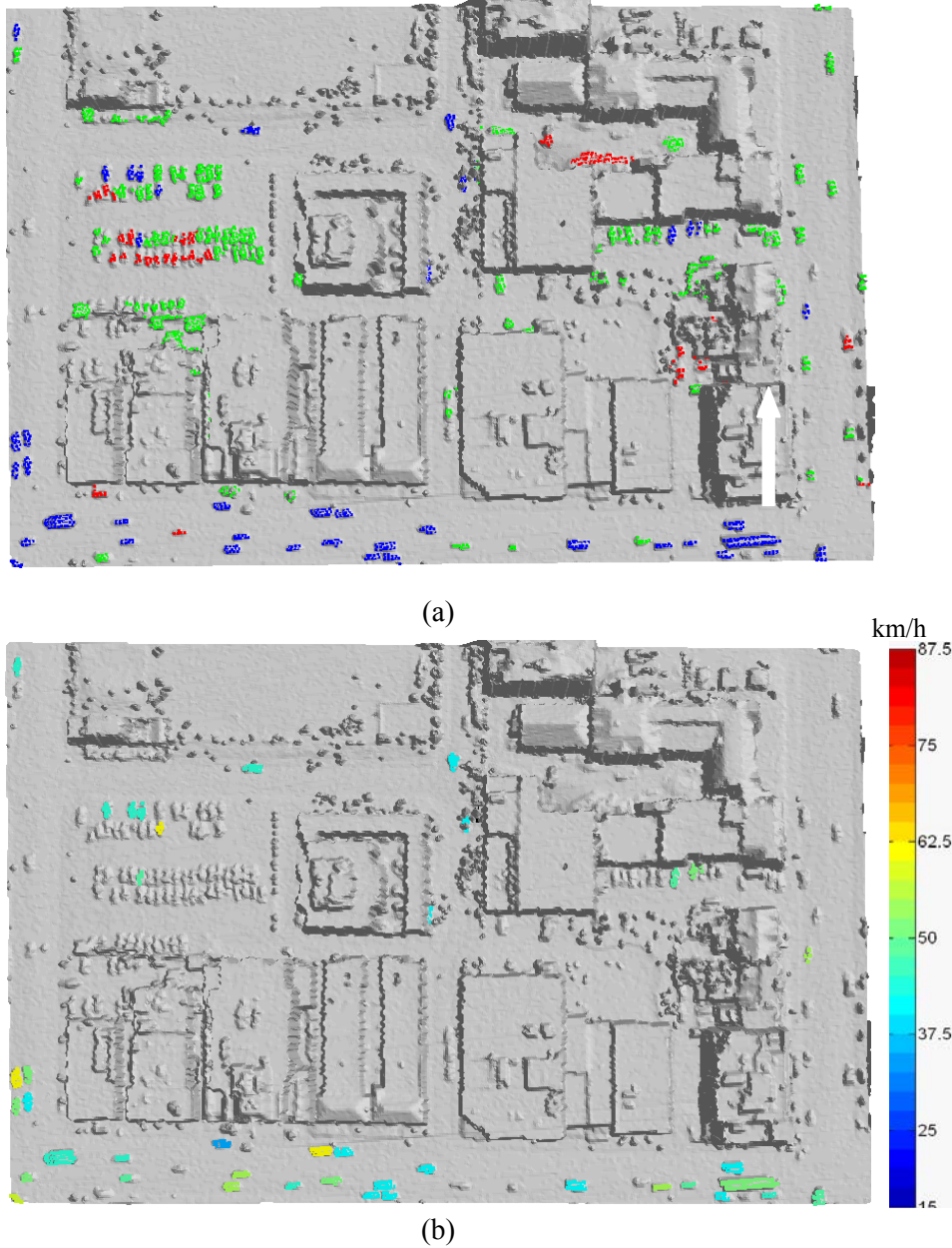
Figure 47: Vehicle motion analysis results for first dataset based on vehicle extraction method I; (a), motion detection (b) velocity estimation of moving vehicles

In Figure 47, Figure 48 and Figure 49 the results of vehicle motion analysis for both datasets using two different vehicle extraction methods are visualized (the result for dataset Enschede using vehicle extraction method I can be found in Figure 46(b)). The quantitative external evaluation of the results can be found in the Table 9 and Table 10. It is to be found out that the strategy based on method I to extract vehicles achieved a better vehicle detection rate, e.g. for dataset Toronto III, with the completeness of ca. 79%, being 6 percentage points higher than that achieved by the OBPA method, while the reliability is almost equivalent for both methods. However, the method II shows a high

geometric precision of extracted vehicles with a RMS{$H(E,R)$} of 0.34m implying a increase in shape preservation accuracy by 15%. For both datasets there is a better agreement between the motion classification results and the reference for most of vehicles when using method II (e.g. approx. 83% vs. 87% for detection rate of moving class in dataset Toronto III). The detection reliability has also increased to some extent (e.g. approx. 74% vs. 82% for detection rate of moving class in dataset Toronto III). Both travelling vehicles on roads and parking ones are successfully extracted (e.g. approx. up to 80% of moving vehicles in dataset Enschede have been detected), including those that were placed along the road margins or in parking areas.

The velocities of moving vehicles of two datasets derived based on two vehicle extraction methods are denoted or listed in Figure 47(b), Figure 48(b) and Table 11, respectively. For dataset Toronto III, moving vehicles derived based on method I has the mean velocity of ca.45 km/h and the velocity range of 15 - 79 km/h; moving vehicles derived based on method II has the mean velocity of ca.52 km/h and the velocity range of 25-66 km/h. For dataset Enschede, the mean error of estimated velocity for moving vehicles has reduced from ca. 3.9km/h using method I to ca. 3.5 km/h using method II.

	AAOE	*EAOE*	*RMS{H(E,R)}*
Method I	78.8%	78.1%	0.39m
Method II	73.2%	79.8%	0.34m

(a)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	48	8	9	65		
	Moving	5	23	4	32	Type I	17%
					97	Type II	26%

(b)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	34	6	5	45		
	Moving	4	27	2	32	Type I	13%
					77	Type II	18%

(c)

Table 9: Evaluation for dataset Toronto III: (a) vehicle extraction, (b) motion classification using vehicle extraction method I, (c) motion classification using vehicle extraction method II.

(a)

(b)
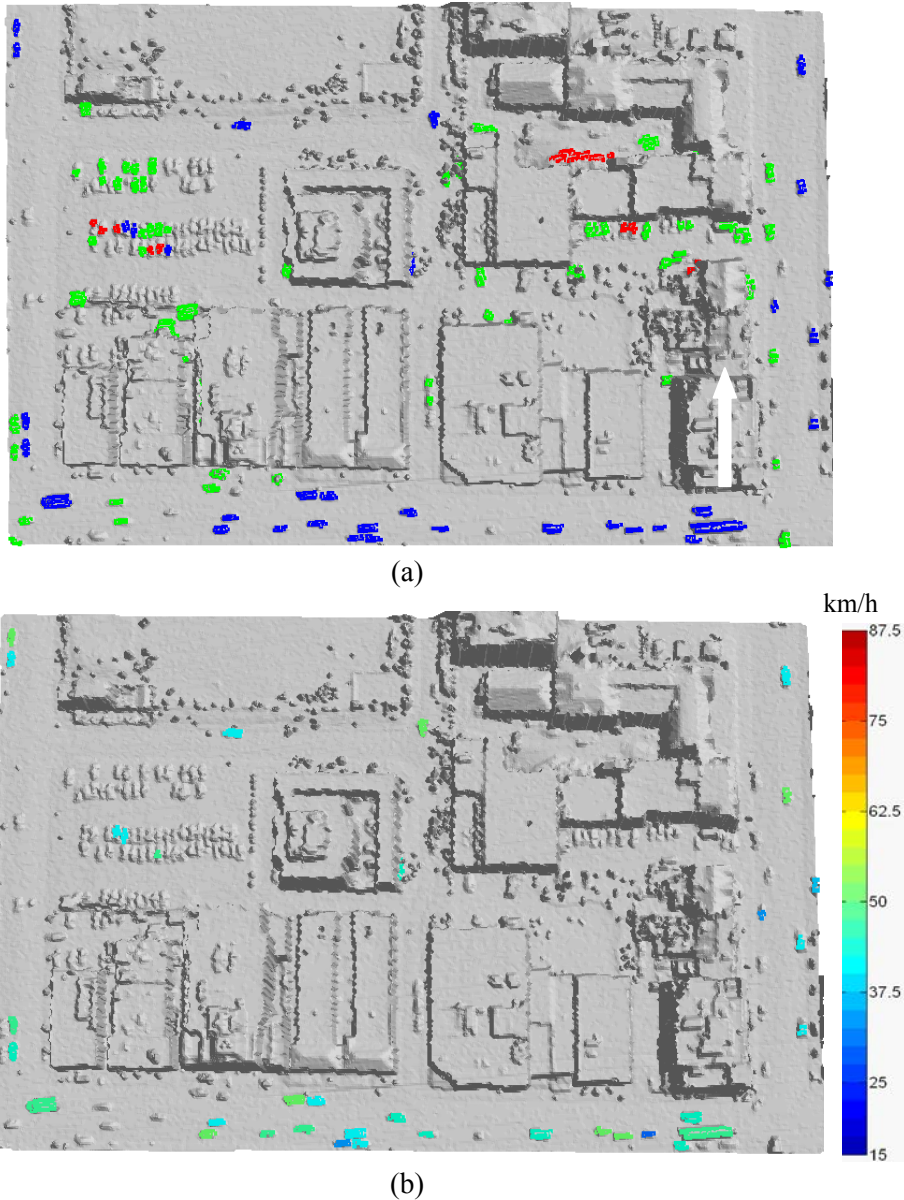
Figure 48: Vehicle motion analysis results for dataset Toronto III based on vehicle extraction method II: (a) vehicle motion classification, (b) velocity estimation of moving vehicles

	AAOE	*EAOE*	*RMS{H(E,R)}*
Method II	82.1%	85.6%	0.19m

(a)

		Classified					
		Stationary	Moving	Uncertain		Errors	
Reference	Stationary	61	2	5	68		
	Moving	2	8	2	12	Type I	20%
					80	Type II	20%

(b)

Table 10 Evaluation for vehicle motion analysis from dataset Enschede based on vehicle extraction method II: (a) vehicle extraction, (b) vehicle motion classification
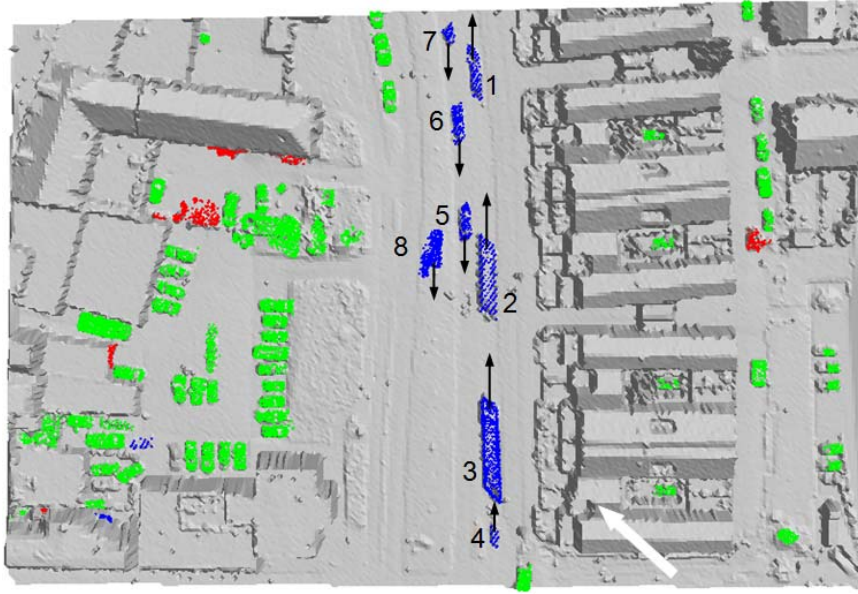
Figure 49: Vehicle motion analysis results for dataset Enschede using vehicle extraction method II

Target #	v_r km/h	v_e km/h	$\lvert \Delta v \rvert$ km/h
1	61.2	63.3	2.1
2	62.2	57.1	5.1
3	54.6	57.8	3.2
4	46.5	50.1	3.6
5	58.3	55.3	3.0
6	57.7	55.5	2.2
7	62.9	59.5	3.4
8	22.9	29.4	6.5

Table 11: Comparison of estimated velocities with reference for dataset Enschede based on vehicle extraction method II

6 Discussion and performance analysis

In this chapter the performance of the proposed strategy for traffic monitoring will be discussed and analyzed. The first part of this chapter (Section 6.1) is directly dedicated to the discussion of the experimental results obtained in Chapter 5, and the rest parts (Section 6.2 - 6.4) are further dedicated to the essential aspects of the analytic and empirical accuracy analysis for both motion classification (detection) and velocity estimation based on simulated and real data. It is helpful to understand inherent relations between respective acquisition parameters and predict the performances of future airborne LiDAR campaigns towards traffic monitoring. The preferences and weaknesses of each method are identified. The different properties of the datasets also allow an analysis of the methods with respect to their requirements for the data.

6.1 Discussion of experimental results

6.1.1 Dataset Toronto I

Through the visual analysis in Figure 43, the overall performance of the strategy seems to be encouraging. Both travelling vehicles on roads and parking ones are successfully extracted prior to be delivered to motion state analysis, particularly including many vehicles that were placed in the large parking lots. Those vehicles which are not extracted or extracted as false alarms are usually caused by ambiguous laser reflection against certain objects' surface or deficient topological relation to the surroundings. Most of these vehicles are stationary and would appear in the parking lots where they are located very close to each other. Some bulged objects above the ground such as flowerbeds and traffic signs are also incorrectly detected as vehicle due to the similar height relation to the ground surface.

The 3D segmentation step can increase the extraction completeness by introducing the constraint on height discontinuity, while the extraction correctness has been kept on a high level. The misdetection is greatly alleviated by adaptively determining the bandwidth for the MS operation controlled by the spatial edge. An abrupt change in the extraction correctness seems not to exist. Additionally, the refinement step by introducing the local context analysis into the classified results can further notably increase the completeness. Especially, small vehicles located densely to each other in the parking lots or on the roads can be detected (Figure 43(a)). Actually, this dataset is characterized by the most challenging scene among the test datasets, as its point density is relatively low and it features a high probability of occurrence of objects with unstable laser sampling. It is found out that the strategy tends to work well with these vehicles showing significant height boundary to surroundings. There could be a higher accuracy of geometry preservation for large and high vehicles such as bus, truck and those vehicles which are located sparsely to each other. For those small vehicles placed close to each other, the geometric accuracy could be degraded. This is the reason why the dataset Toronto I shows a relatively poor performance concerning vehicle shape preservation in comparison to others. Equivalently, the point density seems to have an impact on resolving the accurate outlines of vehicles. The proposed OBPA method can generate a

vehicle-oriented partition of urban areas from airborne LiDAR data by using adaptive MS clustering, though which vehicle instances can be extracted. In addition, the approach can achieve an improved classification performance on ALS data towards semantic object extraction, because MS is able to provide more robust and expressive features at object level considered.

As can be seen in Table 3(b) and Figure 43(a), there is an overall good agreement between the motion classification result and the reference measurement for most of vehicles (approx. 80% detection rate for moving class), so that the theoretical background of motion detection seems justified. It can also be found from the evaluation results that most of moving vehicles were detected in flyovers and main roads of the city and the vehicles classified as motionless are primarily found in the parking lots or along road margins. The red class indicates the vehicles of uncertain status subject to the shape ambiguity. They are mostly placed densely to each other in parking lots or contaminated by neighboring clutter objects. Misdetections and false alarms for the moving class usually appear either when vehicles travel very slowly in the along-track direction or they consist of ambiguous sampled points caused by unstable laser reflection. However, it is noteworthy that both of commission and omission errors of motion classification for the moving class have already been greatly alleviated by identifying and removing the uncertain class prior to the classification step. This kind of vehicles primarily corresponds to those extracted vehicles whose shape cannot be modeled as parallelogram. They usually appear amid parking lots and are false alarms for vehicle extraction (red points in Figure 43 (a)).

As shown in the Figure 43(b), for dataset Toronto I the mean speed of moving vehicles travelling on the flyover on the top is about 75 km/h, which corresponds to the velocity limit on the city express road very well. It is noted that some vehicles on the flyover on the bottom were moving at speeds being quite different from the mean value. Actually, it is found that most of the vehicles with slow velocity on the flyover were either turning at the sharp curve or just leaving the exit of. For other moving vehicles, the slow ones were either on the road shoulder or in the slow lane; the fast ones either were passing other vehicles or had just passed on other vehicle. It is interesting to note that some vehicles were in the slow lane, but moving at high speed.

6.1.2 Dataset Toronto II

Through the visual analysis in Figure 44, the overall performance of the strategy seems to be promising. The good visual impression of the results is also reflected in the quantitative external evaluation (Table 4). The extraction achieved a vehicle extraction completeness of about 80% and a reliability of near 84%. The extraction shows its geometric precision with a RMS$\{H(E,R)\}$ of 0.36m implying a mean geometric inaccuracy of the value corresponding to less than the data resolution. The ambiguous laser reflection against certain objects' surface and deficient topological relation to the surroundings still result in the misdetection and false alarms for vehicle extraction. Most of these vehicles are small-sized or would appear densely in the parking lots. Some bulged objects above the ground such as low trees and traffic signs are incorrectly detected as vehicle due to the similar height relation to the ground surface.

The 3D segmentation based classification method is used to extract vehicles only, which provides high extraction completeness and correctness rates except for some small-sized vehicles above the

open areas in the middle part of the test site whose height edges in relation to ground surface are vague. Actually, this dataset is characterized by the same challenging scene as the dataset Toronto I, apart from the similar properties concerning vehicle extraction, however, it can be found out that the strategy tends to work well with this dataset and provide an even better result. Moreover, there is overall a higher accuracy of geometry preservation for extracted vehicles in dataset Toronto II. The reason why this situation could come out may remain in the fact that much fewer vehicles in dataset Toronto II appear in the large parking lots which make the occurrence of vehicles with unstable laser sampling and deficient topological relations less probably. The proposed OBPA method can be directly performed on the dataset to generate a vehicle-oriented partition of urban areas, though which vehicle instances can be distinguished from other objects.

As can be seen in Table 4(b) and Figure 44(a), there is a good agreement between the motion classification results and the reference for most of vehicles (approx. 78% detection rate for moving class). It can also be found from the evaluation results that most of moving vehicles were detected on the flyover and the main road on the southern side of the major buildings and the vehicles classified as motionless are primarily found along the road margins or underneath the flyover. The red vehicles of uncertain motion state are found to be placed densely to each other in parking lots and contaminated by neighboring objects, while some small vehicles on the flyover are also classified as uncertain. Misdetection (e.g. vehicles just passing through the tunnel underneath the flyover surface with a cautious velocity, on the right border of the test site) and false alarms (e.g. vehicles located along road margins occluded by vegetation on the bottom part or contaminated by closely adjacent objects on the top-right part) for the detected moving class appear either when vehicles travel very slowly in the (near) along-track direction or they consist of ambiguous sampled points caused by unstable laser reflection or undersized vehicle. However, the commission error of motion classification for the moving class have been greatly alleviated by identifying and removing the uncertain class prior to the classification step, while the omission error has been slightly uplifted by identifying the moving vehicles on the flyover as uncertain class. However, since missing a few moving vehicle points is not equivalent to inserting non-moving elements into the moving vehicles and the commission error has obviously more impact on the results in view of traffic monitoring, it can be still justified to identify vehicles of uncertain motion class and remove them prior to performing motion classification.

According to Figure 44(b), the velocities of moving vehicles derived based on the motion artifacts effect are representative for the cars travelling either on the road or on the flyover. Especially, the values for the moving vehicles on the main carriageways are reasonable for this section of the urban express road and correspond very well to the speed limit allowed for moving vehicles. Although there might be some inaccuracies included in the measurements the results show a good match of theory and real measurements. There are only two anomalies among the moving vehicles on the flyover which were travelling with a distinctly slow velocity (<= 50km/h). By further observation it can be found that the two vehicles were leaving the main carriageways of the flyover and diverted to the side way. At the time when the airborne LiDAR sweeping over, they were slowing down for changing the lanes and thus show the different to others on the same flyover. Additionally, the moving vehicles on the normal urban road (across the bottom part) show a moderate velocity or a bit beyond the velocity maximum allowed in urban areas such as Toronto. It is assumed that those vehicles were heading for a traffic intersection and wanted to pass through it as quickly as possible. It can also be noted that the vehicle estimated with slow velocity is found to appear in the margin

area of the road.

6.1.3 Dataset TUM

Through the visual analysis in Figure 45, the overall performance of the strategy still seems to be encouraging. It is observed that many vehicles parked on the margins of the roads that are oriented toward the flight direction and few ones on the road lanes are successfully extracted. Even several vehicles beneath the trees are recovered, which are usually unable to be detected in optical data. However, further inspection of the result could indicate that still many vehicles on the along-track roads are not detected probably because of the lack of sufficient sampled points for each vehicle object. Most energy of the laser pulses which were emitted out and hit these vehicles were finally scattered away due to the oblique incidence angle. The vehicles on the road that is oriented perpendicular toward the flight direction and behind the building are completely missed. The ambiguous laser reflection against certain objects' surface and deficient topological relation to the surroundings still result in the misdetection and false alarms for vehicle extraction. Most of these vehicles are small-sized or would appear very close to adjacent clutters. The bulged objects above the ground such as traffic signs are incorrectly detected as vehicle due to the similar height relation to the ground surface as described in the vehicle model.

The visual impression of the results can be reflected in the quantitative external evaluation (Table 5) to some extent. The extraction achieved a vehicle extraction completeness of about 70% and a reliability of near 78%, which is the worst among the five test datasets. The extraction shows its geometric precision with a comparable RMS$\{H(E,R)\}$ of 0.43m implying a relative stability of geometric outlining capability of the proposed vehicle extraction methods against ALS data acquired under different conditions. The local context guided method is used to extract vehicles only, which provides moderate extraction completeness and correctness rates. However, it can be found out that the strategy tends to work well with many parking vehicles except for some small-sized vehicles beneath the vegetation on the road margins whose height edges in relation to ground surface are vague. This dataset is characterized by the most challenging scene due to the forward-looking mode; the local-maxima detection could be degraded by oversmoothed or incomplete height edges of vehicles and cannot provide sufficient and accurate seed for flooding simulation used for watershed transformation to outline the vehicle shape. Moreover, the accuracy of geometry preservation for extracted vehicles in dataset TUM is becoming worse compared to other datasets. The reason may be tracked back to the fact that the background-markers for watershed transformation which are primarily responsible for the accurate vehicle delineation could not be reliably determined, since some vehicle objects themselves have degenerated from independent and distinct semantic entities to connected and ambiguous clutters, leading to making the indication of the back-ground markers by extracting ridge lines impossible.

According to Table 5(b), it can be also observed that dataset TUM exhibits a relatively moderate performance in terms of vehicle motion classification, which might prove that the acquisition geometry of oblique view LiDAR could further impose negative impacts on the accuracy of motion indication and velocity estimation based on the vehicle shape analysis. It can be easily inferred that ALS data acquired under oblique-view condition could undermine the scanning and reflection mechanism of laser pulse upon vehicle by producing more vehicles with ambiguous or incomplete

sampled points, thereby making the motion classification based on shape features more difficult.

As can be further seen in Table 5(b) and Figure 45, the performance for motion detection becomes deteriorated (approx. 65% detection rate for moving class). Only five moving vehicles were detected on the lanes of main road on the southern side of the building block and the vehicles classified as motionless are primarily found to appear along the road margins of two along-track roads. The red vehicles of uncertain motion state are found to be placed densely to each other in parking lots and contaminated by neighboring objects. Misdetection and false alarms for the detected moving class appear either when vehicles travel very slowly in the (near) along-track roads or they consist of ambiguous sampled points. Equivalently, the commission error of motion classification for the moving class have been alleviated by identifying and removing the uncertain class despite of the strongly varied structure of vehicle points caused by oblique-view mode, while the omission error has been also uplifted to some extent due to the boundary condition of data acquisition used in this experiment.

From Table 6 it can be seen that the velocities derived from the motion artifacts effect are travelled obviously slower than normal ones but still representative for the cars travelling on that section of the urban road and correspond very well to the speed value of vehicles moving towards the traffic light (Figure 45), e.g. vehicle #1 and #2 were approaching the road crossing when the traffic light was red. Moving vehicle #1 exhibits a little bit higher velocity than vehicle #2, as vehicle #2 travelled behind a heavy truck that is not detected and had to brake early and strongly in order to avoid potential danger. Vehicle #5 had a low averaged velocity since it were coming out from a small road and had to turn right onto the main road. Vehicle #3 and #4 became nearly stagnant, as it is supposed that they had then just started the engine preparing to get to the road lane from the parking strips. However, the accuracies of their velocity estimates are even worsen by partial occlusion of trees and poor sampled points through the view mode and scene complexity. Although inaccuracies are definitely included in the measurements the results show a good match of theory and real measurements.

6.1.4 Dataset Enschede

With the first glance at Figure 46, the overall performance of the strategy on this dataset is very promising. The unique main road in the scene on which vehicles were traveling is not oriented perpendicular toward the sensor flight direction. Most of vehicles which were either travelling on roads or parked are successfully extracted, even including several ones partially occluded by trees (within red ellipse of Figure 46(a)), which are usually unable to be detected in optical data; only few ones are not which are undersampled due to ambiguous laser reflection or deficient topological relation to surroundings. Most of these vehicles are small-sized or would appear very close to adjacent clutters or vehicles themselves.

As shown in the quantitative evaluation in Table 7, there is already a considerable improvement in terms of both the completeness and correctness for vehicle extraction. The results achieved a vehicle extraction completeness of almost 87% and reliability from near 85%, which is the best among the five datasets. The extraction shows its geometric precision with a comparable RMS$\{H(E,R)\}$ of 0.23m implying a fairly increase of the vehicle extraction method in geometric outlining capability

thanks to the high point density and quality of the dataset. The local context guided method used here to extract vehicles has demonstrated the capability to provide both high completeness and correctness. It tends to show that the method work well with parking vehicles of some small-sized vehicles or those partially occluded the vegetation. This dataset is characterized by very high point density; the local-maxima detection could therefore provide sufficient and accurate seed for flooding simulation used for watershed transformation to outline the vehicle shape. Moreover, the accuracy of geometry preservation for extracted vehicles in Enschede dataset becomes to get increased to such a high level that also corresponds to the point spacing of the dataset.

According to Table 7(b), it can be also observed that Enschede dataset exhibits a good performance in terms of vehicle motion classification, which might prove that the point density of airborne LiDAR could further impose positive impacts on the accuracy of motion indication and velocity estimation. It can be easily inferred that ALS data acquired with high point density could strengthen the reflection mechanism of laser pulse upon vehicle by acquiring vehicle objects with more complete sampled points and distinct silhouette features, thereby making the motion classification based on shape features more easy. As can be further seen in Table 7(b) and Figure 46, almost all the moving vehicles (7 vehicles) on the lanes of main road in the scene have been successfully detected except only one which was moving through a tree occlusion, and the vehicles classified as motionless are primarily found to appear in the parking lots or along the margins of small roads. The red vehicles of uncertain motion state are found to be densely placed underneath trees and contaminated by neighboring clutters. Misdetection and false alarms for the detected moving class appear have been reduced, they would most probably happen when the vehicles consist of ambiguous sampled points due to unfavorable scanning angles. Finally, the commission error for the moving class has been reduced by identifying and removing the uncertain class in spite of the usage of such a high density dataset.

From Table 8 it can be seen that the velocities for the moving vehicles derived from the motion artifacts effect are representative on that section of the urban road and correspond very well to the speed limit in a European city (Figure 46(b)), e.g. vehicle #1, #2 #3 were passing through this road section uninterruptedly with proper and similar velocities. Moving vehicle #4 that was moving in the same lane exhibits a little bit lower velocity than vehicle #3, as vehicle #4 travelled behind a heavy truck/bus and had to brake early in order to keep an adequate distance. Since the vehicles were moving along neither along-track nor across-track direction, they all were imaged as parallelogram with distinct elongation and shearing, leading to a clear conclusion to the motion state. Vehicles #5 and #6 had a little bit lower averaged velocity. Vehicle #7 was moving faster than the two precedent moving vehicles, as it is supposed that they had intended to pursue two moving ones. The accuracies of their velocity estimates are much better than those derived for TUM dataset according to the comparison with the reference data for velocity estimation. Only the velocity estimates for vehicles #4 and #5 show a relative poor accuracies, which could be tracked back to the fact that the undersized vehicle (vehicle #4) and degraded laser point sampling due to occlusion of (vehicle #5).

To summarize the overall performance, the experiments tend to show that the motion detection rate as well as the velocity estimation accuracy could drop down as vehicles become small or move slowly. This could be mainly caused by the limited spatial resolution and the high sensitivity of the shape parameterization process to the noise and anomaly. Although Toronto dataset delivered relatively good performance in terms of motion classification when considering the results

quantitative evaluation only, it does not necessarily mean that the low point density could not hinder the motion classification. Current airborne LiDAR data have the nominal resolution ranging from 0.5m to 1m, so that, in certain situations, vehicle's points are mixed with the background objects. Hence, the point sampling rate as well as the backscattered energy is biased towards smaller values, which makes motion detection and velocity estimation less robust and reliable. Furthermore, according to the quantification of the motion artifact model in Section 3.3.1.2, vehicles moving with a low velocity would induce minor shape deformations and be more likely to have an inaccurate parameterized shape which can degrade the estimated velocity too. Fortunately, the problem is also to some extent alleviated, since vehicles with unambiguous shape have already been removed prior to the motion classification step and directly labeled as uncertain class. Although there are still some inaccuracies in the estimated velocities, they are able to provide a good representation of the overall traffic situation and the velocity distribution in the scanned scenes.

6.1.5   Comparison of vehicle extraction method towards motion analysis

Through the visual analysis in Figure 47, Figure 48 and Figure 49, the overall performances of the strategies using two methods of vehicle extraction - local context guided method (vehicle extraction method I) and (vehicle extraction method II) appear to be promising and make no huge differences. It is to be found out that the strategy based on method I to extract vehicles achieved a better vehicle detection rate, e.g. for dataset Toronto III, with the completeness of ca. 79%, being 6 percentage point higher than that achieved by the OBPA method, while the reliability is almost equivalent for both methods. However, the method II shows a high geometric precision of extracted vehicles with a RMS$\{H(E,R)\}$ of 0.34m implying a increase in shape preservation accuracy by 15%. For both datasets there is a better agreement between the motion classification results and the reference for most of vehicles when using method II (e.g. approx. 83% vs. 87% for detection rate of moving class in dataset Toronto III). The detection reliability has also increased to some extent (e.g. approx. 74% vs. 82% for detection rate of moving class in dataset Toronto III). Both travelling vehicles on roads and parking ones are successfully extracted, including those that were placed along the road margins or in parking areas.

The visual impression of the results is also reflected in the quantitative evaluation (Table 9 and Table 10). By further observing the quantitative results of evaluations in, it is to be found out that the strategy based on the local context guided method to extract vehicles achieved a better vehicle detection rate, while the reliability is almost equivalent for both methods (Table 9(a) and Table 10(a)). However, the OBPA method shows a higher geometric precision of extracted vehicles. Most of additional vehicles extracted by method I originated from the areas where vehicles are located very close to each other, such as parking lots. The method II tend to be more likely to extract fewer vehicles that are placed as scattered despite of having vague height edges in relation to ground surface, such as those on the road lanes. It may be owing to the fact that the method II using 3D segmentation based classification is able to accurately detect scattered geometric modes at local level by adaptive MS clustering, while the vehicle extraction method I based vehicle extraction relies more on the quality of detected makers for potential single vehicles controlling the segmentation. Clearly, vehicle extraction directly in 3D ALS point cloud results in overall a higher accuracy of geometry preservation. However, according to the evaluation results for dataset Enschede used for comparative study the increase of the point density seems to help to some extent

relieve the difference of motion classification accuracy between two methods, as it can be easily imaged that the high point density of ALS data could facilitate the vehicle extraction operation by high geometric fidelity and improve the extraction results in terms of the shape accuracy whatever kind of methods is used.

As shown in the results and their evaluation (Table 9(b) and (c) and Table 10(b) and(c)), for both datasets there is a better agreement between the motion classification results and the reference for most of vehicles when using method II. The detection reliability has also increased to some extent. The geometric accuracy of extracted vehicles can directly indicate the difference between two methods, since the distinction of motion state is performed based on measuring and examining the geometric properties of vehicle points. False alarm and misdetection rates of classified moving vehicles are reduced for method II, since either additional moving vehicles have been newly extracted (e.g. vehicles on the road on the right border of Figure 47(a) and Figure 48(a))or the high geometric accuracy of extracted vehicles enable the motion classification step to be performed more accurately (e.g. vehicles on the road of Figure 49). Some small-sized cars slowly moving can be detected only using method II where their shape parameterization process is very sensitive to the noise. Apart from the increased quality of moving class method II has actually yielded fewer vehicles detected as stationary. That means that it has not proved to provide the same improvement in the classification quality for the stationary class, although the traffic monitoring research focuses on the extraction of dynamical information. That is also to say that method II has improved the performance of motion indication using airborne LiDAR by sacrificing the detection rate and reliability of stationary vehicles, especially in the places where vehicles are located very close to each other, such as large parking lots.

Contrarily, for the motion analysis using method I, although the number of detected motionless ones increased, the false alarm and misdetection rates for moving ones still remain unimproved. On the one hand, the completeness for the detection of motionless vehicles is increased when the cardinal number of extracted vehicles expands. On the other hand, new detection of vehicles using method I usually appear in parking lots and somewhere vehicles are densely placed to each other, which also increases the risk of false or missing classification of the motion state due to the poor geometric outlining ability. It is more likely for method I to generate a more complete vehicle extraction result with relatively poor geometric accuracy which has potential to deliver the promising data basis for the stationary vehicle class other than moving class. Since airborne traffic monitoring intends to provide the complete overview on traffic scenario and information, not only moving vehicles but also stationary ones are required to be extracted for intelligent traffic management. The complete and accurate detection of stationary vehicles is not a trivial thing. Therefore, the selection of two vehicle extraction methods for motion analysis using airborne LiDAR is the tradeoff between the completeness and diversity of traffic information and the reliability and accurateness of dynamical component of traffic information. Moreover, the number of vehicles of uncertain motion state has increased by averagely 20% when using method I, particularly concentrating in the parking areas, it has again proved that method I could yield more outliers of vehicle extraction for the stationary class.

The velocities of moving vehicles of the two datasets derived based on two vehicle extraction methods are denoted or listed in Figure 47(b), Figure 48(b) and Table 11, respectively. Since the reference data is unavailable for velocity estimation in dataset Toronto III, the objective evaluation

is impossible. Nevertheless, the interrelationship between the spatial distribution of estimated velocity and the spatial contexts of vehicles in the scene could help to infer the quality of velocity estimation to some extent. Generally, the estimated velocities are representative for the vehicles travelling on this section of the urban road and correspond well to the speed limit. According to empirical analysis, there are some anomalies among the moving vehicles on the road margins which were detected as travelling with distinctly too fast velocities (>50km/h) in the result obtained using method I (Figure 47(b)). Although these vehicles were leaving the parking strips for the main lanes of roads or traveling densely behind another, they were not possible to travel with such high velocities at that time. The velocity estimates derived via method II (Figure 48(b)) showed more reasonable values ((<= 40km/h)) for these vehicles. Additionally, the moving vehicles on the normal urban road (across the bottom part) using method I show the mean velocity being a bit lower than the moderate tempo limit and the maximum of the velocity range being far beyond the allowed in urban areas such as Toronto. For dataset Enschede there is a reference which can be used to evaluate the velocity estimates more objectively. The quantitative evaluation results presented in Table 11 has further indicated the similar relationships of velocity estimation accuracy between two vehicle extraction methods. The residues of velocity estimates by using method I are generally larger than that of method II by averagely 3%, which is rooted in the better ability of method II in delineating the vehicles' geometry. Based on empirical and quantitative evaluations concerning velocity estimation, it can be summarized that the vehicle extraction method I tends to reduce the estimation accuracy and under/over-estimate the velocity compared to method II. Furthermore, the variation of point density of datasets appears to have an inconsistent influence on the estimation accuracy. When the point density becomes low, the difference of the estimation accuracy between two methods would become sufficiently large needed to be considered. However, this difference would be reduced to such level (<10%) that can be neglected when the point density becomes larger than 7 points/m^2.

6.2 Performance analysis for motion detector

6.2.1 Analytic performance analysis

To analyze the quality and performance of the vehicle motion detection approach a theoretic analysis method is firstly applied. There are two hypotheses for the vehicle motion state used in our applications – stationary and moving vehicles. The probability density functions (PDFs) of two motion states regarding vehicle shape parameters are to be inferred and summarized as below:

➢   **Stationary vehicles**

For all stationary targets the values of sensed aspect ratio Ar_s and shearing angle θ_{SA} are assumed to be statistically distributed around the expectation value $E\,[Ar_s] = 2 \;\wedge\; E\,[\theta_{SA}] = 90°$. Using the underlying assumption of two dimensional jointly Gaussian-distributed data for the space of shape parameters, the joint probability density function (PDF) $f_s\,(Ar, \theta_{SA})$ of the sensed aspect ratio and the shearing angle of a detected stationary vehicle is given by bivariate normal distribution as

$$f_s(Ar_s, \theta_{SA}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{z}{2(1-\rho^2)}\right] \tag{61}$$

where $z = \frac{(Ar_s - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(Ar_s - \mu_1)(\theta_{SA} - \mu_2)}{\sigma_1\sigma_2} + \frac{(\theta_{SA} - \mu_2)^2}{\sigma_2^2}$, $\rho = cor(Ar_s, \theta_{SA})$, which is the correlation coefficient between Ar_s and θ_{SA} μ_1, μ_2 and σ_1, σ_2 are mean values and standard deviations of Ar_s and θ_{SA}, respectively.

For airborne LiDAR data of medium resolution and above the jointly Gaussian assumption can be easily validated for traffic in most urban areas (Yao et al., 2008). Figure 50(a) shows a typical example of the PDF $f_s(Ar_s, \theta_{SA})$ assuming a point density of 4.5 points/m^2. Figure 50(a) also illustrates the typical behavior of the stationary class that large fluctuations of the aspect ratio are associated with small-sized vehicles while the shearing angle variations are drastically reduced for large sized vehicles.

Based on this PDF of the hypothesis for the stationary vehicle class a constant false alarm rate (CFAR) detector can be designed that groups all extracted vehicle instances into two classes. Class I, called 'stationary vehicle only', should contain all vehicle objects being motionless while scanning-over. Class II, called 'non-stationary vehicle', and should contain all vehicles apart from class I. This second class includes objects that contain moving vehicles but also some sort of outliers. Since the PDF of this second class is at the moment not known, it is assumed to be equally distributed over a large area. With this assumption a likelihood ratio can be computed. Classification is done by comparing that likelihood ratio with thresholds α. This approach provides curves of separation between the two classes, which are actually isolines on $f_s(Ar_s, \theta_{SA})$. The chosen curve of separation determines the probability of false alarm (P_{fa}); often also referred to as 'false alarm rate' (**FAR**). It is simply the integral of $f_s(Ar_s, \theta_{SA})$ over the area where $f_s(Ar_s, \theta_{SA}) < \alpha$.

➢ **Moving vehicles**

The moving target shape is assumed to have a peak value of θ_{SA} other than 90° and a dispersion of the aspect ratio Ar_s from 2 to more or less. By considering various groups of moving vehicles appearing in scenario of urban traffic and assuming them to equivalently move along all the possible directions (usually consists of two mutually perpendicular directions due to the structure of road network in urban areas) with representative velocities (e.g. 60 km/h), a new class describing the superposition of shape deformation effects of multiple moving vehicles, called 'moving vehicle' can be introduced here. Unfortunately, a PDF $f_m(Ar_s, \theta_{SA})$ describing the probability density of this class cannot be explicitly determined. An approximation for $f_m(Ar_s, \theta_{SA})$ valid for single-pass ALS data, can be given by the kernel density estimation technique based on a large amount of samples. Approximated PDF of the moving vehicle class is estimated based on an extensive range of real-life datasets acquired in large urban areas. It is stated that it provides sufficient representativeness from the viewpoint of statistical sampling and can generalize the analytical description ability for the alternative hypothesis within the framework of vehicle motion classification.

• Approximation for the PDF $f_m(Ar_s, \theta_{SA})$ of moving vehicle class

Kernel density estimation (or Parzen window method) is a non-parametric way of estimating the probability density function of random variables. Given some data about a sample of a population, kernel density estimation makes it possible to extrapolate the data to the entire population. For a bivariate random sample **X₁, X₂,** … **Xₙ** drawn from a density *f* of moving vehicle samples, the kernel density estimate for PDF of moving class $f_m(Ar_s, \theta_{SA})$ is defined by

$$\hat{f}_m(x, \mathbf{H}) = n^{-1} \sum_{i=1}^{n} K_{\mathbf{H}}(x - \mathbf{X}_i) \tag{62}$$

where $x = (Ar_s, \theta_{SA})^T$ and $\mathbf{X}_i = (Ar_s^i, \theta_{SA}^i)^T$, $i = 1,2,...,n$. Here $K(x)$ is the kernel for the estimation which is a symmetric probability density function, $\mathbf{H}$ is the bandwidth matrix which is symmetric and positive-definite, and $K_{\mathbf{H}(x)} = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}x)$. The sample $\mathbf{X}_1$, $\mathbf{X}_2$, ... $\mathbf{X}_n$ can also be drawn from the population generated by numerical simulation based on the motion artifacts model (Equations (27) and (29)) by adapting relevant parameters to urban traffic scenarios. For instance, the velocity of vehicle v takes values drawn from a population obeying a normal distribution with the mean of 60km/h(stochastic sampling), and the intersection angle θ_{SA} take values obeying a discrete uniform distribution as vehicles could be moving in any directions in a broad sense when attempting to derive a general PDF. The good agreement of theoretic PDF and real cases can be verified through this estimation process. The choice of K is not crucial: we take $K(x) = (2\pi)^{-1} \exp(-\frac{1}{2} x^T x)$ the standard normal throughout. In contrast, the choice of $\mathbf{H}$ is crucial in determining the performance of $\hat{f}_m$. The most common parameterizations of the bandwidth matrix are the diagonal and the general or unconstrained which has no restrictions on $\mathbf{H}$ provided that $\mathbf{H}$ remains positive definite and symmetric, that is

$$\mathbf{H} = \begin{bmatrix} h_1^2 & 0 \\ 0 & h_2^2 \end{bmatrix} \text{ or } = \begin{bmatrix} h_1^2 & h_{12} \\ h_{21} & h_2^2 \end{bmatrix} \tag{63}$$

This latter parameterization allows kernels to have an arbitrary orientation whereas the former only allows kernels which are oriented to the coordinate axes. Here, the bandwidth is determined by the rule of thumb bandwidth (Bowman and Azzalini, 1997), the number of samples n is set to 2000.

Once the probability density functions of both hypotheses (stationary and moving vehicle) are given in analytic form, applying a CFAR detector of the given design for detecting moving vehicles is a natural way to make a binary decision. Upon the estimation of approximated PDF $\hat{f}_m$ and using this approximation as an alternative hypothesis, $\hat{f}_m(Ar_s, \theta_{SA})$ allows to define a likelihood ratio to which a threshold can be applied. These lines are not isolines of $f_c(Ar_s, \theta_{SA})$ anymore, but they separate the class 'moving vehicle' better from the class 'stationary vehicle only', which can also be verified by the ROC curves of two applied motion detectors in Figure 51. Figure 50(b) shows an example of the shape of $f_m(Ar_s, \theta_{SA})$ together with $f_m(Ar_s, \theta_{SA})$ and a corresponding curve of separation.

However, the classification scheme based on the joint PDFs of two hypotheses can be potentially performed in optimal cases when the sensed aspect ratio and shearing angle of a possible moving target in scanned scenes is uniformly distributed. This does not hold for many military applications, where vehicles are not bound to roads and can move in any arbitrary direction. In case of public traffic in urban areas, where a-priori restrictions on velocity and movement direction of vehicles is available to a certain degree depending on the topology of road network and tempo limit on that road section, the use of a CFAR motion detector is assumed to perform better. It is possible to derive the expectation values for velocity and moving direction of vehicles using ancillary data such as GIS and road database. Furthermore, this a-prior information can be integrated into the estimation of the empirical PDF for the alternative hypothesis of moving vehicle prior to applying the CFAR detector. The incorporation of a-priori information into the vehicle detector improves the amount of detected targets and also reduces the number of false alarms.
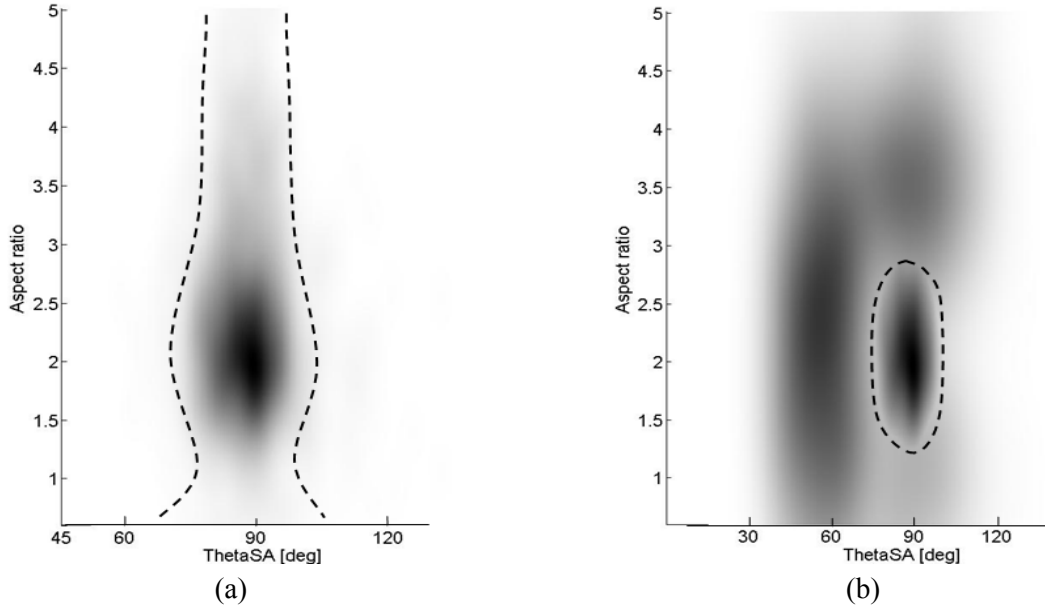
Figure 50: PDF for vehicles of two motion states(a) Theoretical PDF $f_s(Ar_s, \theta_{SA})$ of sensed aspect ratio and shearing angle with $E[Ar_s] = 2$ and $E[\theta_{SA}] = 90°$. The point density is set to 4 points/m^2, (b) approximation of theoretical PDF of $f_m(Ar_s, \theta_{SA})$ and its position relative to the hypotheses $f_s(Ar_s, \theta_{SA})$. The dashed lines are examples for curves of separation. θ_{SA} is normalized into the acute angle range ($\leq 90°$) for simplicity

If the probability density functions of both the Null-hypothesis (stationary vehicle) and the alternative hypothesis are given in analytic form, Receiver Operating Characteristic (ROC) curves can be calculated analytically for every set of parameters. By that, the performance of the detector can be estimated and assessed for all traffic situations. As shown in this section, an analytical description of the Null-hypothesis exists, if homogeneous distribution of vehicle size can be assumed (c.f. Equation (61)). Unfortunately, only approximations for the PDF $f_m(Ar_s, \theta_{SA})$ corresponding to the alternative hypothesis are available up to now (c.f. Equation (62)), which are obtained from the viewpoint of statistical sampling theory. Thus, the use of analytical descriptions for the alternative hypothesis is to some extent applicable for this study. With preliminary knowledge of the alternative hypothesis the performance of the system for motion can be quasi-fully described, as not only the probability of detection can be accurately estimated but also the probability of false alarms remains known.

To evaluate the performance of the motion detector based on analyzing theoretic PDFs in the 2D Euclidean feature space spanned by two shape parameters, the threshold is varied and the probability of motion detection and probability of false alarm are determined for each step of this variation, finally resulting in ROC curves. Figure 51 show the theoretic ROC curve of the motion detector derived based on two estimated analytic PDFs for motion detection. For motion detection using joint PDFs of two vehicle classes, the risk of falsely detecting an outlier is reduced and the probability of false alarms can be decreased as well compared to using the PDF of the class of stationary vehicle only to do the separation. Nevertheless, the nonlinear variability of vehicle shape features defined for resolving the motion state still dominate characteristics of the feature space used for the classification where the large and irregular overlapped area between two hypotheses leads to severe difficulties in distinguishing between moving and stationary vehicle classes. The ROC curve obtained in this way give us an impression that common classification method based on analyzing

the likelihood ratio of moving class to stationary class using the Euclidean distance metric could not yield promising and reliable results towards the accurate distinction of vehicle motion states, despite that it is only limited to a binary classification task.
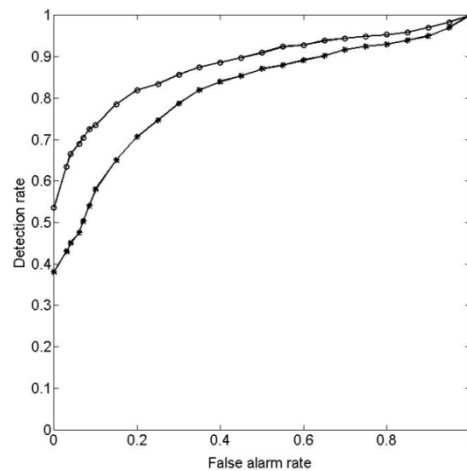

Figure 51: ROC curves of a CFAR- motion detector based on analyzing one PDF of stationary class and joint PDFs of two motion classes, respectively

6.2.2 Experimental performance analysis

To evaluate the quality of the vehicle motion detector (since binary classifier is actually a detector) in a numerical way, one can implement a Lie group distance based motion detector based on simulating LiDAR vehicles points, whose outputs should be ROC curves for the detection of moving vehicle class. For each simulation, a set of 300 simulated instances of vehicle points consisting of 150 stationary and 150 moving vehicles whose velocity are sampled using stochastic technique (Monte Carlo simulation) from a normally distributed population controlled by the tempo limit in urban areas are generated as experimental dataset to which the motion classification strategy is applied. By varying the thresholds, detection rates and false alarm rates are calculated from motion classification results, respectively, and further delivered to yielding ROC curves. This simulator can be parameterized in such a way that acquisition conditions and configurations of simulated LiDAR points of vehicles such as the point density and the moving direction of vehicle relative to the sensor flight path can be flexibly adjusted according to user's settings. Hence, the influences of varying characteristics of LiDAR data on the performance of vehicle motion detection can be quantified. Configuration for LiDAR data simulation is referred to Table 12 which is one of typical sensor and flight configurations used for city mapping tasks using airborne LiDAR.

To evaluate the performance of the detector, the threshold is varied and the probability of detection and probability of false alarm are determined for each step of this variation, finally resulting in ROC curves. Figure 52 show the ROC curves of the motion detector for different point densities of vehicle points (Figure 52(a)) and for different relative moving directions (Figure 52(b)). The comparison of the detector with respect to data characteristics and vehicle behaviors shows a clear advantage for the one using data of high point density. Especially in bad detection scenarios the improvements tend to be larger. It is easy to understand that more detailed information of the vehicle shape can be preserved for accurate classification. Datasets whose point density is less than 4points/m^2 can hardly provide reliable results. As shown in Figure 52(b), that tends to tell us that the

shearing effect of vehicles is more robust than the shape stretching effect of vehicles with respect to distinguishing motion states, as the shape deformation of vehicles moving in the intersection angle θ_v of ca. 60° is dominated by the shearing effect (maximum of the θ_{SA} occurs when $\theta_v \approx 60°$) while elsewhere it is more or less dominated by the stretching effect (as indicated in the quantification of motion effects in Section 3.3.1.2). The inherent shortcomings of vehicle's original aspect ratio *Ar* that cannot be accurately determined and is sensitive to noise account for relative poor performance of the motion detector in such case that the shape deformation of vehicles is dominated by the stretching effect.

Figure 53 illustrates the detection probability using the proposed binary Lie classifier for motion detection (a) over different vehicle velocities for certain vehicle sizes as well as (b) for different relative moving direction, both with fixed false alarm rate. As can be seen, for high velocities and large vehicles (equaling to high point density) the proposed motion detector delivers generally better results. The reason for this behavior is that the motion detector purely relies on the shape information, i.e. for data of low point density the detection is strongly influenced by noise or shape ambiguity, which leads to the significant decrease in performance. In contrast, high velocities evoke more significant motion artifacts effect deviating from normal vehicle structures as assumed, so that, feature values with more distinguishing ability are delivered to detecting a moving vehicle. For moving vehicles with different intersection angles θ_v and yet the same point density, the best performance of motion detection is found to be depending on the vehicle velocity. For instance, the detection rate of the best case that vehicles move in direction of θ_v =30° at high speed (e.g. 120 km/h) is almost improved by 15% compared to the worst case that vehicles move in along-track (θ_v =90°), where this improvement is likely to disappear if the velocity has decreased to ca.60 km/h. It could be traced back to the fact that the shearing effect of vehicle shape artifacts obtained under airborne LiDAR scanning mechanism is maximized towards the along track direction as the velocity of moving vehicles is approaching the flying velocity of the sensor platform.
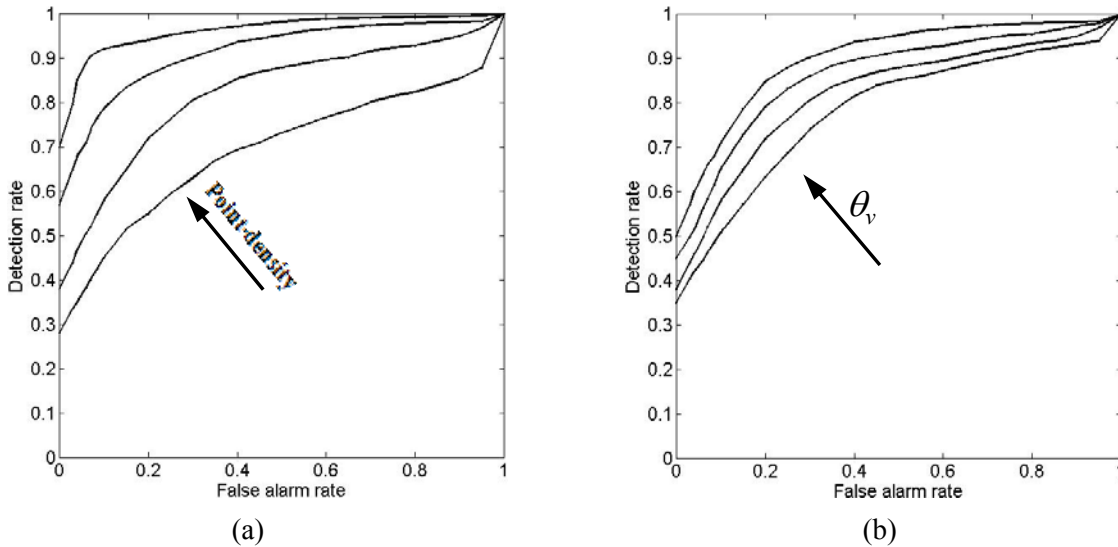

(a)   (b)

Figure 52: ROC curves of the motion detector using the Lie group metric (a) for point densities of (from innermost to outermost) − 2 points/m^2, 4 points/m^2, 7 points/m^2, and 10 points/m^2, (b) for different relative moving directions of vehicles (from outermost to innermost: $\theta_v = 0°$, $\theta_v = 30°$ $\theta_v = 90°$, and $\theta_v = 60°$). Here, the point density is fixed at 4 points/m^2
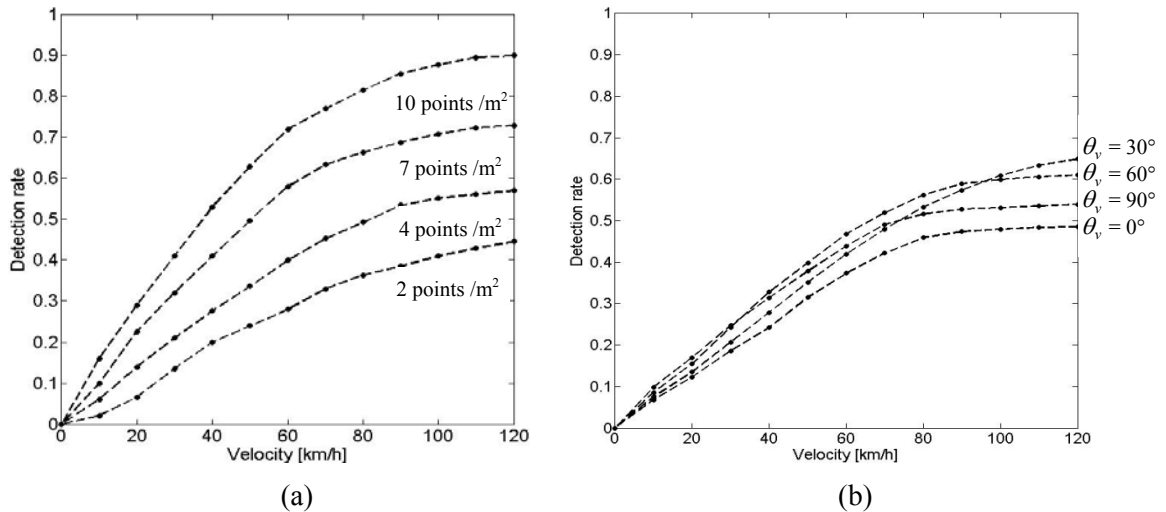
Figure 53: Numerical Detection Characterization: detection probabilities for given fixed false alarm rate (10e–2): (a) results calculated for varying vehicle sizes (i.e. point density), (b) results calculated for varying relative moving directions; Here, the point density is fixed at 4 points/m^2

6.3 Accuracy prediction for velocity estimation

Finally, to demonstrate the quality of the velocity estimation for real-life scenarios and enable to deliver a valuable and quantitative guidance on the planning of future flight campaigns for airborne LiDAR data acquisition towards traffic analysis, true data of road networks in urban areas are chosen to undergo simulated experiments on the prediction of velocity estimation accuracy. It is useful for exploiting boundary conditions of applying the proposed strategy in real airborne LiDAR campaigns for traffic analysis. Generally, it can be stated that this experiment is designed by considering following points:

- validate the feasibility and repeatability of velocity estimation results
- verify the velocity estimation scheme which can be used to provide rational estimates with sufficient accuracy in wide range of datasets acquired over urban areas
- show the potential of velocity accuracy estimator developed in this work that it would have a great impact on guiding and assisting the flight planning

Flight height	h	420 m
Pulse repetition rate	PRR	110 kHz
Sensor velocity	v_L	120 km/h
Scan angle	α_s	60°
Point density	PD	4.5 points/m^2
Swath	S_w	450 m
View mode		Nadir
Scan pattern		parallel line

Table 12: Configuration parameters for airborne LiDAR acquisition used in the simulation

The standard deviation of the estimated velocity σ_v is simulated for two road network sections north of Munich which are supposed to be able to characterize most of typical scenarios in urban

areas. In this area, several main roads and large express roads are situated which are highly frequented during rush hours. Four velocity estimators are applied to this test: (i) shearing-based estimator, (ii) stretching-based estimator, (iii) integrated shape deformation based estimator and joint estimator, (iv) joint estimation of velocity and moving direction. For each test site, two general schemes are assumed to exist for velocity estimates:

(1) The moving direction of vehicle relative to sensor flight path is known
(2) The moving direction of vehicle relative to sensor flight path is unknown

Figure 54: Simulation of standard deviation of velocity estimates σ_v on two road networks north of Munich using the velocity estimation schemes: (a) and (b) show the estimation accuracy for the first road network in % of the absolute velocity using the first and second schemes, respectively, (c) and (d) show the estimation accuracy for the second road network in % of the absolute velocity using the first and the second scheme, respectively

As three methods within the first scheme complement each other in terms of performance we finally combined the estimators depending on the relative orientation between road and flight path to get

optimal results. For every relative orientation such the estimator that provides the best results is chosen. That means that the maximum of estimated velocity accuracies is assumed to be selected as the accuracy value for velocity estimate of that road location. Real flying and sensor parameters using Riegl LMSQ560 LiDAR sensor have been used in this simulation and an average speed of 120 km/h was assumed (configuration specifications can be found in Table 12). The average velocity of moving vehicles on the roads is set to 60 km/h. The error measures for the shearing angle and intersection angle of moving vehicles can be assessed empirically from shape parameterization , for our case, $\sigma_{Ars} = \sigma_{\theta SA} = 2°, \sigma_{\theta v} = 2°$. The orientation of the roads relative to the planned flying path and the resulting σ_v values obtained by combining the estimators in the first scheme are shown in Figure 54(a) and (c), while the resulting values σ_v using second scheme for the same sites are shown in Figure 54(b) and (d). σ_v is given in % of the absolute velocity as relative accuracy.

(a)   (b)
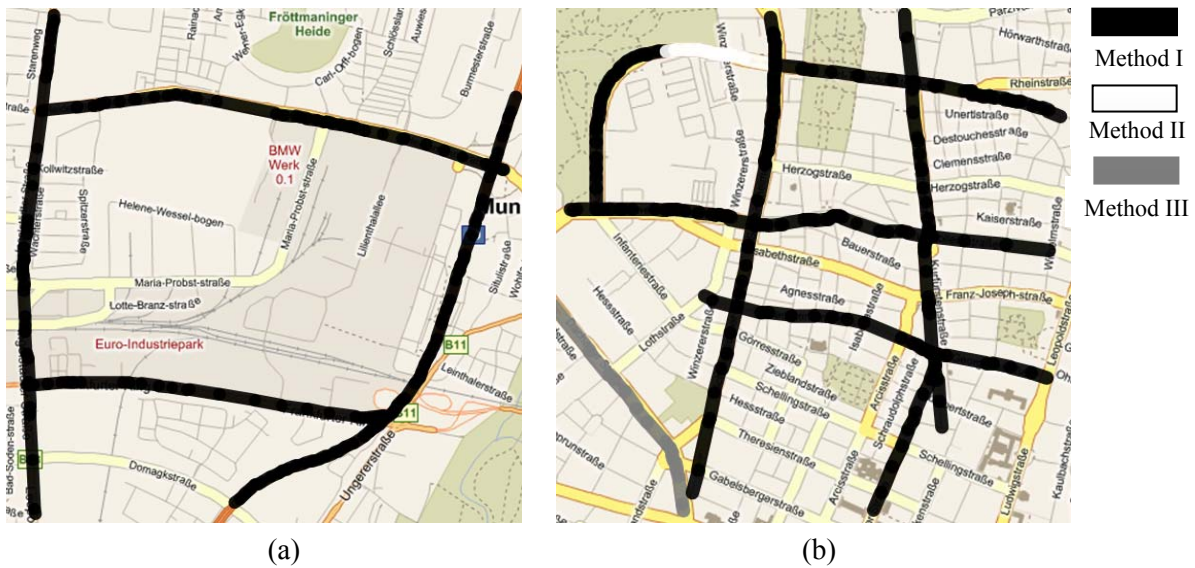
Figure 55: Indication of velocity estimation methods used for the two road networks under the first velocity estimation scheme: (a) indicates which estimation method is chosen in which parts of the first road network, (b) indicates which estimation method is chosen in which parts of the second road network.

With the presented algorithm velocities can be estimated with an accuracy better than 10% for ca. 80% of the investigated road networks. Figure 55 indicates which estimator is chosen in which parts of the road network. It shows that the across-track shearing-based estimator (method I) provides best results for large parts of the road network. The along-track stretching-based (method II) / combined (method III) estimator outperforms the across-track shearing-based approach only in areas where the road is extended nearly in the along-track direction (i.e. $\forall \theta_v \leq 25°$), e.g. in the second test site, Dachauer street (in the bottom-left part) requires the method III to be used for velocity estimation, whereas one part of Ackermann street (curved, in the top-right part) requires the method II to be performed. Moreover, in most parts of the road networks, the accuracy of velocity estimation using the first scheme is generally higher than that obtained using the second scheme, especially when vehicles move along a direction which is close to the across-track. This is due to the fact that the joint estimation of velocity and moving direction angle could incorporate additional error sources caused by the unknown relative moving direction of vehicles to sensor flight path, leading to an accumulative error for final velocity estimates.

7 Conclusions and outlook

The overall objective of this thesis was to develop and investigate a procedure for automatic extraction and motion estimation of vehicles from airborne LiDAR data with a focus on urban areas. It is assumed that the point spacing of datasets should be better than 1m. The dominant consideration in designing such a system that vehicles can be extracted with their motion states simultaneously is how to manipulate the nonlinearity and vulnerability of vehicle shape in ALS point cloud of urban scenes. In order to be able to extract vehicles with sufficient completeness and geometric accuracy the system builds on a comprehensive context-guided approach and explicit 3D handling of the object "vehicle" in the point cloud. The vehicle model integrates essential knowledge about geometric structures such as areas, planarity, elongatedness and vertical extension with abstracted knowledge about the contexts. Moreover, local context modeling can acquired links between a vehicle and the properties of immediately adjacent objects while the modeling of the global context can take the dependence of the shape and appearance of a vehicle on different scene structures in a large-scale environment in to account.

Another feature of the presented approach is the development and application of a systematic scheme for motion detection and estimation of vehicles. The scheme is established based on the motion artifacts model which is introduced in this thesis as analytic functional relations between observed shape parameters and the velocity of vehicles. And criterion that characterizes whether the direction of an object motion is known or not divides the motion estimation method into two general categories. First, the motion direction is assumed to be derived in advance; and in the latter case, the motion direction is required to be estimated together with the velocity quantity. Provided that the criterion of the two categories can independently represent knowledge about moving direction of objects, the accuracy assessment based on the motion artifacts model allows to rate the confidence of velocity estimates for the object motion. Particularly in the treatment of complex scenes, the velocity estimation is of great importance according to their level of confidence, because it can be inclined to indicating the promising estimation method, thereby making the motion estimation more reliable and robust.

The experiments are given and discussed to apply and validate the whole strategy to a variety of real-life scenarios. The usability of the proposed strategy has been successfully demonstrated with the extraction and motion estimation of vehicles in ALS data of urban areas. The fact that satisfactory results for both vehicle extraction and motion analysis can be achieved even in the case of very complex inner city areas also confirms the efficiency of the system. As difficulty exposed to the strategy, the poor or ambiguous geometric shape of scanned vehicles and their vague spatial relations to surroundings in the point cloud are identified, which could degrade the vehicle extraction results and accuracy of motion estimation. One of possible ways to remedy this problem is to perform the vehicle extraction module upon the non-vehicle areas in an iterative way. The extracted vehicle points could therefore serve as valuable approximate or a-prior knowledge for vehicle extraction in the next step.

The implementation of the strategy for extraction and motion estimation of vehicles comprises the

three stages: (1) contextual analysis, (2) extraction of vehicle objects, and (3) motion classification and velocity estimation. For each stage useful comments and considerations are summarized as follows

(1) The global and local background information is obtained by the contextual analysis, which is very important for the workflow of the vehicle extraction. The contextual analysis includes, on the one hand, the distinction of the global context areas between "areas with elevated road" and "areas without elevated road" for evoking the proper *Route* of vehicle extraction*,* and on the other hand, the extraction of local context objects to facilitate vehicle extraction by providing the **RoI**. Besides the accurate determination of ground surface, the exclusion of objects with similar spatial contexts but not belonging to the ground surface such as low vegetations and traffic signs as well as the definition and modeling of knowledge about vehicles' shape and positional properties plays a key role in the phase of vehicle extraction towards reduction of false alarms.

(2) In the extraction of vehicle objects the goal are followed to extract most reliable parts in point clouds corresponding to the vehicle object with the help of the integrated usage of multiple kinds of information. To this end, on the one side, the extraction of ground surface that should construct most potential areas where most of vehicles should appear is applied prior to isolation of singe vehicles by morphological segmentation; on the other side, preliminary point sets of single vehicles are directly detected as local geometric modes from point clouds which are further delivered to a binary classification to acquire the semantic label. For more challenging scenes the two vehicle extraction are combined to generate an improved result especially in view of the completeness for those vehicles of dense placement. The experiments showed that the context-guided method is more likely to extract vehicles of dense placement such as in parking lots more completely whereas the OBPA method tends to be more capable of extracting vehicles with high geometric accuracy and complement the context-guided method by extracting vehicles on elevated roads. Vehicle extraction is built upon the pure geometric information of ALS data without utilizing and calibrating the laser reflection intensity, thereby giving us a great flexibility and extendibility in data handling. Forward-looking ALS data seem to be inappropriate for vehicle extraction due to the unstable point sampling and shadow effect, which can further hinder the motion analysis in the next step.

(3) The motion classification process based on 3D shape categorization consists of two steps: (i) the parameterization and generation of vehicle shape description measures in view of motion artifacts model in ALS data, (ii) integrating them into a nonlinear classification framework working on the Lie group space. It could be stated that the classifier based on Lie group metric outperforms other common classification method working on the linear Euclidean space concerning distinguishing vehicle motion state, since the shape deformation measures defined towards the motion artifacts are of nonlinear essentiality to construct a heterogeneous feature space and can better be resolved by classifier working on manifold. The experiments also showed that the motion detection using the OBPA method for vehicle extraction can improve the performance of motion analysis compared to that using the context guided method; it is due to the fact that the OBPA method can extract vehicle objects with higher shape fidelity by handling point cloud directly instead of gridded data. However, despite of the lower percentage classification results of motion detection, the context guided method can extract more vehicles which can be classified as stationary, especially in areas where vehicle objects are concentrated such as parking lots, to provide more complete traffic information in view of monitoring applications. It can be found out by theoretic accuracy analysis and simulation study that point

density of ALS data, vehicle velocity, moving direction of vehicles could have different impacts on the performance of motion detection. On the one hand, the denser the point cloud is and the faster vehicles move, the more accurately moving vehicles can be detected, on the other hand, neither vehicles moving along-track nor across-track can be detected with the largest probability. The velocity of moving vehicles is estimated by inverting the analytic model of motion artifacts. Last but not least, the accuracy of velocity estimates can be derived by exerting the error propagation law onto the inverted functional equations for the motion artifacts model. It has been found that none of the velocity estimation methods can always guarantee the best quality for single vehicles under various observation conditions. The method by combing two motion components is most reliable method due to its generality to various circumstances, which could provide optimal results for a scanned scene at global level. The accuracy of velocity estimation seems to be proportional to the shape accuracy and point density of single vehicles. Moreover, the theoretic accuracy analysis of the motion artifacts model can be used to empirically predict the performance of velocity estimation using airborne LiDAR depending on the orientation of road networks, tempo limit and flight configurations, therefore enabling the optimal planning of the flight campaign in view of traffic monitoring applications.

The research work of this thesis is driven by the basic idea that traffic-related information like traffic density and vehicle velocity could be directly derived relying on the so-called motion artifacts in single-path airborne laserscanning data. Our research work is focused on the design and realization of a traffic monitoring concept under LiDAR acquisition condition for city surveying and modeling tasks, which is able to provide the value-added product along with common products such as building models, DEM and canopy height models. Therefore, no extra expenses will be produced, even if the strategy would fail to extract adequate vehicles. The purpose of the vehicle extraction algorithm intends to automatically provide the data basis which can further be delivered to the step of motion classification and velocity estimation based on shape deformation (artifacts) analysis.

In addition to possibilities for further development of the presented approach directly derived from the analysis of the results, such as improvement of the quality of vehicle extraction and motion analysis by introducing more advanced methods, there are some more general questions concerning the scientifically deep handling of this topic. Below are some selected open questions that can be taken up in future:

➢ Most of airborne LiDAR systems can operate in different modes concerning data acquisition. Scan mode plays an important role in our application. The motion artifacts model is constructed based on the assumption that ALS data are scanned by linear motion of the laser beam, which is nowadays most widely used. All the experimental datasets used in this thesis were acquired by "parallel line" scan pattern. It remains as an open problem that whether the proposed strategy for vehicle motion analysis can be adapted to other scan patterns such as "zigzag", "elliptic" and "burst line", without loss of the generality and accuracy for motion detection and estimation. For elliptic scan pattern generated by palmer scan it is obvious that the presented motion artifacts model is not valid yet and needs to be revised according to the scanning trajectory of laser beam. However, data of palmer scanners are expected to show a more distinct artifacts effect for moving objects which could be used to improve the motion detection rate and velocity estimation accuracy.

➢ New possibilities also arise due to diversity and advances in airborne LiDAR sensor technology. First, the full-waveform LiDAR has emerged as a standard instrument and it would be interesting to investigate the question of whether the analysis of the waveform information of laser echo signals can compensate for the lack of geometric information caused by partial occlusions to contribute to improving vehicle extraction beneath trees. Additionally, extracted vehicles with more complete point samplings are expected to yield more complete and accurate traffic dynamical information. Second, airborne LiDAR system which is simultaneously equipped by three (forward, nadir, backwards) scanners and can scan the scene with a constant time delay to form three LiDAR snapshots sampled along the temporal axis has recently emerged, one example can be found in Oude-Elberink and Vosselman, (2009) and Vosselman (2009). It is deserved to investigate the capability of such systems to detect moving objects and estimate their velocities, since moving objects can be recorded at different positions which leads to a direct conclusion for motion. The correspondence between multiple instances of a vehicle in three datasets appears to be a challenging task for estimating their motion. Finally, for ALS systems of next generation, it is suggested that the scan pattern and trajectory are controllable and can be programmed in advance, so that the corridor mapping campaign can be realized for promoting traffic monitoring applications by focusing on roads of interest which is guided by maps. The motion sequence of the laser beam could be adaptively adjusted and optimized for extracting object motion in accordance with the orientation of local road networks and tempo limit.

➢ Another potentially challenging task in future research remains in the quantification of qualitative and quantitative impacts of the geometric accuracy measure of vehicle extraction on the motion analysis. The establishment of the relationship between the shape preservation measure for extracted vehicles and the error measures for shape properties defined towards the motion artifacts is helpful to deepen the understanding of the performance of motion detection and estimation by specifying the error sources. Moreover, the normalized geometric accuracy for vehicle extraction is expected to serve as approximation for initial error values of shape deformation observations which are delivered to the error functions of the velocity estimator, therefore making the performance prediction of the motion analysis towards traffic monitoring in real-life urban scenarios more accurate and practical.

The unsolved problems discussed here can be treated in future research to enhance the proposed strategy for deriving traffic-related information from ALS data. Summarizing the points mentioned above objectively, it can be stated that the airborne LiDAR using the presented approach is fairly competent towards the urban traffic analysis compared to other sensors. Extracted vehicles can enable us to extract semantic information about vehicles (i.e. for vehicle counts and traffic density analysis) as well as the motion state and dynamical parameters of vehicles (i.e. for deriving traffic flow information) without changing data acquisition configurations commonly used for city mapping. The use of airborne LiDAR for traffic monitoring is consequently justified by such a great superiority to other data sources that not only spatial but also temporal information about the scanned scene is simultaneously encoded and resolvable.

References

Ahlberg J, Klasén L, Grönwall C, Ulvklo M, Jungert E (2003) Automatic target recognition on a multi-sensor platform. Proceedings of the Swedish Symposium on Image Analysis (SSAB'03), 93–96

Alexander C, Tansey K, Kaduk J, Holland D, Tate N J (2010) Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, In press

Axelsson P (1999) Processing of laser scanner data – algorithm and applications. ISPRS Journal of Photogrammetry & Remote Sensing, 54(2-3): 199-214

Bae K-H and Lichti DD (2008) A method for automated registration of unorganised point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 63(1):36-54

Baltsavias EP (1999) Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry & Remote Sensing, 54(2-3): 199-214

Barber CB, Dobkin DP, Huhdanpaa HT (1996) The quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software, 22(4): 469-483

Bartels M, Wei H, Mason DC (2006) DTM generation from LIDAR data using skewness balancing. Proceedings of 18th International Conference on Pattern Recognition, Hong Kong, Vol.1, 566-569

Blair JB, Rabine DL, Hofton MA (1999) The laser vegetation imaging sensor (LVIS): a medium-altitude, digitations-only, airborne laser altimeter for mapping vegetation and topography. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3): 115-122

Bogenberger K, Ernhofer O, Schütte C (1999) Effects of telematic applications for a high capacity ring road in munich. In: Proceedings of the 6th World Congress on Intelligent Transportation Systems, Toronto

Bowman AW and Azzalini A (1997) Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations, London, Oxford University Press

Chang L-D, Slatton KC, Anand V, Liu P-W, Lee H, Campbell MV (2010) Automatic forest canopy removal algorithm for underneath obscure target detection by airborne lidar point cloud data. Proc. SPIE 7664, 766424

Chapelle O, Vapnik V, Bousquet O, Mukhherjee S (2002) Choosing multiple parameters for support vector machine. Machine Learning, 46:131–159

Chen Q (2010) Assessment of terrain elevation derived from satellite laser altimetry over mountainous forest areas using airborne lidar data. ISPRS Journal of Photogrammetry and Remote Sensing, 65(1): 111-122

Chen Q, Baldocchi D D, Gong P, Kelly M (2006) Isolating individual trees in a savanna woodland using small footprint LIDAR data. Photogrammetric Engineering and Remote Sensing, 72(8): 923-932

Chevalier T, Steinvall O, Larsson H (2007) Performance of laser penetration through forest vegetation. Proc. SPIE 6550, 65500Q

Clodea S, Rottensteiner F, Kootsookosc P, Zelnikerd E (2007) Detection and vectorization of roads from lidar data. Photogrammetric Engineering and Remote Sensing, 73(5): 517-535

Comaniciu D (2003) An algorithm for data-driven bandwidth selection. IEEE Transaction on Pattern Analysis and Machine Intelligence, 25(2): 281–288

Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence, 24(5): 603 – 619

Doneus M, Briese C, Fera M, Janner M (2008) Archaeological prospection of forested areas using full-waveform airborne laser scanning. Journal of Archaeological Science, 35(4): 882-893

Dubuisson-Jolly M.-P, Lakshmanan S, Jain A (1996) Vehicle segmentation and classification using deformable templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(3): 293–308

Duistermaat J J, Kolk J A C (2000) Lie Groups, Berlin: Springer

Eikvil L, Aurdal L, Koren H (2009) Classification-based vehicle detection in high-resolution satellite images. ISPRS Journal of Photogrammetry and Remote Sensing, 64(1): 65-72

Ferryman J, Worrall A, Maybank S (1998) Learning enhanced 3-d models for vehicle tracking. The British Machine Vision Conference (BMVC)

Fletcher P T, Lu C, Joshi, S (2003) Statistics of shape via principal geodesic analysis on Lie groups. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Proceedings, Vol.1, 95-101

Fletcher P T; Lu C, Pizer S.M; Joshi S (2004) Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Transactions on Medical Imaging, 23(8): 995-1005

Fukunaga, K (1990) Introduction to Statistical Pattern Recognition, 2nd ed. New York, Academic Press

Grejner-Brzezinska D, Toth C, Paska E, (2004) Airborne remote sensing supporting traffic Flow estimates. Proceedings of 3rd International Symposium on Mobile Mapping Technology, on CD-ROM

Grenander U, Miller M. I. and Srivastava A (1998) Hilbert-Schmidt lower bounds for estimators on matrix Lie groups for ATR. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(8): 790–802

Grönwall C (2006) Ground object recognition using laser radar data: geometric fitting, performance analysis, and applications. Doctoral thesis, Linköping, Sweden, Linköping University

Grönwall C, Steinvall O, Gustafsson F, Chevalier T (2007) Influence of laser radar sensor parameters on range-measurement and shape-fitting uncertainties. Optical Engineering, 46(10): 106201

Gronwall C, Tolt G, Karlsson D, Esteki H, Rydell J, Armstrong E E, Woods J (2010) Threat detection and tracking using 3D FLASH ladar data. Proc. SPIE 7696, 76960N

Grün A, Li H, (1997) Linear feature extraction with LSB snakes. In: Grün et al., 287–298

Haag M, Nagel H-H (1999) Combination of edge element and optical flow estimates for 3D-model-based vehicle tracking in traffic sequences. International Journal of Computer Vision, 35(3): 295–319

Hinz S (2003) Integrating local and global features for vehicle detection in high resolution aerial imagery. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol.34(3W/8): 119–124

Hinz S (2004) Detection of vehicles and vehicle queues in high resolution aerial images. Photogrammetrie - Fernerkundung – Geoinformation (PFG), 2003(4): 201-213

Hinz S, Bamler R, Stilla U (2006) Editorial Theme Issue: Airborne und spaceborne traffic monitoring. ISPRS Journal of Photogrammetry and Remote Sensing, 61(3-4): 135-136

Hinz S, Baumgartner A (2001) Vehicle Detection in aerial images using generic features, grouping, and context. In: Pattern Recognition (DAGM 2001), Lecture Notes on Computer Science 2191, Berlin, Springer, 45–52

Hinz S, Lenhart D, Leitloff J (2008) Traffic extraction and characterization from optical remote sensing data. Photogrammetric Record 23(124): 424-440

Hinz S, Stilla U (2006) Car detection in aerial thermal images by local and global evidence accumulation. Pattern Recognition Letters 27(4): 308-315

Hofton MA, Minster JB, Blair JB (2000) Decomposition of laser altimeter waveforms. IEEE Transactions on Geoscience and Remote Sensing, 38(4): 1989–1996

Hug C, Ulrich A, Grimm A (2004) Litemapper-5600-A waveform-digitizing lidar Terrain and vegetation mapping system. International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, 36(8/W2): 24-29

Johansson B, Moe A (2005) Object Recognition in 3D laser radar data using plane triplets. Technical Reports: LiTH-ISY-R-2708, Linköping, Sweden, Linköping University

Jutzi B, Stilla U (2006) Range determination with waveform recording laser systems using a Wiener Filter. ISPRS Journal of Photogrammetry & Remote Sensing, 61(2): 95-107

Kaasalainen S, Hyyppa H, Kukko A, Litkey P, Ahokas E, Hyyppa J, Lehner H, Jaakkola A, Suomalainen J, Akujarvi A, Kaasalainen M, Pyysalo U (2009) Radiometric calibration of LIDAR intensity with commercially available reference Targets. IEEE Transactions on Geoscience and Remote Sensing, 47(2): 588-598

Kirchhof M, Stilla U (2006) Detection of moving objects in airborne thermal videos. ISPRS Journal of Photogrammetry and Remote Sensing, 61(3-4): 187-196

Klassen E, Srivastava A, Mio M, Joshi SH (2004) Analysis of planar shapes using geodesic paths on shape spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(3): 372-383.

Klausmann P, Fries S, Willersinn D, Stilla U, Thönnessen U (1999) Application-oriented assessment of computer vision algorithms. In: Jähne B, Haußecker H, Geißler P (eds) Handbook of computer vision and applications: Systems and applications, Vol. 3, 133-152

Koksal A.E, Shapiro J.H, Wells W.M, (1999) Model-based object recognition using laser radar range imagery. In: Firooz A. Sadjadi, (Ed), Automatic target recognition IX; Proceedings of the Conference, Orlando, FL, 3718, 256-266

Larsen SØ, Koren H, Solberg R (2009) Traffic Monitoring using very high resolution satellite imagery. Photogrammetric Engineering and Remote Sensing, 75(7): 859-869.

Laszlo, JM (1996) Computational Geometry and Computer Graphics in C++, New York: Prentice Hall

Leitloff J, Hinz S, Stilla U (2010) Vehicle extraction from very high resolution satellite images of city areas. IEEE Trans. on Geoscience and Remote Sensing, 48(7): 2795-2806

Lenhart D, Hinz S, Leitloff J, Stilla U (2008) Automatic traffic monitoring based on aerial image sequences. Pattern Recognition and Image Analysis, 18 (3): 400–405

Lin Y-C, Mills JP (2010) Factors influencing pulse width of small footprint full waveform airborne laser scanning data. Photogrammetric Engineering & Remote Sensing, 76(1): 49–59

Mahlisch M, Schweiger R, Ritter W, Dietmayer K, (2006) Sensorfusion using spatio-temporal aligned video and Lidar for improved vehicle detection. Intelligent Vehicles Symposium, IEEE, 424-429

Mallinis G, Koutsias N, Tsakiri-Strati M, Karteris M (2008) Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS Journal of Photogrammetry and Remote Sensing, 63(2): 237-250

Marquardt DW (1970) Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. Technometrics, 12(3): 591-612

Melzer T (2007) Non-parametric segmentation of ALS point clouds using mean shift. Journal of Applied Geodesy, 1(3): 159 –170

Meyer F, Hinz S, Laika A, Weihing D and Bamler R (2006) Performance analysis of the TerraSAR-X traffic monitoring concept. ISPRS Journal of Photogrammetry and Remote Sensing, 61(3/4): 225–242

Michaelsen E, Stilla U (2000) Assessing the computational effort for structural 3D vehicle recognition. In: Ferri FJ, Inesta JM, Amin A, Pudil P (eds) Advances in Pattern Recognition: Joint IAPR International Workshops SSPR 2000 and SPR 2000, LNCS 1876. Berlin: Springer, 357-366

Michaelsen E, Stilla U (2002) Probabilistic decisions in production nets: An example from vehicle recognition. In: Caelli T, Amin A, Duin RPW, Kamel M, Ridder Dde (eds) Structural, syntactic, and statistical pattern recognition: Joint IAPR International Workshops SSPR 2000 and SPR 2000, LNCS 2396, Berlin:Springer, 225-233

Michaelsen E, Wankmüller U, Stilla U (1998) Wissenserwerb für Produktionsnetze zur 3D-Erkennung von Fahrzeugen. In: Levi P, Ahlers RJ, May F, Schanz M (eds), Mustererkennung 1998, Berlin: Springer, 507-514

Nieves RD, Reynolds WD (2010) Three-dimensional transformation for automatic target recognition using lidar data. Proc. SPIE 7684, 76840Y

Olson C, Huttenlocher D, Doria D (1996) Recognition by matching with edge location and orientation. In: Image Understanding Workshop '96, San Francisco, CA, Morgan Kaufmann Publishers

Oude Elberink S, Vosselman G (2009) 3D information extraction from laser point clouds covering complex road junctions. Photogrammetric Record, 24(125): 23-36

Papoulis A (1984) Probability, random variables, and stochastic processes. Tokyo: McGraw-Hill

Persson Å, Söderman U, Töpel J, Ahlberg S (2005) Visualization and analysis of full-waveform airborne laser scanner data. In: Vosselman G, Brenner C (Eds) Laser scanning 2005. International Archives of Photogrammetry and Remote Sensing, 36 (3/W19): 109-114

Reitberger J, Schnorr Cl, Krzystek P, Stilla U (2009) 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6): 561-574

Rossmann W (2002) Lie Groups: An introduction through linear groups. London, Oxford University Press

Rutzinger M, Höfle B, Hollaus M, Pfeifer N (2008) Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. Sensors, 8(8): 4505 – 4528

Sampath A, Shan J (2007) Building boundary tracing and regularization from airborne lidar point clouds. Photogrammetric Engineering and Remote Sensing, 73(7): 805–812

Schenk T (1999) Photogrammetry and laser altimetry. In: Beata M. Csatho. (Ed) Joint Workshop of ISPRS III/5 and III/2 on Mapping Surface Structure and Topography by Air-borne and Space-borne Lasers, La Jolla, San Diego, CA

Shan J, Toth C.K (2008) Topographic Laser Ranging and Scanning: Principles and Processing. CRC Press

Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8): 888-905

Sithole G, Vosselman G (2003) Automatic structure detection in a point cloud of an urban landscape. Proceedings of 2nd Joint Workshop on Urban Remote Sensing, 67–71

Sithole G, Vosselman G (2005) Filtering of airborne laser scanner data based on segmented point clouds, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W19): 66-71

Skaloud J and Lichti DD (2006) Rigorous approach to bore-sight self calibration in airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 61(1): 47-59

Söderman U, Persson Å, Töpel J, Ahlberg S (2005) On analysis and visualization of full-waveform airborne laser scanner data. Laser Radar Technology and Applications X. In: Kamerman W (Ed), SPIE Proceedings, 5791: 184-192

Sohn G, Dowman I (2007) Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing, 62(1): 43-63

Soille P (2003) Morphological Image Analysis. 2nd Ed. Berlin: Springer

Soille P, Pesaresi M (2002) Advances in mathematical morphology applied to geoscience and remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 40(9): 2042-2055

Srivastava A, Klassen E (2002) Monte-Carlo extrinsic estimators of manifold-valued parameters. IEEE Trans. Signal Processing, 50 (2): 299–308

Steinvall O, Klasen C, Grönvall U, Söderman S, Ahlberg A, Person M, Elmqvist H, Larsson D, Letalick P, Andersson T, Carlsson, Henriksson M (2004) 3D laser sensing at FOI--overview and a system perspective. In: Gary W Kamerman (Ed), Laser Radar Technology and Applications IX Proc. SPIE 5412, 294-309

Steinvall O, Larsson H, Gustavsson F, Chevalier T, Persson Å, Klasén L (2004) Characterizing targets and backgrounds for 3D laser radars. Military Remote Sensing. In: Kamerman W, Willetts DV (eds) SPIE Proc, 5613: 51-66

Stilla U, Michaelsen E (2002) Estimating vehicle activity using thermal image sequences and maps. Symposium on geospatial theory, processing and applications. International Archives of Photogrammetry and Remote Sensing. 34(4)

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

Stilla U, Rottensteiner F, Hinz S (2005)(eds) CMRT05 Object Extraction for 3D City Models, Road Databases, and Traffic Monitoring - Concepts, Algorithms, and Evaluation. International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, Vol 36(3/W24)

Stilla U, Rottensteiner F, Paparoditis N (2009)(eds) CMRT09 Object Extraction for 3D City Models, Road Databases, and Traffic Monitoring - Concepts, Algorithms, and Evaluation. International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, Vol 38(3/W4)

Suchandt S, Runge H, Breit H, Steinbrecher U, Kotenkov A, Balss U (2010) Automatic extraction of traffic flows using TerraSAR-X along-track interferometry. IEEE Transactions on Geoscience and Remote Sensing, 48(2): 807-819

Sullivan G, Worrall A, Ferryman J (1995) Visual object recognition using deformable models of vehicles. IEEE Workshop on Context-based Vision, 75–86

Szarvas M, Sakai U, Ogata J (2006) Real-time pedestrian detection using LIDAR and convolutional neural networks. Intelligent Vehicles Symposium, IEEE, 213-218

Tan T, Sullivan G, Baker K (1998) Model-based localisation and recognition of road vehicles. International Journal of Computer Vision, 27 (1): 5–25

Tao WB, Hai H, Zhang YM (2007) Color image segmentation based on mean shift and normalized cuts. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 37(5): 1382-1389

Toth C, Grejner-Brzezinska D (2005) Traffic flow estimation from airborne imaging sensors: a performance analysis. Proceedings of Workshop on High Resolution Earth Imaging and Geospatial Information, Hanover, Germany, on CD ROM

Toth C, Grejner-Brzezinska D (2006) Extracting dynamic spatial data from airborne imaging sensors to support traffic flow estimation. ISPRS Journal of Photogrammetry and Remote Sensing, 61(3-4): 137-148

Turin GL (1960) An introduction to matched filters. IEEE Trans. Info. Theory, 6(3): 311-329

Vosselman G (2009) Advanced point cloud processing. Photogrammetric Week 2009, Fritsch D, (Ed.), Wichmann Verlag, Heidelberg, ISBN 978-3-87907-483-9, 137-146

Vosselman G, Maas H-G (2010) Airborne and terrestrial laser scanning. London: CRC Press

Wagner W, Ullrich A, Ducic V, Melzer T, Studnicka N (2006) Gaussian decomposition and calibration of a novel small- footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing, 60 (2):100-112

Wehr A, Lohr U (1999) Airborne laser scanning: an introduction and overview. ISPRS Journal of Photogrammetry & Remote Sensing, 54(2-3): 199-214

Yao W, Hinz S, Stilla U (2008) Traffic monitoring from airborne full-waveform LIDAR – Feasibility, simulation and analysis. XXI Congress, Proceedings. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 37(B3B): 593-598

Yao W, Hinz S, Stilla U (2009) Automatic estimation of vehicle activity from airborne thermal infrared video of urban areas by trajectory classification. Photogrammetrie - Fernerkundung – Geoinformation (PFG), 2009(5): 393-406

Yao W, Hinz S, Stilla U (2010) Automatic vehicle extraction from airborne LiDAR data of urban areas using morphological reconstruction. Pattern Recognition Letters, 31(10): 1100-1108

Yao W, Stilla U (2010) Mutual enhancement of weak laser pulses for point cloud enrichment based on full-waveform analysis. IEEE Transactions on Geoscience and Remote Sensing, 48(9): 3571-3579

Yarlagadda P, Ozcanli O, Mundy J (2008) Lie group distance based generic 3-d vehicle classification. ICPR 2008. 19th International Conference on Pattern Recognition, 1-4

Zhang C, Baltsavias E, Grün A (2001) Updating of cartographic road databases by image analysis. In: E P Baltsavias, A Gruen, L V Gool (Eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), 243–253

Zhang L, Huang X, Huang B, Li P (2006) A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. IEEE Transaction on Geoscience. Remote Sensing, 44(10): 2950–2961

Zhao H, Shibasaki R (2005) A novel system for tracking pedestrians using multiple single-row laser range scanners. IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans, 35(2): 283-291

Zhao T, Nevatia R (2001) Car detection in low resolution aerial image. The Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings., vol.1, 710-717

Zheng Q, Der S Z, Mahmoud H I (2001) Model-based target recognition in pulsed lidar imagery. IEEE Transactions on Image Processing, 10(4):565-572

Curriculum Vitae

Persönlichen Daten

Name	Wei Yao
Geburtsdatum	10.1981
Geburtsort	Nanjing, China
Staatsangehörigkeit	Chinesisch

Bildung

05/2004 – 05/2007	Geodäsie und Geoinformation (Dipl.-Ing) Technische Universität München	München
09/1999 – 06/2003	Geodäsie und Geoinformation (Bachelor. Eng) Wuhan Universität	Wuhan, China
09/1996– 06/1999	No.45 Gymnasium der Tianjin Stadt Abitur	Tianjin, China
09/1993 – 06/1996	No.7 Mittelschule der Nanjing Stadt	Nanjing, China

Beruf

Seit 01/2009	Fachgebiet Photogrammetrie und Fernerkundung Technische Universität München *Stipendiat des Bayerischen Eliteförderungsprogramms*	München
05/2007-12/2008	Lehrstuhl für Methodik der Fernerkundung Technische Universität München *Wissenschaftlicher Mitarbeiter*	München

Acknowledgements