MOUNTAIN-PLAINS CONSORTIUM

MPC 21-427 | M.A. Rahman and M.M. Mekker

Uses and Challenges of Collecting LiDAR Data from a Growing Autonomous Vehicle Fleet: Implications for Infrastructure Planning and Inspection Practices

MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Technical Report Documentation Page

1. Report No. MPC-577	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Uses and Challenges of Collecting LiDAR Data from a Growing Autonomous Vehicle Fleet: Implications for Infrastructure Planning and Inspection Practices		5. Report Date March 2021
		6. Performing Organization Code
7. Author(s) M. Mekker, M.A. Rahman		8. Performing Organization Report No. MPC 21-427
9. Performing Organization Name and Address Utah State University, Logan, UT		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address Mountain-Plains Consortium, North Dakota State University, PO Box 6050, Fargo, ND 58108		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code
15. Supplementary Notes Supported by a grant from the US DOT, University Transportation Centers Program		

16. Abstract

Autonomous vehicles (AVs) that utilize LiDAR (Light Detection and Ranging) and other sensing technologies are becoming more prevalent in the transportation industry. Concurrently, transportation agencies are increasingly challenged with asset management, traffic operations, and safety assessments. The affordability of LiDAR technology continues to increase. Given this, there will be substantial challenges and opportunities for the utilization of big data resulting from the growth of AVs with LiDAR. A proper understanding of the data size generated from this technology will help agencies in making decisions regarding storage, management, and transmission of the data.

This study explored the point cloud data size generated from a 16-beam LiDAR sensor unit at different operating modes and driving speeds with a mobile platform in the highway environment. Two types of roadway (collector and arterial) with three different sections (travel lanes, right of way, and unfiltered/total) were considered in this study for data collection. Conservative projections show an estimated total data size ranging from 1 to 3.5 gigabytes for 10,000 vehicle-miles traveled using the selected LiDAR unit. AVs will likely use multiple and/or higher resolution units. This study also developed simple models for estimating data size based on LiDAR operating mode and collection period duration.

17. Key Word data collection, data files, intelligent vehicles, laser radar, vehicle fleets		18. Distribution Statement Public distribution	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 58	22. Price n/a

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Uses and Challenges of Collecting LiDAR Data from a Growing Autonomous Vehicle Fleet: Implications for Infrastructure Planning and Inspection Practices

Md Ashikur Rahman
Michelle M. Mekker, Ph.D.

Department of Civil and Environmental Engineering
Utah State University
Logan, Utah

March 2021

Acknowledgements

First and foremost, I am thankful to my supervisor Dr. Michelle Mekker, Assistant Professor, Civil and Environmental Engineering Department at the Utah State University for her patience, guidance, and continuous motivation. Without her kind support and direction, it would have been a difficult journey for me to conduct this research. I am also grateful to Dr. Scott Budge for his valuable time and guidance about the technical aspect of my thesis. I would like to express my gratitude to Dr. Alfonso Torres for his cordial comments and help for the hard questions related to data processing.

This work was supported by the Mountain-Plains Consortium, a University Transportation Center funded by the USDOT.

ABSTRACT

Autonomous vehicles (AVs) that utilize LiDAR (Light Detection and Ranging) and other sensing technologies are becoming more prevalent in the transportation industry. Concurrently, transportation agencies are increasingly challenged with asset management, traffic operations, and safety assessments. The affordability of LiDAR technology continues to increase. Given this, there will be substantial challenges and opportunities for the utilization of big data resulting from the growth of AVs with LiDAR. A proper understanding of the data size generated from this technology will help agencies in making decisions regarding storage, management, and transmission of the data.

This study explored the point cloud data size generated from a 16-beam LiDAR sensor unit at different operating modes and driving speeds with a mobile platform in the highway environment. Two types of roadway (collector and arterial) with three different sections (travel lanes, right of way, and unfiltered/total) were considered in this study for data collection. Conservative projections show an estimated total data size ranging from 1 to 3.5 gigabytes for 10,000 vehicle-miles traveled using the selected LiDAR unit. AVs will likely use multiple and/or higher resolution units. This study also developed simple models for estimating data size based on LiDAR operating mode and collection period duration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND NOTATIONS

ADAS	Advanced Driver Assistance Systems
ALS	Airborne Laser Scanning
ASKF	Augmented State Kalman Filter
AVs	Autonomous Vehicles
CAVs	Connected Autonomous Vehicles
DATMO	Detection and Tracking of Moving Objects
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DDA	Discrete Discriminant Analysis
DEM	Digital Elevation Model
DPM	Deformable Part Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
ENU	East, North, and Up
FPFH	Fast Point Feature Histogram
GIS	Geographic Information System
GPS	Global Positioning System
GRF	Geographic Reference File
GMM	Gaussian Mixture Model
GNSS	Global Navigation Satellite System
GVF	Gradient Vector Flow
HADMs	High Accuracy Driving Maps
IDW	Inverse Distance Weighted
IMU	Inertial Measurement Unit
IMM	Interacting Multiple Model
INS	Inertial Navigation System
ITV	Iterative Tensor Volting
K-NN	K-Nearest Neighbor
LDA	Linear Discrimination Analysis
LiDAR	Light Detection and Ranging
MLS	Mobile Laser Scanning
MMS	Mobile Mapping System
MSTV	Multiscale Tensor Volting
NURBS	Non-Uniform Rational B-Spline
PCA	Principal Component Analysis
PPHT	Progressive Probabilistic Hough Transform
PMS	Pavement Management System
RANSAC	Random Sample Consensus
RAG	Region Adjacency Graph
ROI	Region of Interest
SVM	Support Vector Machine
TIN	Triangulated Irregular Network
TSR	Traffic Sign Recognition
UTM	Universal Transverse Mercator
VHM	Vehicle Height Model
WGS	World Geographic System
WNDH	Weighted Neighboring Difference Histogram

Executive Summary

The uses of geospatial technologies for a wide variety of transportation applications have expanded rapidly in recent decades. As Light Detection and Ranging (LiDAR) technology has matured, it has brought groundbreaking change in the transportation industry for infrastructure management, safety assessment, and other uses with greater precision, resource efficiency, and flexibility. The potential applications of LiDAR are expected to expand with the continued advancement of sophisticated sensor technologies and development of autonomous vehicles. It is important to consider and understand the big data generated from this technology in the context of a growing autonomous vehicle fleet for effective field use and data management practices.

This study explored the point cloud data size generated from a 16-beam LiDAR sensor unit in ASPRS recommended (. las) file format at different operating modes and driving speeds with a mobile platform in the highway environment. Two types of roadway (collector and arterial) with three different sections (travel lanes, right of way, and total) were considered for data collection in this study. In total, five different regression models were developed to project the data size generated from the sensor. Conservative projections show an estimated total data size ranging from approximately 1000 megabytes to approximately 3500 megabytes for 10,000 vehicle-miles traveled using the selected LiDAR unit. Autonomous vehicles will probably use multiple and/or higher resolution units. For transportation agencies to make appropriate decisions regarding future data management, it is imperative to provide proactive estimations of data size.

1. INTRODUCTION

The purposes of this research were to understand the large-scale point cloud data size generated from a mobile LiDAR (Light Detection and Ranging) platform and to outline associated challenges and previous applications of LiDAR data in the transportation industry. The outcome of this research will help transportation agencies make decisions regarding data storage and management and understand the utilization of the big data generated from LiDAR in the context of a growing autonomous vehicle fleet.

1.1 General Background

The advent of autonomous vehicles (AVs) and connected-autonomous vehicles (CAVs) have brought about significant changes in the transportation system in recent years. The technology is expected to have an even greater impact as it develops further in the coming years. Autonomous vehicle technology has the potential to reduce many negative aspects of personal automobile use and is expected to pave the way toward certain benefits. According to the National Highway Traffic Safety Administration (2016), 94% of the car crashes are attributed to human error, such as driving too fast, distraction, or poor judgement. It is widely accepted that AVs have the potential for tremendous safety benefits since they can lessen or eliminate the impact of human error. With the widespread deployment of automated driving systems, an approximate annual social benefit of $800 billion dollars is expected for congestion mitigation, reduced road causalities, minimized energy consumption, and increased productivity caused by the reallocation of driving time (Montgomery, Mudge et al., 2018).

The trend toward vehicle automation is escalating daily. At present, the majority of AV manufacturers are considering LiDAR as the fundamental sensing component of AVs. Mounted on an AV, LiDAR provides the means to gain a 3D representation of the surrounding environment with a 360-degree view. LiDAR has also become a mainstream technology for highway data collection. It can collect large-scale, dense data in the form of 3D point clouds for various transportation applications with greater flexibility, resource efficiency, and accuracy. Unlike traditional sensors, such as cameras, it can collect data in the dark with no impact on quality.

The use of LiDAR technology has been growing in the transportation industry in recent decades. This emerging technology is most commonly used via three different platforms: mobile, airborne, and stationary. LiDAR measures the distance of the target objects from the reflection time of a light pulse. The mobile and airborne LiDAR units usually integrate an inertial measurement unit (IMU) and a global navigation satellite system (GNSS). The GNSS provides the locational information (x, y, and z) of the point clouds in a global frame while the IMU informs the orientation (roll, pitch, and yaw) of the platform. One of the major advantages of LiDAR is that, unlike traditional manual survey methods, it can produce a high-resolution 3D point cloud of the surveyed area within a fraction of the time, causing minimal disruption to the traffic operations.

The direct ownership and application of LiDAR sensors has increased significantly among transportation agencies and organizations at the local, state, and national level. Although the demand for LiDAR is growing, it poses certain challenges and drawbacks in terms of applicability. One of the major challenges regarding the use of LiDAR technology is use and management of the big data generated from this sensor technology. It sends thousands of laser pulses per second. Having rotating mirrors, these 3D laser scanners can take millions of measurements over a scene in just a few seconds or minutes by Kemeny and Turner (2008). Tupas et al. (2016) indicated that “In LiDAR acquisition and processing operations, terabytes to petabytes of disk storage used is the norm.” An appropriate method to estimate the large-scale data size generated from the LiDAR technology is crucial for transportation agencies to understand the considerable data size and also to consider certain challenges and solutions.

1.2 Problem Statement and Scope

Remote sensing technologies for highway asset inventory are undergoing rapid development. The possibility of acquiring 3D information over a large area, such as a roadway, with survey grade accuracy and decreased survey time is creating new opportunities for innovation and unique management of resources. Advanced Driver Assistance Systems (ADAS) based on LiDAR sensors is one of the most innovative and game-changing technologies for the operation of AVs. The use of mobile LiDAR for rapid data collection to make high resolution informative maps has made the decision-making process for transportation agencies quicker and easier. Extensive research is ongoing in the transportation industry to understand the application of LiDAR technology for various purposes. Since the current use of LiDAR technology is increasing rapidly, there is a need to evaluate potential challenges and opportunities associated with the application of this sensing technology.

One of the most considerable disadvantages of the current LiDAR data collection process is that a significant amount of time and computing power is required to prepare raw data for possible analysis. Although LiDAR has become more affordable for average users, the effective process and extraction of useful information from the raw point clouds still poses a technical challenge (Chen, 2007). Cao et al. (2015) stated that due to the volume of the data, the computational and technical requirements to process, manage, and store the data can be significant. The processing and analysis of LiDAR data involves different software platforms, which may reduce data quality, be proprietary, and require varying levels of training to utilize. As the prevalence of this technology in the context of an autonomous vehicle fleet grows, transportation agencies will soon be faced with the challenges of technical complexities and high-volume data size. The rapid increase of LiDAR applications has created the need for storage and online distribution of collected 3D cloud data (Kulawiak et al., 2019).

Few studies have been conducted to understand the data size generated from LiDAR sensors at a large scale. This study was designed to demonstrate the potential data sizes of large-scale LiDAR data collection, possible uses of the collected data, and outline the related challenges and solutions for transportation agencies interested in collecting such data from a growing fleet of AVs. Different scenarios, such as the travel lanes, right of way, and total data, were considered to understand the data size. Linear regression models were developed to project the large-scale data size under the selected different scenarios. The model used data collected from different roads (arterial and collector), at different vehicle speeds and with different operating modes of the LiDAR sensor. The LiDAR operating mode (horizontal resolution and refresh rate), and duration of the sensor run time (related to vehicle speed) were considered as the explanatory variables to forecast the data size generated from this sensor. This research will help the transportation agencies to understand the possible data size they can expect from autonomous vehicles using LiDAR and the associated challenges with collecting data. Moreover, the outcome of this research will help agencies that are considering investments in the collection of LiDAR data to understand the benefits and challenges.

1.3 Research Objective

The main goal of this research was to understand the potential uses and challenges of large-scale LiDAR data from a growing fleet of AVs by transportation agencies. An extensive literature review was conducted to understand the existing applications of LiDAR technology in the transportation industry. Then, LiDAR data was collected from a real-world highway environment at different vehicle speeds and LiDAR operating modes in order to understand the possible data size generated from the sensor.

The specific objectives of this project were to:

- Present a model developed to estimate the size of LiDAR data sets collected by an AV.
- Summarize previous studies of LiDAR in the transportation industry
- Outline the possible challenges and future research scope of using this type of sensor.

1.4 Justification of the Research

Transportation infrastructure plays a significant role in economic, health, societal, environmental, etc., aspects of life in the United States and across the globe. The high spatial resolution and mapping accuracy of LiDAR has made it popular for transportation planning and infrastructure asset inspection and maintenance. This sensor has also become popular in the AV industry for its affordability and reliability. AVs rely on their fundamental sensors to retrieve and process the information about their immediate surroundings, which informs their decision-making and actions. It is necessary for transportation agencies to understand the wide range of applications of this emerging technology and the corresponding opportunities. A method for estimating the large-scale LiDAR data size and related challenges will help transportation agencies make necessary decisions and policies. Results of this research will provide a baseline for future data management, data policies, and resource management.

2. LITERATURE REVIEW

This chapter covers an extensive literature review regarding applications of LiDAR technology in the transportation industry to highlight the potential of the 3D point cloud data. Most of transportation agencies that use LiDAR use it for inventory and asset management purposes. However, with the emergence of the autonomous vehicle, the potential of this sensor technology is extended beyond asset management. The review outlines the algorithms used to extract information from the LiDAR data and the relevant challenges.

2.1 Road Sign Detection

Road signs or traffic signs provide important information to road users. They are mainly erected at the side or above the road. These signs usually use shapes, colors, words, or different symbols to deliver certain message to users. As per state and federal regulations, the vast majority of road signs are covered with reflective material to reflect light in all directions to be more visible to users and readable during nighttime to increase traffic safety. A LiDAR sensor can be used to detect the road signs by extracting reflection intensity information about the reflective objects.

Chen, Kohlmeyer et al. (2009) used mobile LiDAR data to automatically detect traffic signs. After significant filtering of the data, a threshold value was used to extract point clusters with high point density and a random sample consensus (RANSAC) algorithm was applied to geometrically fit a plane to each point cluster. The authors declared an approximate success of 98% of this method to detect traffic signs. Yang, Fang et al. (2012) proposed a method by using mobile LiDAR for the automated extraction of road markings. In this method, inverse distance weighted (IDW) interpolation and discrete discriminant analysis (DDA) were used to develop a georeferenced feature image of the point clouds representing the road surface. Later, filtering was done based on the intensity and elevation to reduce false candidates, such as pedestrians and cars. Finally, the progressive probabilistic Hough transform (PPHT) operator was applied by using the semantic knowledge (shape, pattern) to extract road markings from the segmented points. Landa and Prochazka (2014) used intensity information from LiDAR data to extract and detect road signs. A threshold intensity value of the raw cloud data was considered to cluster the selected points based on the Euclidean distance between two points. The cluster was then filtered based on height, elevation, and point density. The authors reported that this method is only applicable for a single sign extraction with a 93% success rate while attributing the missing road signs to low density in point cloud data.

Wu, Wen et al. (2015) developed a method using the principal component analysis (PCA) and intensity filtering for traffic sign detection and visibility evaluation from mobile LiDAR data. The 3D point clouds were projected onto corresponding 2D images. The vector-based formula was applied by using the geo-referenced relation between the points and the image to extract spatial-related features of the traffic signs and image features. Finally, the extracted image and spatial related features were combined to evaluate visibility. It was reported that the average deviation between the calculated results and subjective evaluation were under 5%. Li, Shinohara et al. (2016) proposed an automatic method to detect and recognize road signs using images and 3D point clouds acquired by a mobile mapping system (MMS). An object-based image analysis approach was used to classify road signs from the high-resolution raster image acquired by a camera mounted on MMS, where the point cloud was projected for matching laser points with pixel. Then, false positive (over detection) candidates were filtered out. Later, a template matching method followed by shape normalization was used to recognize and cluster the road signs from the false negatives (false classification). Field tests of this method showed an estimated accuracy of

98.4% to detect and recognize road signs. The authors mentioned the efficacy of this method under the challenging conditions, such as discoloration, deformation, and partial occlusions.

Wen, Li et al. (2016) proposed a method using an algorithm based on identifying linear structure (traffic sign pole), terrain, and intensity filtering for detecting and recognizing traffic sign using LiDAR point clouds and imagery. After filtering the road surface and boundary, the traffic sign position and placement inspection were conducted by analyzing the geo-spatial relationship between the traffic sign and the road elements. An image-based traffic sign recognition (TSR) was implemented to obtain the traffic sign type. The authors reported the experimental results of Mobile Laser Scanning (MLS) point clouds and images showed a detection precision of 91.63% and 92.61%, respectively. The results also showed a precision of 96.32% for traffic sign type recognition. Soilán, Riveiro et al. (2016) proposed a method for the automated detection and recognition of vertical traffic signs from a combination of 3D point clouds and imagery data integrated by MMS techniques. After preprocessing and filtering, a density-based cluster method, density-based spatial clustering of applications with noise (DBSCAN), was applied to cluster high intensity points, which were further filtered to separate traffic signs from poles, facades, and walls. After syncing and projecting the 3D point clouds on a 2D color image generated by the vehicle camera, a hierarchical approach was applied to recognize the meaning of the traffic sign. According to the authors, the recall result of this proposed method showed an approximate success rate of 98% to detect traffic signs. This method was suitable for identifying the geometric and semantic properties of traffic signs in urban and highway environments. Gargoum and El-Basyouny (2017) used DBSCAN on filtered data to separate and cluster traffic signs from other high reflective objects, such as license plates. This approach was tested on three different highways and the results demonstrated a success rate of 93%–100%. However, the authors reported this algorithm worked better on highways without overhead signs.

2.2 Traffic Monitoring

Traffic monitoring is the study of traffic conditions on a road network. It is important for transportation planners and engineers to ensure sustainable movement of traffic in a road network. LiDAR can easily detect, classify, and track the movement of the vehicles. The produced scans of the LiDAR give the upper and side contours of the vehicles. The profiles of vehicles can be affected by weather conditions like rain and snow. LiDAR surveys help to accurately and cost effectively collect traffic data for ensuring traffic safety and operation.

Toth et al. (2004) used georeferenced airborne LiDAR data to compute traffic flow by extracting vehicle information along with road surface modeling. After extracting and smoothing the road boundary, a simple adaptive parallel height thresholding scheme was applied to extract and cluster vehicle candidate points. The candidate vehicle clusters were further filtered and classified by testing a variety of parameter combinations and, based on the average length of the vehicles, the velocity of the vehicles was approximated. Finally, the computed vehicle locations, categories, and estimated velocities were used to define average vehicle density and velocity as parameters to estimate traffic flow. Yao, et al. (2008) made a comparison among stationary, mobile, and airborne LiDAR data applications to derive traffic monitoring information with a focus on urban areas. In this study, an airborne laser scanning (ALS) simulator was used to address the motion artifact of the moving vehicle in the laser data. It was found the ALS data had a greater advantage in terms of extracting 3D objects. After masking out vegetation and human-made objects, the remaining points in the dataset were transformed to generate a vehicle height model (VHM) following normalized digital surface models (nDSM). Nashashibi and Bargeton (2008) presented a detection and tracking of moving objects (DATMO) approach for the detection, tracking, and classification of multiple vehicles using mobile LiDAR. In this approach, data segmentation was performed by using the Ramer algorithm following heuristic rules. A bicycle model was used following a

Kalman filter to estimate different parameters of the detected surrounding vehicles and track them. The authors reported the proposed method was validated over hundreds of miles. Patlins, et al. (2010) explored the use of LiDAR technology with trams as a distance measuring sensor from obstacles. A control algorithm was developed where the LiDAR provided signal as input to calculate tram braking path. The study recommended placing signal receptors on the right of tram's headlight. It was found that the reflection time of the LiDAR beam from an obstacle is negligible compared to the time of tram braking.

2.3 Pavement Condition Assessment

Pavement condition and performance plays a critical role in the successful operation of the highway network. Up-to-date data on pavement condition are collected by the local and state agencies as a part of the pavement management system (PMS) for decision making and to perform necessary maintenance. The cost of pavement surveying largely depends on the applied methods and survey frequency. LiDAR is preferred for easy collection of highly accurate 3D geospatial information regarding the pavement condition for ensuring flexible, cost-effective, continuous monitoring, and maintenance.

Hernández and Marcotegui (2009) proposed the use of a quasi-flat zone algorithm and a region adjacency graph (RAG) representation to automatically filter artifacts and detect pavement from mobile LiDAR data. Later, the pavement segmentation was carried out by projecting the range image onto the filtered point clouds. The final output of this method was a contour image of the ground foot print. Tsai and Li (2012) attempted to evaluate the feasibility of using mobile LiDAR to detect pavement cracks by employing a dynamic-optimization-based crack segmentation method under different lighting and poor intensity contrast conditions. Ouyang and Xu (2013) used 3D pavement images generated from a modified 3D camera and mobile LiDAR data to detect and obtain pavement crack measurements. The non-uniform rational B-spline (NURBS) surface approximation and Haar transform were applied to a pavement image to locate crack regions and orientations for performing fine crack detections. Finally, alligator cracking (interconnected longitudinal and transverse cracks) was identified by analyzing the crack densities from longitudinal and transverse crack histograms. The authors reported the difference in the total crack length between the multiple runs for both ways were smaller than 2% at a given speed.

Křemen et al. (2014) evaluated the use of static 3D terrestrial laser scanning point clouds (total station) to assess the pavement roughness by comparing it with test field checking points. A digital terrain model (DTM) was created from the laser scanning points and compared with the elevation of the test field checking points. The paper concluded that the accuracy of 3D laser scanning point clouds is sufficient for situations where standard deviation in elevation must be 5 mm or less. Guan et al. (2015) used mobile laser scanning data to propose an automated pavement crack detection method. This method used iterative tensor volting (ITV), inverse distance weighted (IDW) interpolation, and morphological thinning to extract pavement cracks from the noisy geographic reference file (GRF) images. This method was basically focused on extracting the type and location of pavement cracks, not the width. The authors reported that the proposed method is applicable for pavement cracks with low contrast, low signal-to-noise ratio, and bad continuity. Kumar and Angelats (2017) proposed an automated algorithm to detect road roughness from mobile LiDAR point clouds. This process involved extensive filtering, point thinning processing, and interpolation to generate an intensity raster surface. Morphological and multi-level ostu thresholding operations were applied to detect road roughness within candidate regions. Finally, the candidate regions were clustered based on spatial density and standard deviation of elevation to remove the outliers belonging to normal road surfaces or nearby vehicles for detecting roughness along the road surface.

2.4 Geometric Data Extraction and Assessment

The geometric design of roads mainly deals with the dimensions and layout of the physical elements of the roadway, following different standards and constraints. The geometric data of a highway is mainly composed of three parts: horizontal alignment, vertical alignment, and cross section. 3D point clouds can be used for the geometric assessment of the highway due to its high data accuracy and density.

Cremean and Murray (2006) developed a clothoid model to construct and recursively estimate the planar road geometry on extracted road features from single-axis LiDAR range measurements in off-highway environments. Tsai, et al. (2013) proposed a method for cross slope measurement using mobile LiDAR technology. After processing the raw point cloud data, a region of interest (ROI) process was performed on the data to extract the rectangular regions within a single lane between the pavement markings. The width of the ROI was defined by the distance between the pavement markings, and the length was determined by the interval of ROI. After performing linear regression on the extracted point clouds from each ROI, the slope of the regression was used to measure the cross-slope. Dawkins (2014) used an augmented state Kalman filter (ASKF) to estimate the off-road terrain profile as unknown inputs to the 7-degrees of freedom full suspension model. Later, mobile LiDAR was used to collect off-road terrain profiles and validate the model. The paper reported that the method is more effective for capturing low frequency content of the terrain profile. Ai and Tsai (2016) proposed a method for automated sidewalk and curb ramp assessment using 3D mobile LiDAR and image processing. The sidewalk points were extracted from the raw point clouds using the lateral and elevation offset. The author suggested use of a threshold of 7.3 m (24 ft) for lateral offset and 0.6 m (2 ft) for elevation offset for a good balance between reliability and efficiency. Cross section segmentation, k-nearest neighbor (k-NN), and a B-spline algorithm were introduced to smoothly connect different segmentation points to detect the sidewalks. A deformable part model (DPM) was applied to detect and extract all curb ramp candidates from the data using video log images. Later, an interactive tool was employed to further filter the data and detect the curb ramps. Gargoum, et al. (2018) recommended the use of LiDAR data to automatically assess the vertical clearance on the highway. After extensive filtering, the k-Nearest Neighbor (k-NN) search algorithm and DBSCAN clustering algorithm were used to divide point clouds into bridge and non-bridge groups using histogram analysis. This method demonstrated a satisfactory performance for network-level assessment of vertical clearance. Mekker, et al. (2018) explored the integration of connected vehicle data and LiDAR technology for identifying and diagnosing geometry-related capacity problems in recurring work zones considering safety and time efficiency. The connected vehicle data, reported as space mean speed every minute for segments of an average 0.88 miles in length, was used to identify recurring bottleneck locations. Then, by connecting the speed data with the geospatial LiDAR data collected in the selected work zone, it was found that the lane width and/or taper length of selected sites did not follow the recommended design, resulting in recurring traffic congestion.

2.5 Embankment Stability Monitoring

Embankment stability is an important element of highway maintenance for ensuring safety. It is considered under the design process from the beginning until implementation of the project. It is mainly influenced by the imposed load, loading pattern, and the condition of the slope. The detailed and accurate information acquired from a LiDAR sensor can be used to assess, characterize, and map the condition of an embankment.

Lato, et al. (2009) examined the integral use of multiple LiDAR sensors (static, mobile, terrestrial, and airborne) to monitor the geotechnical hazard for linear transport corridors. The authors concluded that the fusion of data from multiple LIDAR sensors can be used to detect small rock block release (sub 15 cm). The paper recommended using LiDAR data for geomechanical structural feature identification and kinematic analysis, rock fall path identification, and differential monitoring of rock movement or failure over time. However, it was found that the mobile terrestrial LiDAR can be more efficient compared to static terrestrial LiDAR in terms of coverage, rate of acquisition, dynamic collection, and integration with corridor operation. Đapo, et al. (2011) used a terrestrial laser 3D scanner to understand the load inefficiency in the rehabilitation project of a railway bridge, "Sava Jakuševac," in Croatia. After filtering the raw point cloud data, the required elements, including columns and load bearing parts of the structure, were taken from the scanned data. It was found that the irregularity of the track on the bridge was caused by an uneven subsidence of column S6.

2.6 Lane Marking and Road Edge Extraction

Lane marking and road edge detection are important for safe driving. The road edge is the boundary between the road surface and the non-road surface. The lane marking conveys messages to road users regarding the purpose of different parts of the road and where they can legally operate their vehicle. Road edge and lane detection are also necessary for the safe operation of autonomous vehicles. Conventional digital maps with road-level resolution are not sufficient for autonomous vehicles to understand their surroundings. To resolve this issue, LiDAR-based road edge and lane marking detection systems, mainly based on elevation and reflectivity are becoming popular among transportation agencies and vehicle manufacturers for its higher accuracy and detailed information.

Hu, et al. (2004) used high resolution optical imagery and airborne LiDAR data to extract a grid-structured urban road network. An iterative Hough transform algorithm and morphological operation were employed on the segmented point clouds to extract candidate road markings and parking areas. The segmented road marking and parking area candidates were verified from shape analysis and vehicle clue detection of the LiDAR data and high-resolution imagery. Finally, the topology analysis was used to form the road network from the verified roads and parking areas. Ogawa and Takagi (2006) used mobile LiDAR data to detect the lane. After filtering raw point clouds, a Kalman filter and auto correlation function were applied considering four lane parameters (width, offset, yaw angle, and curvature) to recognize the lane markings from the candidates. The paper mentioned this method was more suitable for roads with highly reflective lane markings. Jaakkola, et al. (2008) applied an image processing algorithm to intensity and height images to automatically filter road markings and curbstone points, modelling the road surface as a triangulated irregular network (TIN) from mobile mapping. The clustering of the road markings and curbstones was done by segmentation using thresholding and applying morphological operations on the raster image of intensity and elevation. Field tests of this method showed an overall accuracy of 80.6%, 92.3% and 79.7%, respectively, for classifying lines, crosswalks, and curbstones. Chen, et al. (2009) used mobile LiDAR to detect the 3D lane markings from point clouds. First, the road surface was detected by height filtering. Hough transform clustering and RANSAC fitting were then employed to extract the 3D lane markings from the candidate points. However, the authors did not report the success rate of this method to detect lane markings. Zhu and Mordohai (2009) used height and intensity information from LiDAR data to extract the road network. Region segmentation following a normalized cut algorithm (Ncut) and edge detection were performed to generate hypothesis-based road boundaries and interior features. A minimum cover algorithm was employed to find the salient regions that best suited the 2D map and automatically selected the road widths. Finally, the road regions were transformed back to 3D points and roads were extracted. Zhang (2010) proposed a method to identify road regions and road edges using LiDAR point clouds. After elevation filtering, the false alarm mitigation module was performed following a minimum threshold road width to further filter and validate potential road region candidates. Finally, the identified 3D road edge points were further projected and validated on a 2D ground plane to detect the curb. The author stated in the experimental study that the proposed algorithm detected most road points, road curb points, and road edge points correctly with a false alarm rate of 0.83% and a missing rate of 0.55% per scan. Kumar, et al. (2010) used a gradient vector flow (GVF) snake model for the automated extraction of road edges from mobile LiDAR and vehicle trajectory data. The snake model derived its energy from the LiDAR point clouds before the snake contour was developed using the trajectory data. The computation process of these energy terms used surface slope, LiDAR intensity, and the qualitative measurement of the elasticity and rigidity terms. The authors reported that the tested result of this method was satisfactory in terms of detecting the difference between the slope values of curbs and planar road surface. Wang, et al. (2012) proposed an automatic algorithm to extract the road surface and boundary based on trajectory information from mobile LiDAR data. The road boundary and surface were determined by applying a hypothesis testing method based on the local altitude variance of a particular segment. A mean height threshold filter value was used to exclude altitude abnormal points (cars, pedestrians) from the road surface and boundary. Kang, et al. (2012) proposed a method using the probabilistic interacting multiple model (IMM) to detect road boundaries from LiDAR point cloud measurements. This approach involved multiple Kalman filters to track the intersection points between a LiDAR scanned surface and a curb. Experimental results were promising for curb detection and the authors reported the performance of this method increased when the GPS measurement is unreliable. Hata and Wolf (2014) used mobile LiDAR data to develop a road marking detector system following the Otsu thresholding method. After detecting the road curbs with a given threshold following a ring compression method, a deterministic calibration was performed on the LiDAR intensity information to ensure similar intensity values for asphalt and road markings within the road. The Otsu thresholding method was applied on this calibrated intensity information to develop a bimodal intensity histogram with one mode grouping asphalt intensities and another grouping road marking intensity. The authors reported that by adding road markings onto the curb maps the error level of lateral localization was reduced by 53.56% with an average lateral error of 0.3119. The study stated that the proposed method was capable of detecting different types of road markings, including crosswalks, continuous lines, and dashed lines. Liu and Lim (2014) developed a method for extracting roads from airborne LiDAR point clouds and associative vector data. After separating and refining the candidate road points, a fitting was used to transform road points into polylines. The paper stated that the experimental results of this method demonstrated satisfactory performance for road extraction from airborne LiDAR data for structured and unstructured lanes. Guan, et al. (2015) proposed a method for rapid and accurate road marking inventory using mobile LiDAR data. After filtering the road surface, a modified inverse distance weight (IDW) interpolation method was used to interpolate the road surface into a 2-D georeferenced feature (GRF) images. After dividing the GRF image into sub-images, a weighted neighboring difference histogram (WNDH)-based dynamic thresholding and multiscale tensor voting (MSTV) were used to segment and extract road markings from the noisy, corrupted images. The author concluded that the proposed method to detect and extract road markings was more suitable for a subtropical, urban environment.

Liu and Lim (2016) proposed a method to extract roads from fused airborne LiDAR data and aerial imagery. An edge-clustering algorithm and k-nearest-neighbors (KNN) clustering were applied to filter candidate road points. After filtering the cluster, local IDW interpolation and curve fitting was applied to the road segments in order to obtain road centerlines. The authors reported that the quantified completeness and correctness of five test results of this method to extract roads were 82.6% and 87.4%, 89.2% and 91.2%, 80.7% and 87.6%, 84.2% and 90.4%, and 79.5% and 89.5%, respectively. This paper stated that the application of this method to extract roads was hindered for dense vegetation areas, curved roads, and traffic islands. Li, et al. (2016) proposed an automated road centerline extraction method using LiDAR data. After filtering ground points, the adaptive mean shift cluster algorithm was applied to cluster the road center points into linear points. The principal component analysis, least squares multiple lines fitting, and a hierarchical grouping method were employed to connect primitives of road centerlines into continuous road lines to develop a road network. The authors reported that the experimental result of the proposed method demonstrated the same level of performance in less time compared to other methods like the mean shift algorithm, Tensor voting, and Hough transform. Ghallabi, et al. (2018) proposed a lane marking detection method for localization within an HD map using a multilayer LiDAR system. After filtering the road points and projecting them on a 2D intensity image, a Hough transform algorithm was employed on the resulting image to detect lane markings from the high reflectivity road points. After one more filter, the proposed method was validated by localizing the vehicle in a map with lane-level information and absolute accuracy of 2 cm. The authors reported this method was capable of providing a lane-level localization with 22-cm cross-track accuracy.

2.7 Roadside Objects Detection

The detection of roadside objects is important for operation of autonomous vehicle systems to ensure obstacle avoidance and for safety audit purposes. The improvement of the traffic safety and efficiency requires detailed level information about objects located with the road and along the roadside. LiDAR technology can be used to address this issue as it can provide real-time, high-resolution 3D point cloud data containing information about the location and shape of the objects.

Shamayleh and Khattak (2003) conducted an experiment to examine the accuracy of LiDAR data for the collection of roadway inventory information by comparing with the ground truth information. The LiDAR data was transformed into a triangulated irregular network (TIN) by incorporating the height information and overlaid with the corresponding aerial image. Later, related analyses were performed in the ArcGIS environment to obtain the values of selected elements of roadway inventory. To validate the accuracy of this data, a set of ground points were selected for field survey based on an empirical method. After conducting a comparison with the ground data, very minor differences were found between the LiDAR and field data. Lam, et al. (2010) developed a method to extract 3D road data from mobile, ground-based LiDAR point clouds. This approach modeled the road as a dynamic system of connected planes by plane fitting local 3D point data and using Kalman filtering to globally extract these planes. Afterwards, bounding boxes were projected from each extracted road patch to localize objects, such as lamp posts, power line posts, and road signs. These localized urban scenes were filtered from vegetation by using RANSAC followed by least squares fitting.

Kwan & Ransberger (2010) explored the use and analysis of LiDAR data to detect obstructions in the transportation network during emergency response in New Orleans, LA, before and after Hurricane Katrina. First, DEM was developed for all the pre- and post-storm LiDAR data along the network links. The change in elevation between the pre- and post-storm LiDAR data were used to identify obstructions. Lin and Hyyppä (2010) developed a method for the automated detection of culverts in mobile laser, scanning point clouds. A digital terrain model (DTM) was developed from point clouds to remove the shadow influence. Then, possible locations of the culverts were coarsely located after searching the blank zones. Later, topological plot segmentation and an intensity filter were employed to detect real culverts. The geometric parameters of the culverts were calculated from the selected schematic components. The authors reported that the proposed method was tested to detect two pedestrian culverts, and the results showed an estimated error of lengths and widths compared to the real ones of less than 9% and 16%, respectively. This paper recommended the consideration of water reflectance for future culvert detection. Pu, et al. (2011) used mobile laser scanning (MLS) point cloud data to develop a knowledge-based feature recognition theory. The on-ground points were extracted via filtering and a hybrid model was applied to recognize the features. Later, based on certain geometric characteristics of the features and topological relationships with other features, on-ground objects were clustered into more detailed groups like traffic signs, trees, poles, building walls, and barriers. The authors reported an overall accuracy of 86% to recognize poles using this method.

Jeong and Lee (2016) proposed a method to classify different road objects from mobile LiDAR point clouds. This method involved the development of a training data set from the raw point clouds. Later, the geometric features of the data were extracted using a Fast Point Feature Histogram (FPFH). Finally, geometric features were used to develop a classification model by employing support vector machines (SVM). The authors reported an overall accuracy of 98% to classify the objects and a prediction accuracy of 91%. Yang, et al. (2016) proposed a method to extract multiple types of road features from mobile LiDAR data to produce High Accuracy Driving Maps (HADMs) in a highway environment. The Hernández and Marcotegui method were applied to group the points into ground and non-ground points. After detecting and extracting the road surface from the ground points, the non-ground points were classified into individual object candidates by pointwise classification, multi-rule segmentation, and adjacent segments merging. Finally, the support vector machine (SVM) approach was implemented with one versus one classification to classify the features into different classes, such as trees, street lamps, traffic signs, buildings, etc. Quantitative evaluations of this method demonstrated good performance in extracting road features with an average precision and recall of 90.6% and 91.2%, respectively. However, the results also showed that the proposed method's precision and recall vary largely with point densities, positional accuracies, and signal-to-noise ratios of the point cloud data.

Ordóñez, et al. (2017) developed a method to detect and classify pole-like objects from mobile laser scanning point clouds by employing a heuristic segmentation algorithm, linear discriminant analysis (LDA), and SVM. Experimental tests of this method showed an overall success rate of 90% to detect different types of pole-like objects in an urban environment. Zheng, et al. (2017) proposed a method for automatic recognition and extraction of street lighting poles from mobile LiDAR data. After filtering out the ground points, the remaining points were grouped into ground and non-ground objects. Graph cut-based segmentation and a Gaussian-mixture model (GMM) based method were used to model the clusters of the non-ground objects and the street lighting poles were recognized through matching the clusters with a database of sample street lighting pole models. Experimental results showed that this method achieved an overall performance of 90% in terms of true positive rate.

Hůlková, et al. (2018) employed a raster image processing technique to make rough classifications of the highway environment using mobile laser scanning. The proposed method required the prior knowledge of the terrain of the study area. In this approach, the raw point clouds were divided into voxels and matrix using point information, such as class, height, and intensity. These were then projected onto a horizontal plane to create a raster image. An image processing technique that included segmentation, filtering, and object-oriented classification was applied repeatedly on the raster image to classify it according to 10 pre-selected classes of highway environment. The outcome of this method was compared with a semi-automatic classification method using the Terra Scan software, where the classification was inspected by a human expert and later improved manually if needed. This paper reported an overall classification accuracy of 94.5% of the presented method. However, the authors mentioned that the proposed method misclassified tree trunks to the poles class and long vehicles to the crash barriers class.

2.8 Sight Distance Assessment

Sight distance is one of the fundamental components of highway design. It is the length of road surface that driver can see at a point along a roadway. Provided sight distance is important to ensure adequate stopping distance at a specified driving speed. If the available sight distance is less than what is required for a driver action, then the risk of a crash is significantly increased. There are four types of sight distance considered in roadway design: intersection sight distance, stopping sight distance, passing sight distance, and decision sight distance. The traditional methods applied for sight distance assessment are often time consuming, cost- and labor-intensive, and require interference with traffic operations to ensure worker safety. Nowadays, many transportation agencies use LiDAR technology for sight distance analysis considering its efficiency over conventional methods.

De Santos Berbel, et al. (2008) developed ArcGIS add-in software to calculate the available sight distance on roads using LiDAR data. As input, DSM was preferred over DEM and the trajectories followed by a vehicle along the road in each direction were considered. The authors recommended combining the sight distance information with traffic flow data, speed maps, crash data, and alignment consistency information to conduct traffic safety studies. Castro, et al. (2011) applied geographic information system (GIS) tools to estimate the available sight distance from vehicle trajectory and LiDAR data. This method involved the development of a digital terrain model (DTM) and viewshed analysis using ArcMap tools to find all visible points from the position of the observer (driver). All the visible points were transformed into a raster polygon and intersected with the vehicle trajectory obtained from GPS data. The available distance between the observer and the closest intersection was considered as the available sight distance. de Santos-Berbel, et al. (2014) used mobile and airborne LiDAR data to estimate the difference between sight distance derived from a digital surface model (aka triangulated irregular network) and a digital terrain model (bare surface). Probabilistic tests (Kolmogorov Smirnov and Mann Whitney Wilcoxon) were used to measure any differences in sight distance outputs using the two surface models. The paper recommended the use of a DSM model derived from mobile LiDAR data to estimate available sight distance because of its denser point clouds leading to a high-resolution model.

3. METHODOLOGY

3.1 Introduction

This chapter summarizes the equipment's used in this study, the point cloud data collection, and data processing. The data collection section covers basic information related to the software used to collect the data and information about the selected road segments for this study. The data processing section outlines the process of storing and converting the point cloud data from the rosbag file to American Society for Photogrammetry and Remote Sensing (ASPRS) recommended las file format.

3.2 Equipment Used

The point cloud data were collected with an Ouster LiDAR (OS1-16) unit, which has 16 channels or beams. The LiDAR unit has a range of 105 m (344.488 ft) with 80% reflectivity and a sampling rate of 327,680 points/sec. There are five available operating modes for this LiDAR unit used in this study: 512x10, 512x20, 1024x10, 1024x20, and 2048x10. The first value refers to the horizontal resolution, which is the number of times per rotation that each of the 16 beams is emitted. The second value is the rotation rate in hertz. The size of the data per second of recording depends on the mode because the first number determines how many vertical columns there are in one rotation's point cloud and the second number determines how many times per second the LiDAR completes a full rotation. For each mode, the maximum number of points per second recorded in the initial data file is the horizontal resolution times, the rotation rate times 64 for Ouster units. Since a 16-beam Ouster sensor was used for this study, at maximum, only 16 of the 64 points in a column in the data file were valid (i.e. x, y, and z values were non-zero). For example, the operating mode of 512x10 for the 16-beam unit would generate a maximum of 327,680 points per second, with 81,920 non-zero points at max. For 1024x20, the maximum number of points is 327,680 non-zero points per second, which is four times more than the 512x10 mode.

Ouster LiDAR operates in an 850 nm band and, unlike other sensors, this system returns only one measurement for every emitted pulse. The IMU and GNSS data were collected using an Inertial Sense µINS sensor kit. The purpose of using the IMU was to obtain the trajectory of the vehicle at locations with no GPS connection and at a higher frequent rate compared to GPS. The update rate of the three sensors (INS, GPS, and LiDAR) used for data collections were 4 milliseconds, 200 milliseconds, and 10 milliseconds. To process the collected data, an onboard AI computing device (NVIDIA Jetson TX2) was used. A 2019 Subaru Outback model vehicle was used as the mobile platform to collect the data (Figure 3.1).

Figure 3.1 Mobile LiDAR unit setup used for data collection

3.3 Data Collection Procedure

In this study, to understand the point cloud data size generated from the mobile LiDAR, three different road segments were selected (Figure 3.2). Data was collected from the selected segments at different vehicle speeds and available operating modes of the selected LiDAR unit. On survey days, the weather was sunny with a temperature of 85-90°F. The selected segments were operating under normal traffic conditions with no restrictions on vehicular traffic. During the data collection period, because of active traffic, a small change in duration(s) took place to cover the same distance at a constant speed with different LiDAR modes, which may have caused a minor effect on the data size. Additionally, scattered sunlight noise may have had a possible impact on the data collection process by interrupting the signals returning to the sensor. Intense sunlight mainly causes errors by making it difficult to differentiate between the LiDAR pulse and the photon of the sun.

The GPS, INS, and the LiDAR data were collected in the form of ROS (Robot Operating System) messages and then stored in the rosbag file format. The coordinates of the point clouds and the INS unit were stored in the local coordinate system, East, North, and Up (ENU), while the coordinates of the GPS were stored in the word geographic coordinate system (WGS-1984). An external hard drive was used to store the data due to storage capacity limitations of the NVIDIA Jetson TX2. Table 3.1 lists the roadway segments (in Logan, UT), vehicle speeds, and LiDAR operating modes used for data collection in this study.

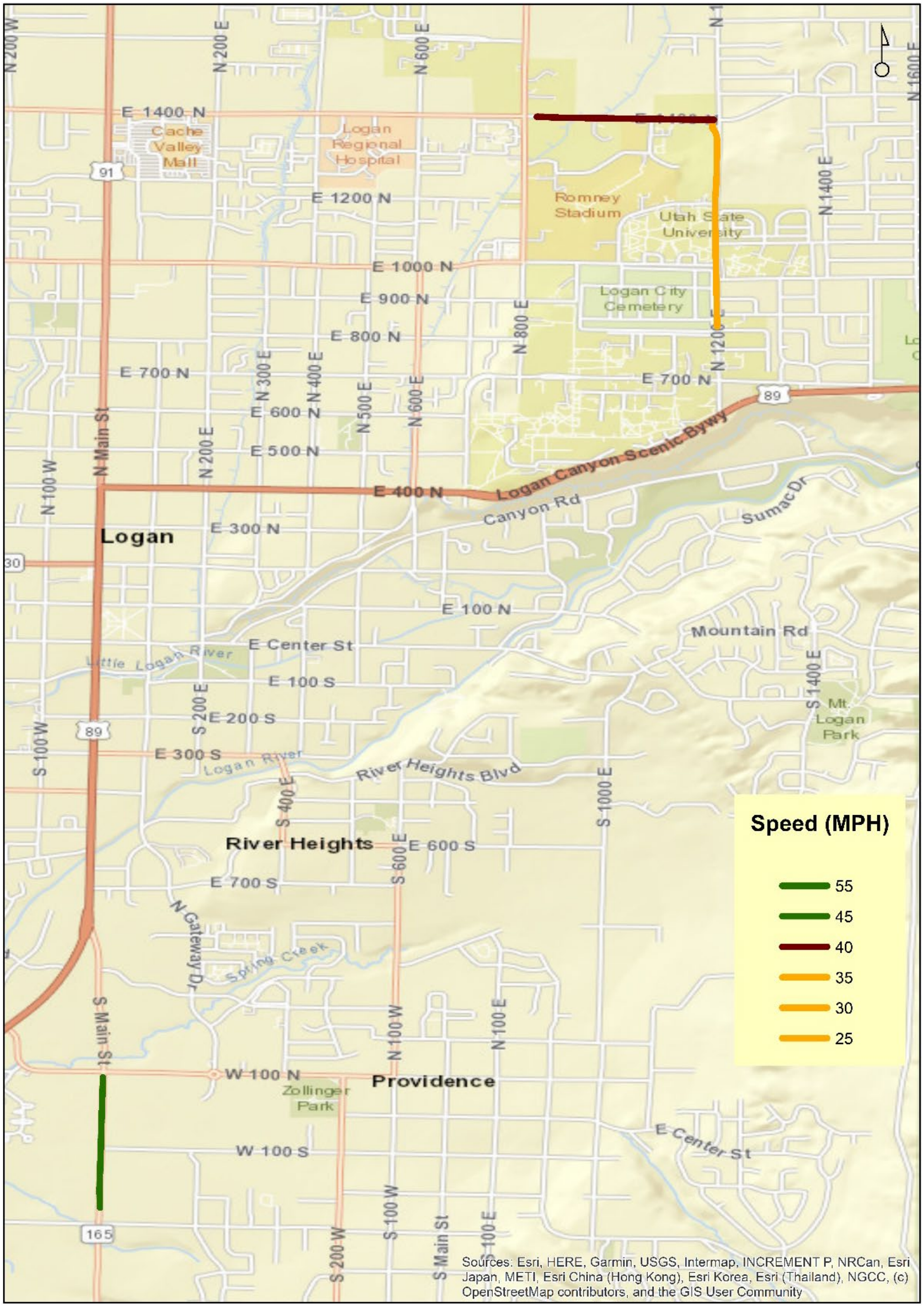


Figure 3.2 Data collection map

Table 3.1 Data Collection Road Segments, Speeds, and Operational Settings

Road Segment	Starting Point	Ending Point	Total Lanes	Length (mi)	Vehicle Speed (MPH)	Operating Mode
N 1200 E	E 850 N	E 1400 N	2	0.71	25	512x10
						512x20
						1024x10
						1024x20
						2048x10
					30	512x10
						512x20
						1024x10
						1024x20
						2048x10
					35	512x10
						512x20
						1024x10
						1024x20
						2048x10
E 1400 N	N 1200 E	N 800 E	2*	0.48	40	512x10
						512x20
						1024x10
						1024x20
						2048x10
UT-165	E 1200 S	E 1700 S	4*	0.5	45	512x10
						512x20
						1024x10
						1024x20
						2048x10
					50	512x10
						512x20
						1024x10
						1024x20
						2048x10

*Includes a 12-ft two-way-left-turn lane.

To understand data sizes for the travel lanes only and the right of way only, a buffer was made following the Cache County, UT, standards for arterial and collector roads. Right of way (ROW) represents the area of land along and around a roadway that is owned, operated, and maintained by the owner of the road. Most assets within this boundary fall under jurisdiction of the owner, which is often a local or state transportation agency. The travel lanes are the areas of the roadway where motorized vehicles operate, not including parking lanes, bike lanes, shoulders, etc. Figure 3.3 represents a generic roadway cross-section, including the right of way and travel lanes. The dimensions in this figure are not the same as those used in this study. Figure 3.4 demonstrates visually how the point cloud is impacted by the implementation of the right-of-way and travel lane buffers, based on both vehicle trajectory and the centerline.

In this study, the width of a travel lane and the right of way for arterial roads are 12 feet and 100 feet, respectively. The values are 11 feet and 80 feet, respectively, for collector roads. The buffer was made based on the vehicle trajectory developed by the INS. However, due to the technical movement of the vehicle on the road, an equal buffer zone for both sides of the road were not possible to make. That's why the buffer zone may have covered more area for one side of the road than the other side for the original vehicle trajectory.

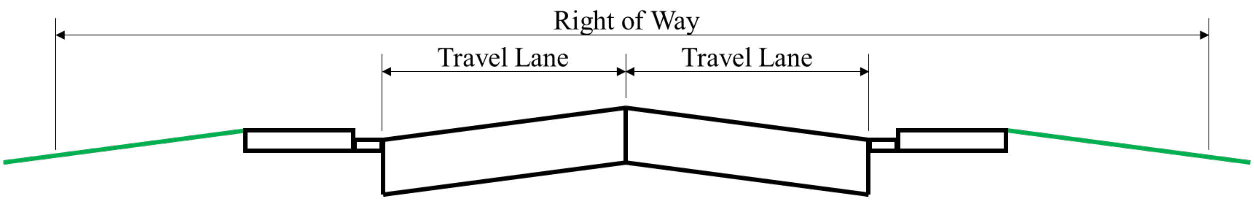
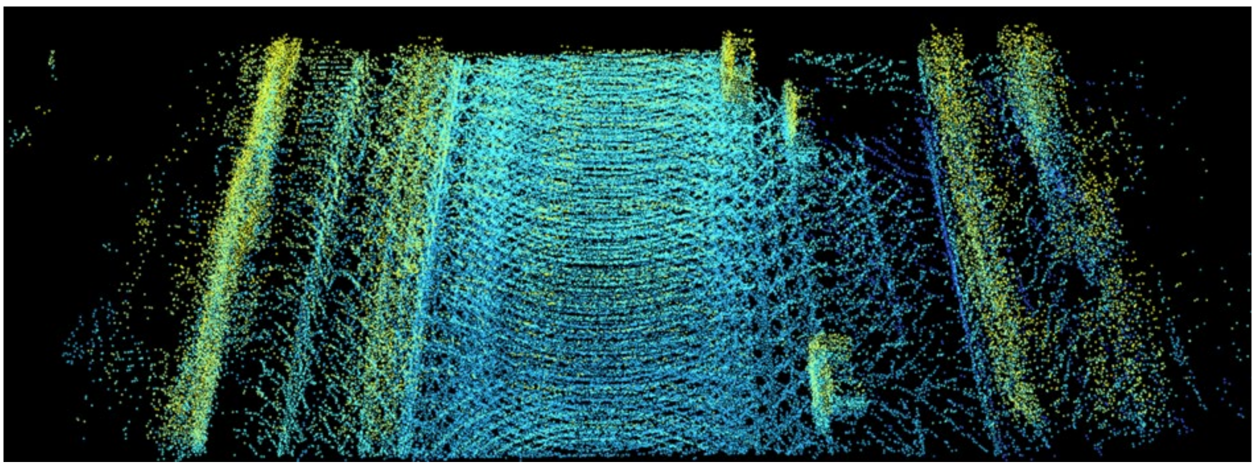
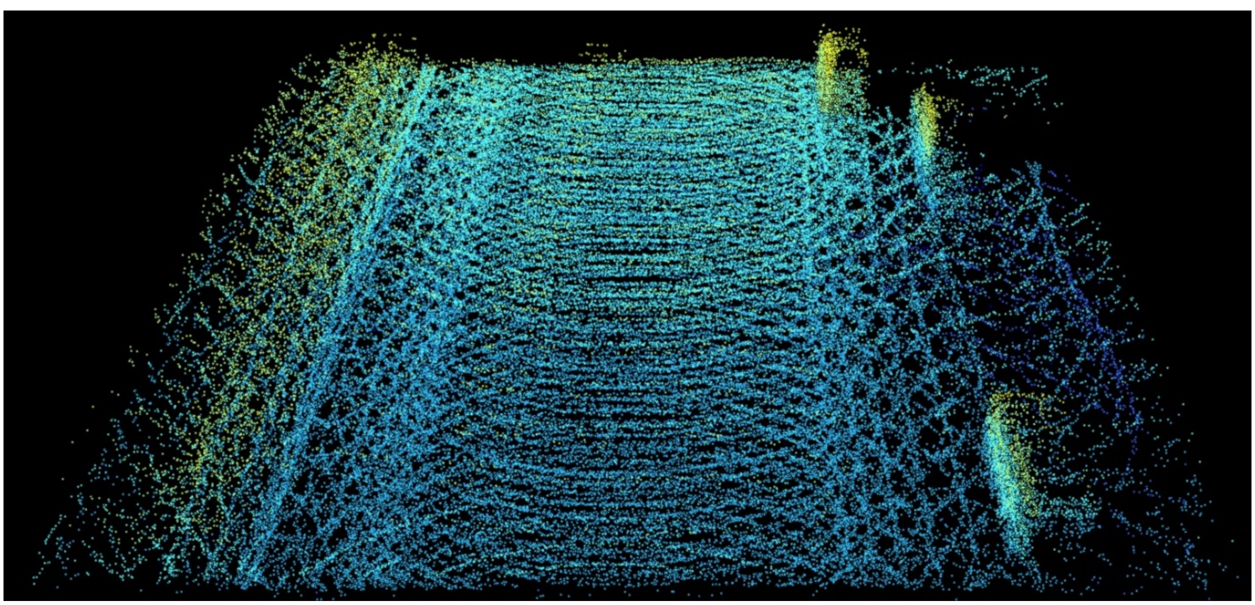


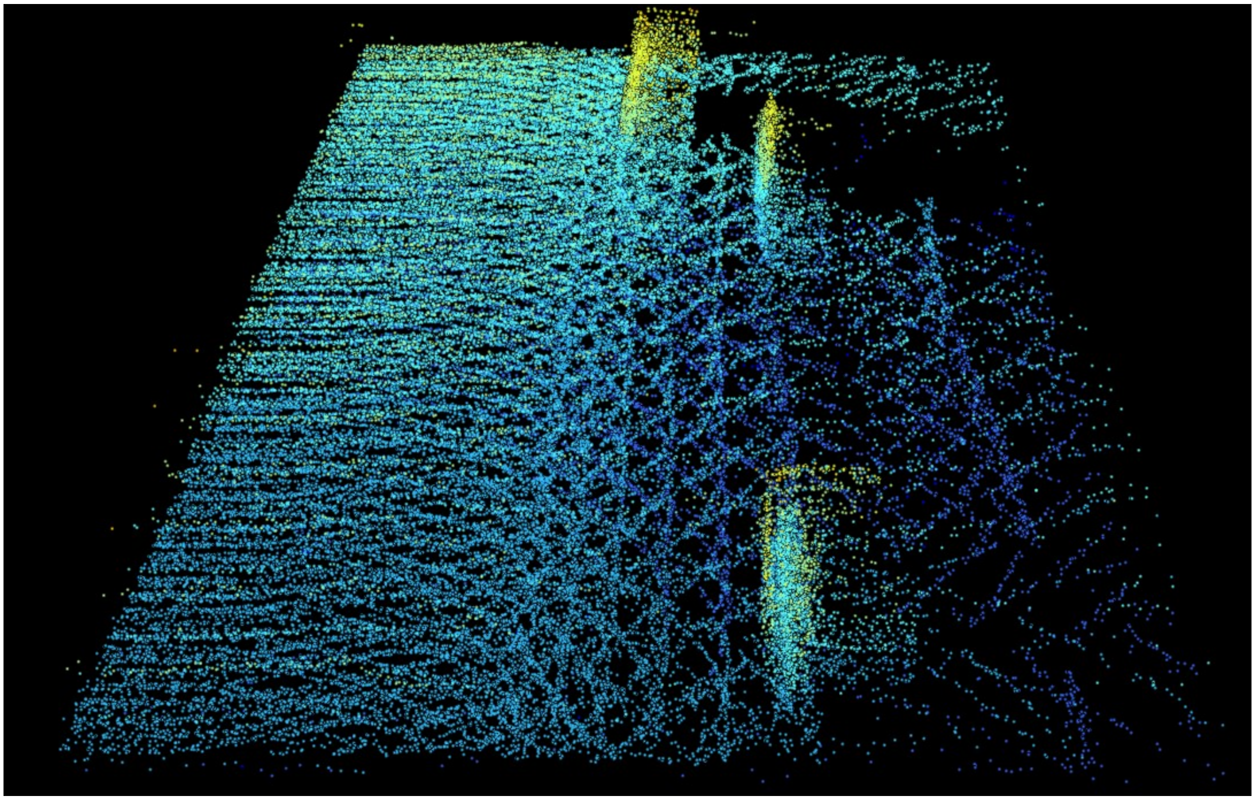
Figure 3.3 Cross section design of a typical highway



(a) Total



(b) Right-of-way with buffer based on centerline



(c) Travel lanes with buffer based on centerline

Figure 3.4 Point cloud data collection sample from an active highway

In order to generate an equal buffer zone for the both sides of the road, the centerline of the road was considered as the vehicle trajectory and relevant data for the selected segments was estimated. The LAStools platform was used to convert the shape file into the las file format, which is a binary file recommended by the ASPRS to store LiDAR data. The LAS version 1.4 was used to save the point cloud data. At the initial collection stage, there were many attributes recorded for each point, but only 4 attributes were included (x, y, z, and the intensity) in the final las file. Finally, to comprehend the data size for the selected segments, the relevant individual las files for each roadway segment and each LiDAR mode were combined into one single file.

For an example of the raw data, Figure 3.5 shows a sample snapshot of the LiDAR unit's view as the vehicle traverses a grocery store parking lot. From top to bottom, the image includes ambient, intensity, range, and point cloud views. It is a simple depiction of the point cloud in one instant during a drive. This may be what an autonomous vehicle "sees" as the passenger travels to the store. Figure 3.6 depicts a compiled point cloud map from a drive around a parking lot lasting approximately five minutes. Figure 3.7 shows a sample point cloud from one of the data collection roadway segments, with some road features called out. For mobile LiDAR platforms, this represents the most common first step in data processing. Table 3.2 provides an example of the raw point cloud data directly generated from the sensor unit.

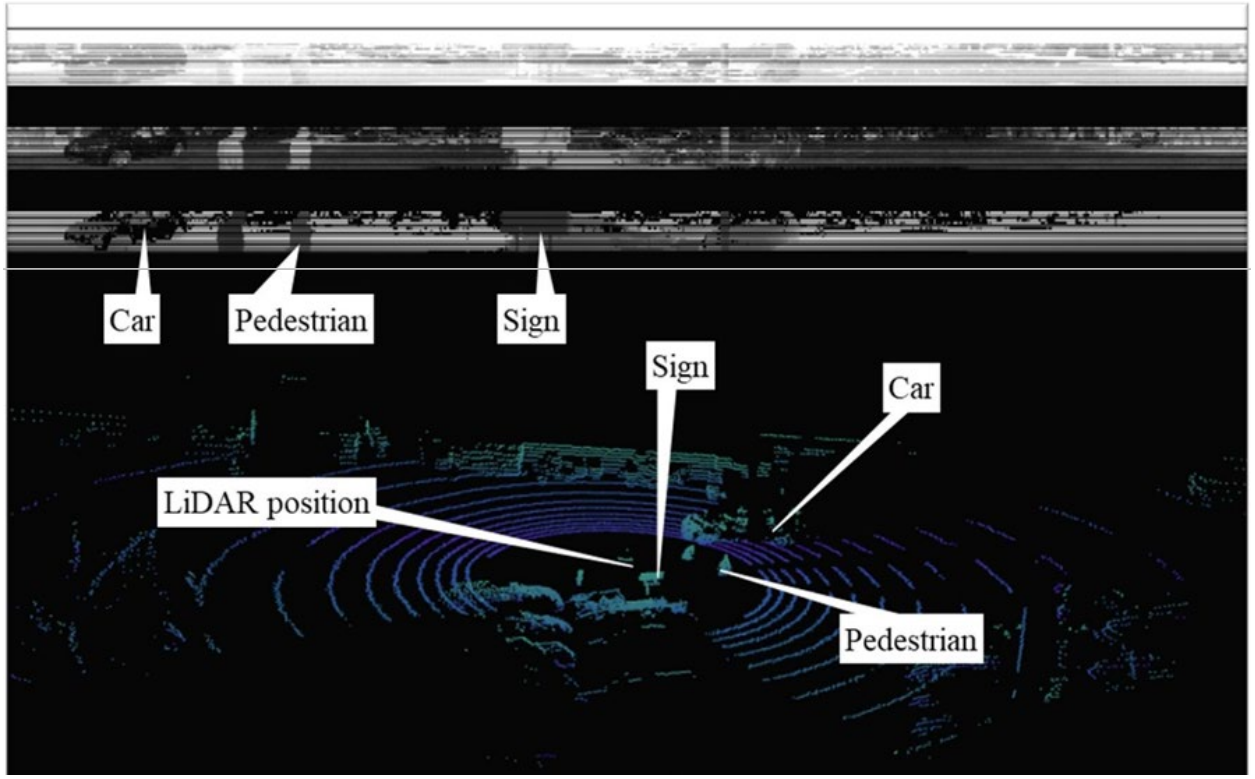


Figure 3.5 Stream LiDAR view at a grocery store parking lot

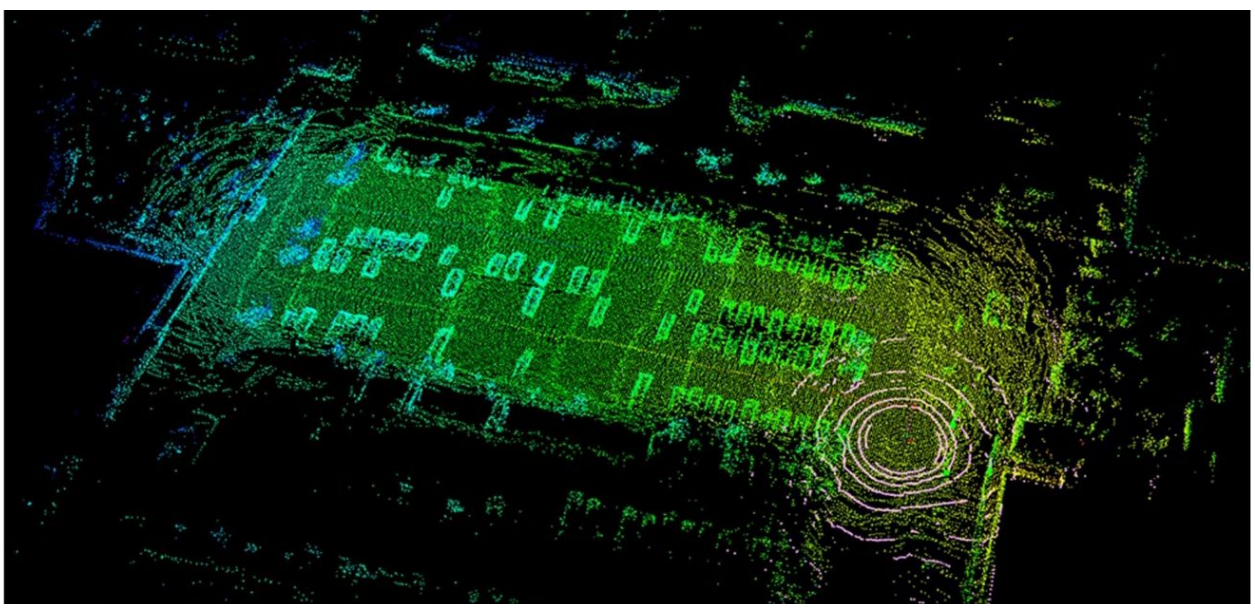


Figure 3.6 Sample point cloud map generated at a parking lot

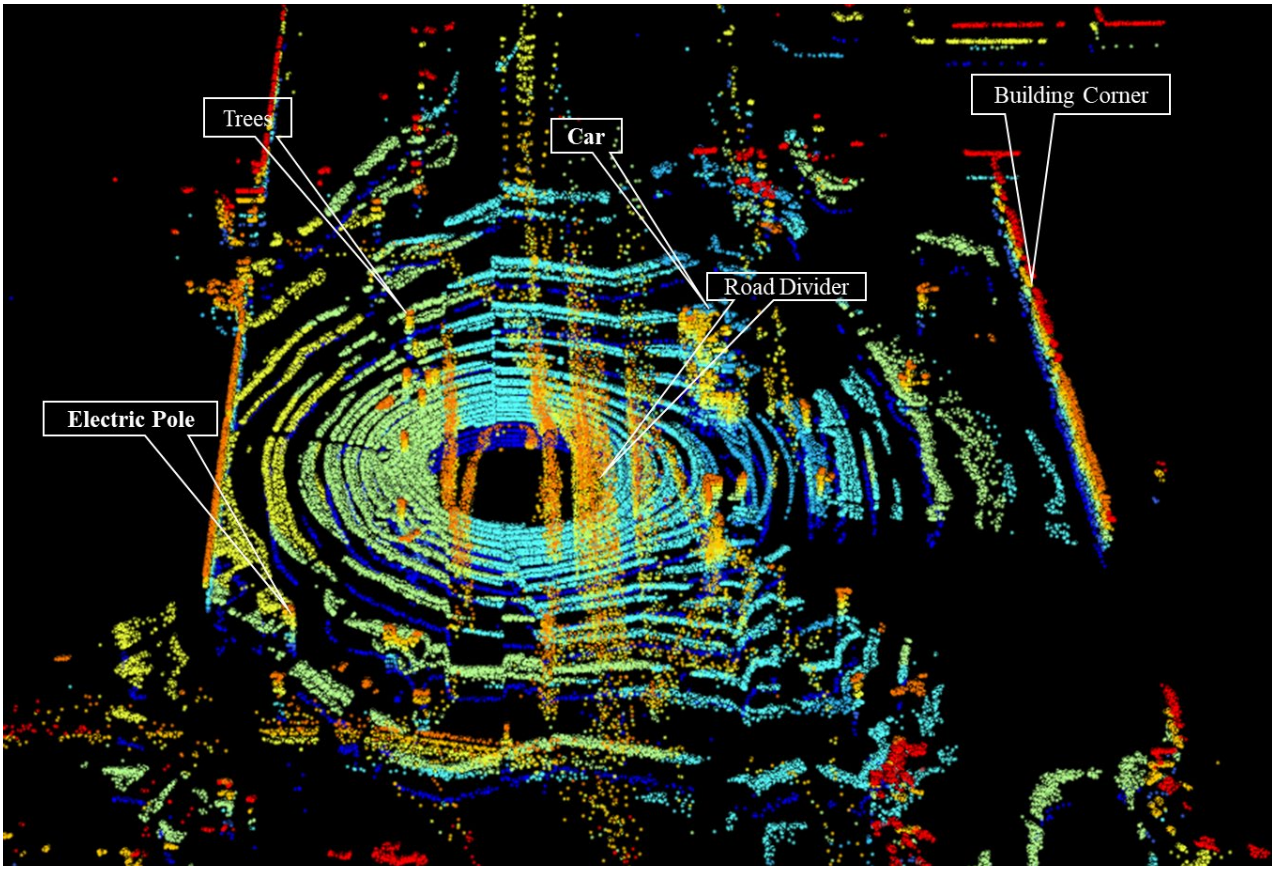


Figure 3.7 Sample point cloud map generated on a road

Table 3.2 Point Cloud Data Sample

X (m)	Y (m)	Z (m)	Intensity	t (ms)	Reflectivity	Ring	Noise	Range (mm)
31.721	0.544	-2.795	38	0	4140	42	302	32845
24.038	-1.349	-2.480	77	0	4426	44	396	24203
19.911	0.330	-2.438	157	0	6272	46	480	20062
16.488	-0.939	-2.315	23	0	642	48	93	16637
13.874	0.225	-2.220	45	0	889	50	114	14052
11.591	-0.667	-2.062	19	0	263	52	126	11792
10.181	0.161	-2.017	18	0	197	54	79	10380
9.033	-0.632	-1.963	41	0	354	56	77	9259
8.123	0.124	-1.925	37	0	260	58	75	8349
7.245	-0.434	-1.862	47	0	266	60	70	7493
6.611	0.099	-1.833	50	0	236	62	108	6861
33.966	-2.004	-2.249	40	49040	4692	40	420	34099
33.129	0.459	-2.830	33	49040	3687	42	306	33253
24.057	-1.424	-2.482	71	49040	4143	44	394	24227

4. DATA ANALYSIS

4.1 Introduction

The increased availability of 3D point cloud data has created new challenges in the transportation industry regarding the analysis and incorporation of this informative big data into different decision-making processes rather than using it only for visualization purposes. The use of big data resulting from LiDAR technology is gaining popularity as the development of AVs continues. LiDAR sends millions of pulses per second to survey the objects within its range. Guan et al. (2013) stated that LiDAR data usually contains tens or hundreds of points per square meter. Therefore, the processing of LiDAR data is highly computational and data intensive even for a small area. Tomljenovic and Rousell (2014) stated that an area of 0.2 square miles can generate a point cloud from 500 GB to 3 TB depending on the selected point cloud density. The point cloud density depends on the scanning capacity of the sensor. Vehicle speed, scanning speed, and system measurement rate all influence the resolution of the collected data according to Guan et al. (2016). Besides the storage, processing, and computation of the LiDAR data for a large-scale area poses a significant challenge.

4.2 Regression Model Development

In this study, the collected data detailed in the previous chapter were extrapolated for large-scale estimation and consideration by public agencies. A linear regression model was developed for the purpose of extrapolation. In the model, the data size was considered as the dependent variable. The duration of the data collection period (*X,* in seconds), the duration of the LiDAR to complete a single rotation (*Z,* in milliseconds), and the number of columns in the point cloud data (*Y*) are considered as the explanatory variables. The following equations estimate the data size for the different sections (travel lanes, right-of-way, and total) as discussed in Section 3.4 and shown in Figure 3.3. Table 4.1 shows the input data used for estimating the linear regression models.

$D_{TL,veh} = -23.240 + 0.344X + 0.061Y + 1.714Z$ with $R^2 = 0.825$ Eq. 4.1

$D_{TL,cl} = -55.460 + 0.936X + 0.052Y + 1.988Z$ with $R^2 = 0.570$ Eq. 4.2

$D_{ROW,veh} = -119.021 + 1.899X + 0.100Y + 3.853Z$ with $R^2 = 0.691$ Eq. 4.3

$D_{ROW,cl} = -302.634 + 4.611X + 0.141Y + 6.390Z$ with $R^2 = 0.671$ Eq. 4.4

$D = -380.913 + 5.774X + 0.179Y\_8.192Z$ with $R^2 = 0.663$ Eq. 4.5

Where:

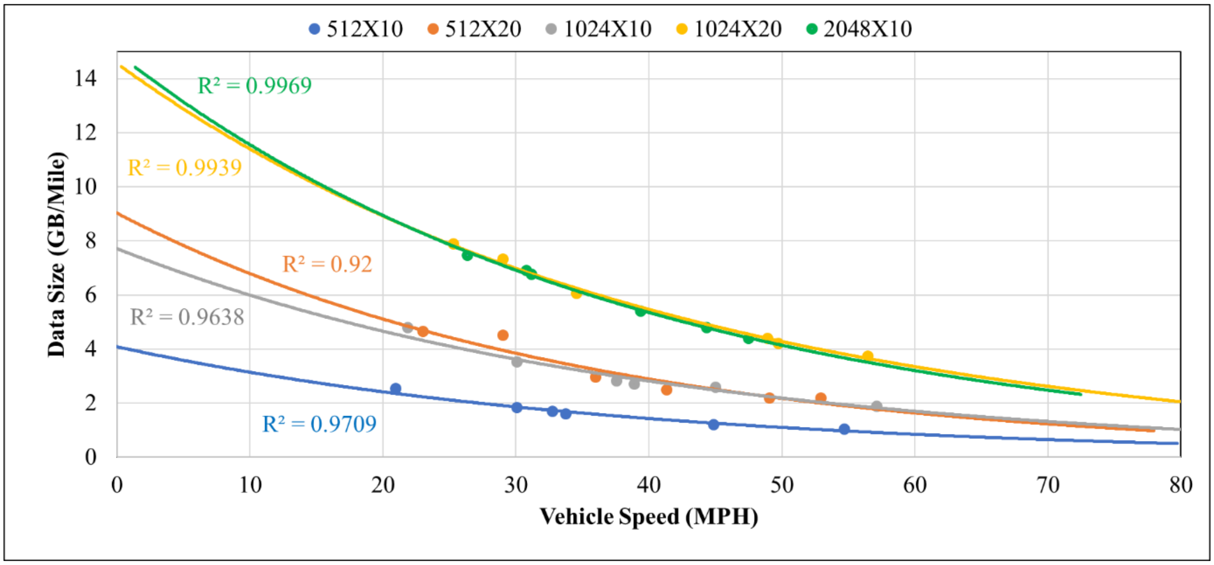
$D_{TL,veh}$ = total data falling within the travel lanes buffer based on the vehicle trajectory (MB)
$D_{TL,cl}$ = total data falling within the travel lanes buffer based on the centerline (MB)
$D_{ROW,veh}$ = total data falling within the right-of-way buffer based on the vehicle trajectory (MB)
$D_{ROW,cl}$ = total data falling within the right-of-way buffer based on the centerline (MB)
D = total data (MB)
X = duration of the data collection period (s)
Y = number of columns in the point cloud data (a.k.a. sampling rate per rotation)
Z = duration of a complete, single rotation of the LiDAR (ms)

Table 4.1 Regression Model Data Input

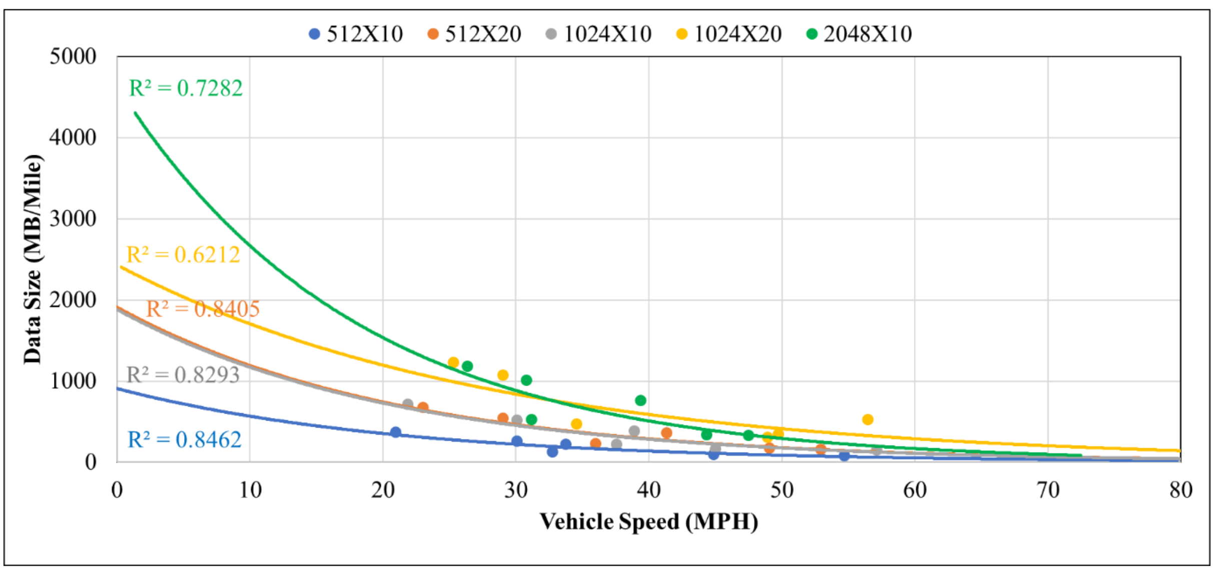
D (MB)	$D_{ROW,veh}$ (MB)	$D_{TL,veh}$ (MB)	$D_{ROW,cl}$ (MB)	$D_{TL,cl}$ (MB)	Time, X, (s)	Number of Columns, Y	Refresh Rate, Z (ms)
264.000	115.000	54.000	213.786	64.653	122.0	512	10
480.000	205.000	98.264	381.098	117.649	111.0	1024	20
508.000	215.000	103.584	399.688	124.018	117.0	512	10
875.000	372.000	178.860	691.554	214.145	101.0	1024	20
842.000	357.000	175.390	663.668	210.000	97.0	2048	10
183.820	80.000	37.626	148.721	45.048	85.0	512	10
384.000	161.000	77.926	299.301	93.299	88.0	1024	20
367.650	156.000	75.250	290.006	90.095	85.0	512	10
761.253	323.000	155.853	600.462	186.600	88.0	1024	20
718.000	305.000	147.000	567.000	176.000	83.0	2048	10
90.000	70.300	19.670	71.968	22.909	78.0	512	10
163.000	128.000	35.000	131.037	40.764	71.0	1024	20
155.346	122.317	32.780	125.219	38.179	68.0	512	10
338.107	266.220	70.000	272.537	81.529	74.0	1024	20
374.660	295.000	76.500	302.000	89.100	82.0	2048	10
106.300	70.600	33.500	80.399	40.400	51.2	512	10
174.000	115.276	56.000	131.277	67.534	41.8	1024	20
184.364	122.446	59.280	139.442	71.489	44.4	512	10
254.123	169.000	162.000	192.458	94.450	30.6	1024	20
365.000	242.500	234.000	276.161	135.000	43.9	2048	10
48.070	39.600	30.480	37.201	26.055	40.1	512	10
88.000	72.490	55.800	68.100	47.700	36.7	1024	20
84.350	79.000	61.188	72.392	54.000	40.0	512	10
155.240	137.500	103.000	126.000	90.200	36.8	1024	20
171.272	152.000	114.000	139.280	99.750	40.6	2048	10
39.640	31.710	21.000	30.254	18.480	32.9	512	10
81.940	66.000	46.000	63.000	40.480	34.0	1024	20
76.300	63.200	48.900	60.300	42.600	31.5	512	10
175.369	145.300	105.000	134.694	79.159	36.2	1024	20
164.070	137.000	102.000	127.000	76.900	37.9	2048	10

4.3 Graphical Analysis

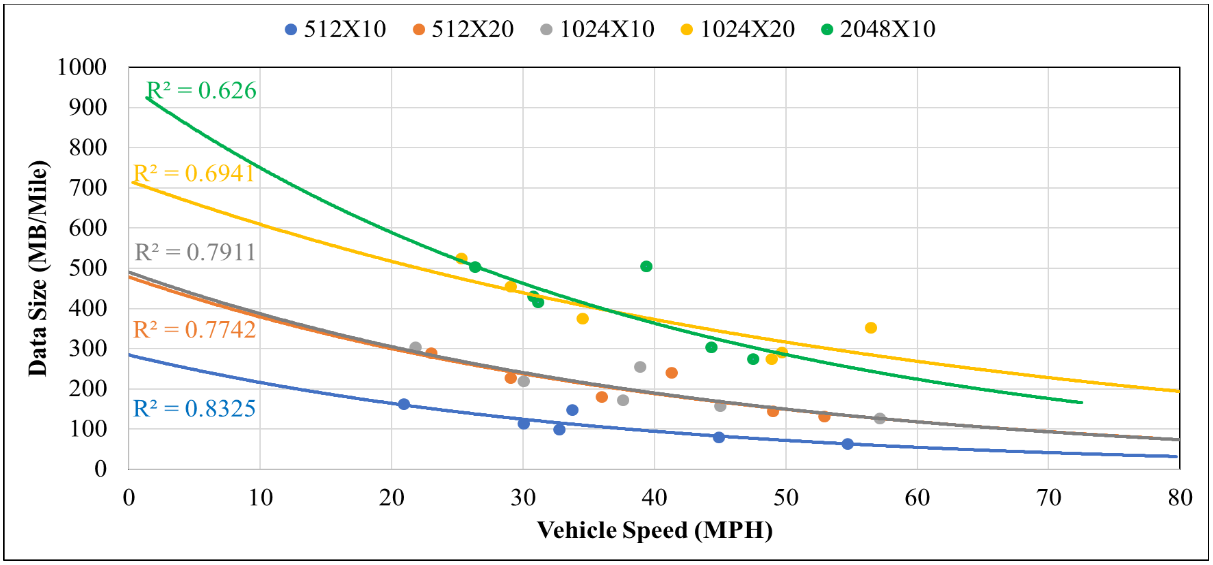
Figure 4.1 represents the projected data sizes in megabytes based on the measured data sizes for six different scenarios: unfiltered total, (Figure 4.1a), filtered total (Figure 4.1b), right-of-way with the actual vehicle trajectory (Figure 4.1c), travel lanes with the actual vehicle trajectory (Figure 4.1d), right-of-way with the centerline (Figure 4.1e), and travel lanes with the centerline (Figure 4.1f). The markers represent measured data sizes from the data collection runs detailed in Chapter 3. The lines represent projections of data size for other speeds using exponential functions. Exponential functions were used instead of other trendline types with higher R-squared values because the data size per mile cannot be negative. Also, as speeds approach zero, the data size per mile will theoretically approach infinity. The data sizes by speed are normalized for distance traveled. As can be expected, data sizes at low speeds are much higher than at higher speeds. Agencies can expect that for a single vehicle, total data sizes collected for low-speed roads will be larger and have greater point density than for high-speed roads.



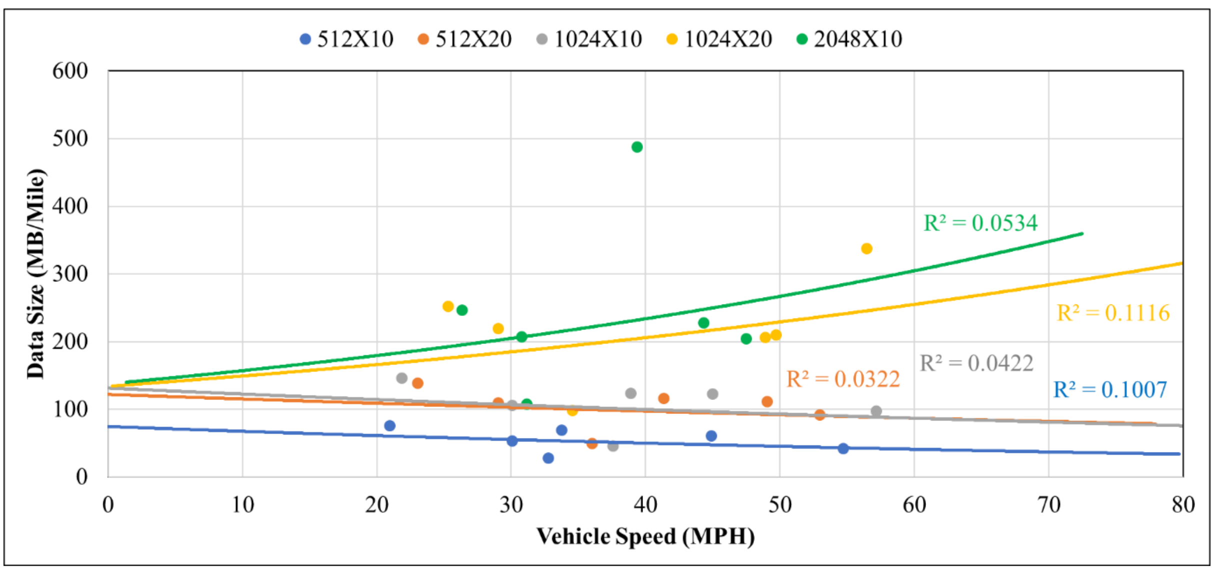
(a) Unfiltered total (including “zero” points)



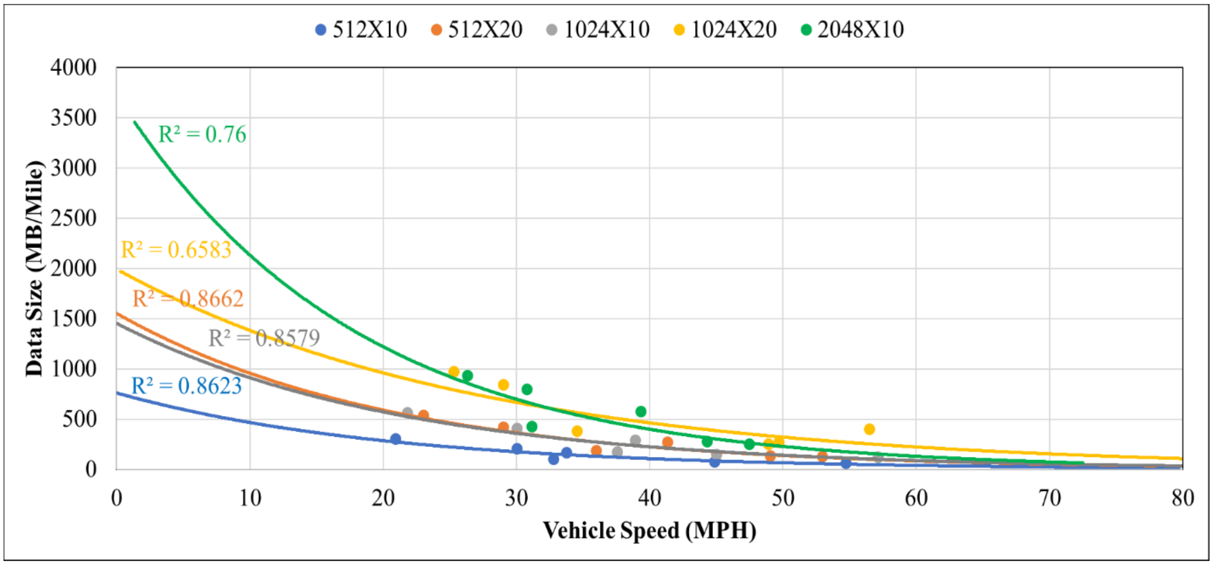
(b) Total



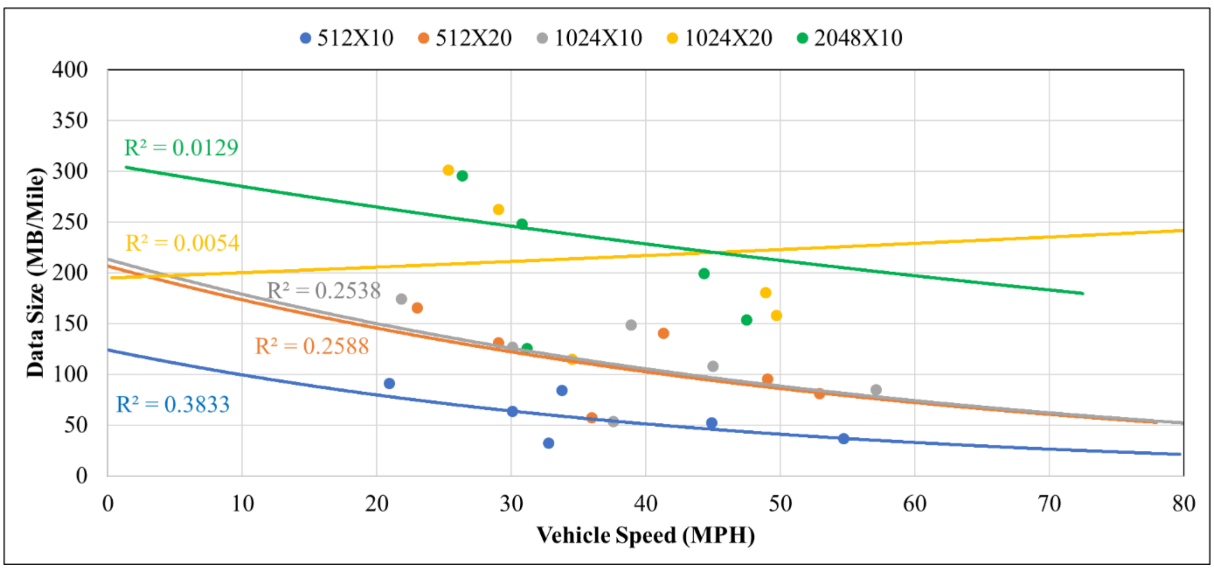
(c) Right-of-way with buffer based on actual vehicle trajectory



(d) Travel lanes with buffer based on actual vehicle trajectory



(e) Right-of-way with buffer based on centerline

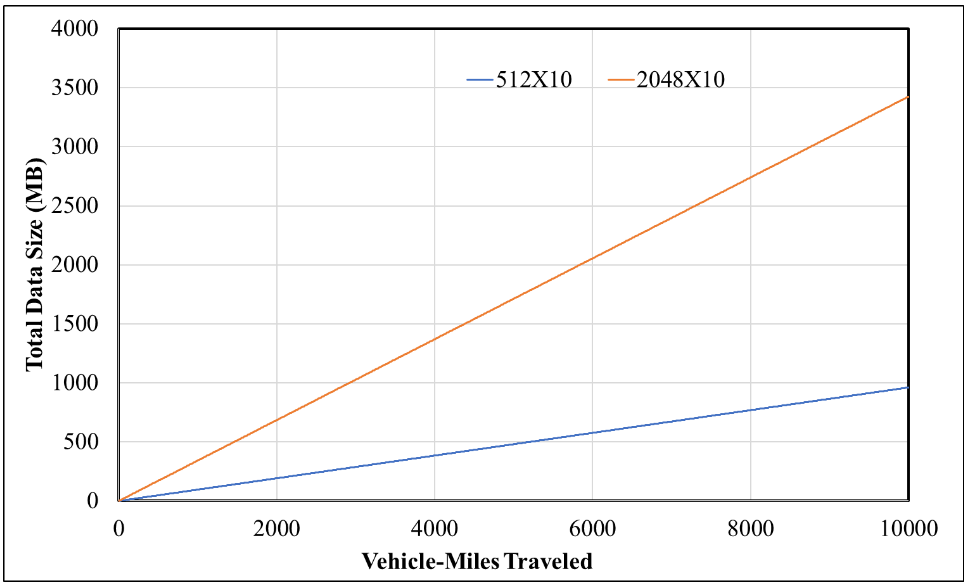


(a) Travel lanes with buffer based on centerline

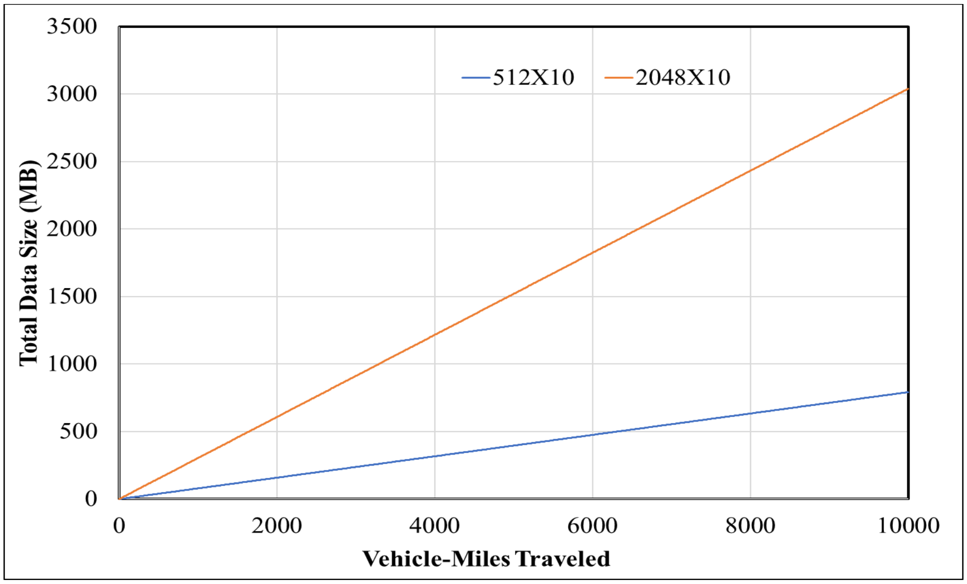
Figure 4.1 Measured and projected data size by speed and LiDAR mode

Figure 4.1a is different than the remaining figures of this section as it resembles the raw data size generated directly from the sensor without any filtering or file format conversion. For the travel lanes scenario, at 80 mph with the lowest operating mode (512x10) of the selected LiDAR, close to 125 MB/mile of data is expected for centerline trajectory while for actual trajectory approximately 75 MB/mile of data can be expected. With the highest operating mode, a little more than 300 MB/mile can be expected for the centerline trajectory and around 360 MB/mile for the actual trajectory. As for the total scenario, at 80 MPH with the lowest LiDAR operating mode (512x10), close to 800 MB/mile can be expected while with the highest operating mode (2048x10) more than 4300 MB/mile can be expected. The range of the LiDAR data size that can be expected for the right-of-way scenario with actual vehicle trajectory is from 300 MB/mile to 900+ MB/mile for a vehicle travelling at 80 mph. However, in terms of the centerline trajectory, the range of the data is from 750 MB/mile to about 3500 MB/mile. Graphs such as these can provide transportation agencies with a range of conservative to liberal outlooks regarding data size expectations. However, it is important to note that these estimations are based on limited runs in a variable environment. The R-squared value in these graphs are not consistent because of the dynamic environment of the study area. This issue is particularly apparent with the travel lanes data extrapolation, since only a small slice of the data is considered, which also happens to have the most variation due to the presence (or lack thereof) of other vehicles. To get more accurate and reliable data size estimations, multiple runs in different environment can help. However, given the limitations of this study, a single run for each LiDAR mode for any particular road segment was conducted.

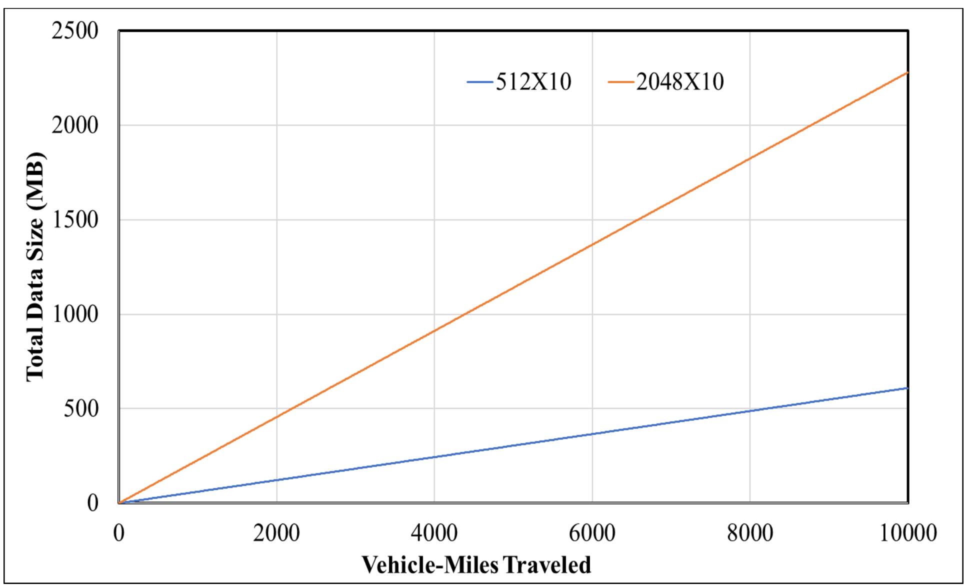
The projected total data sizes depicted in Figure 4.2 were developed considering vehicle-miles traveled (VMT) at 45 mph. The blue line represents the data size that can be expected for the lowest operating LiDAR mode, while the red line depicts the data size for the highest operating mode of the selected LiDAR sensor.



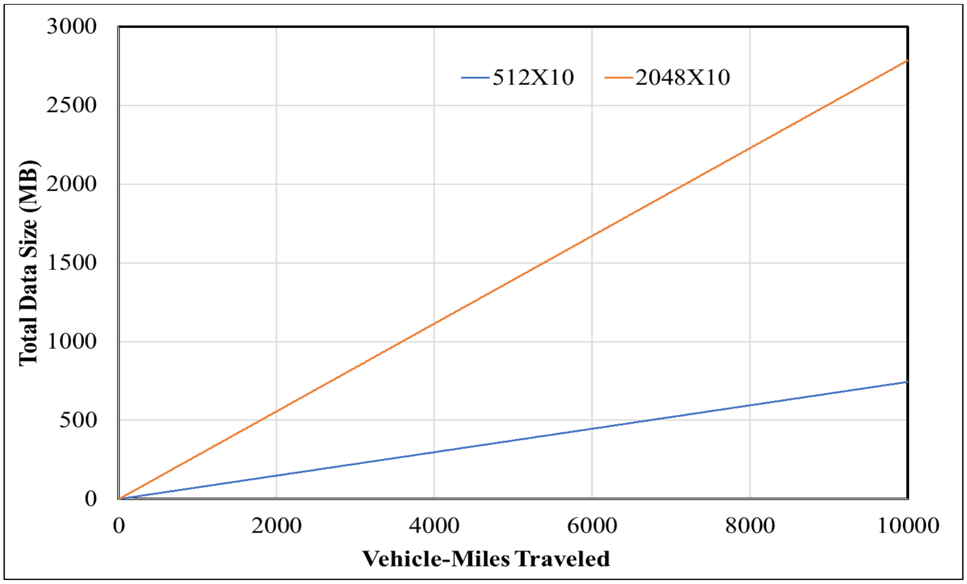
(a) Total



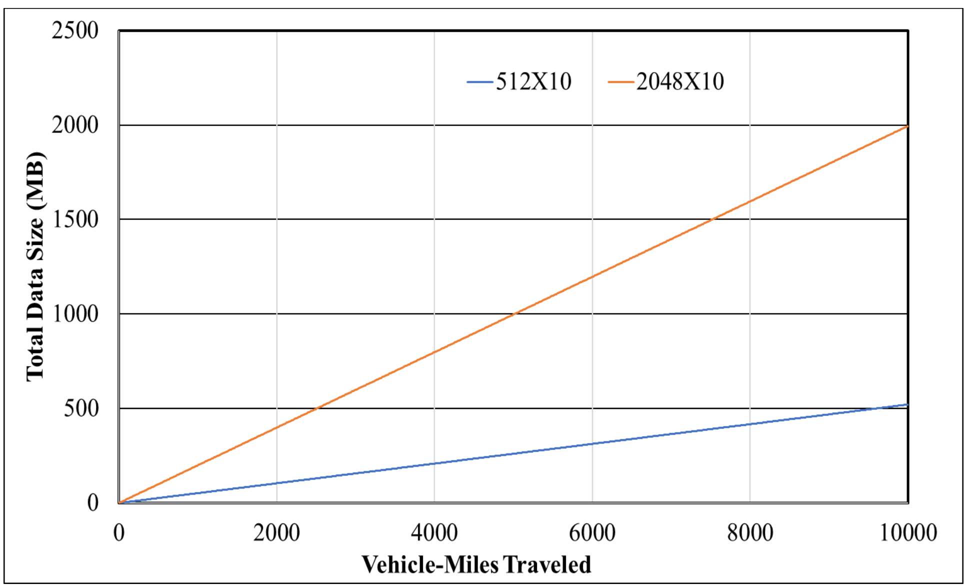
(b) Right-of-way with buffer based on actual vehicle trajectory



(c) Travel lanes with buffer based on actual vehicle trajectory



(d) Right-of-way with buffer based on centerline

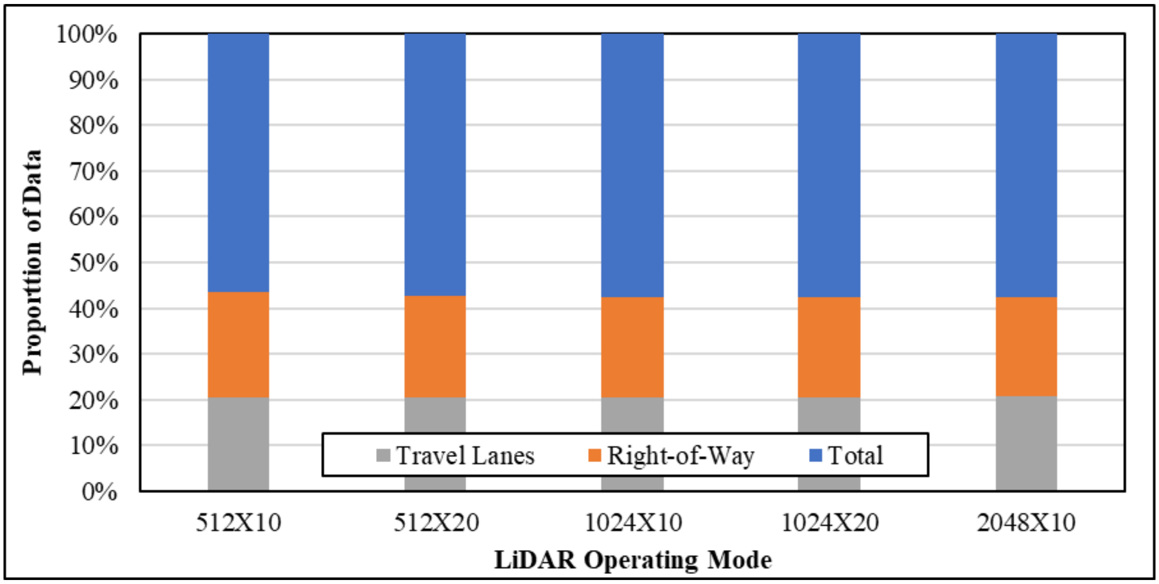


(e) Travel lanes with buffer based on centerline

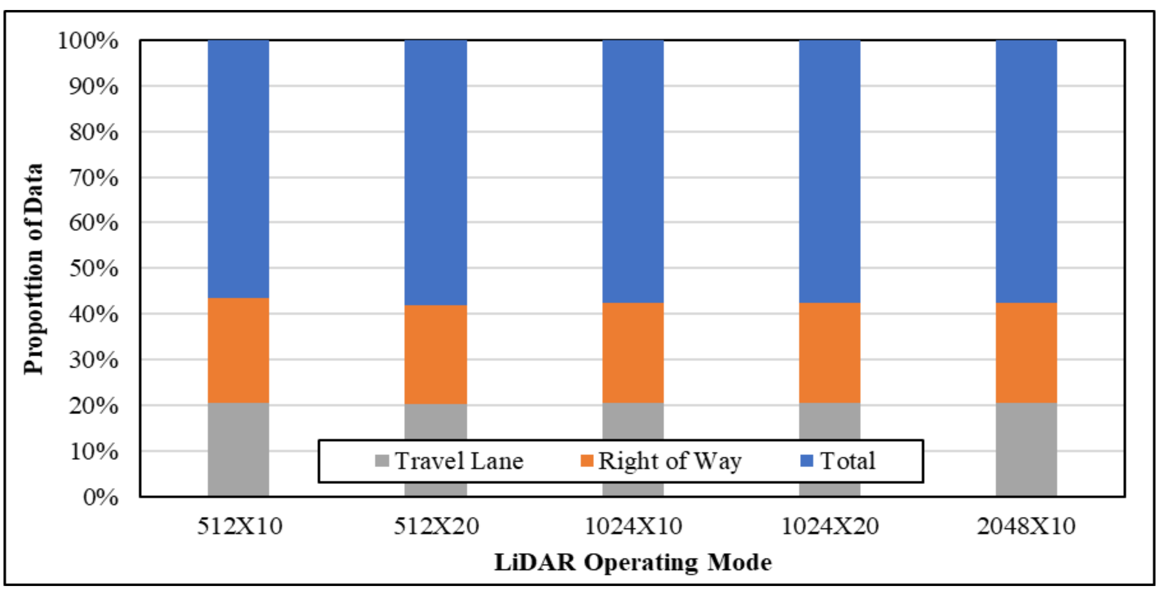
Figure 4.2 Projected data size by vehicle miles traveled

In the case of data generated from the selected LiDAR unit, the total dataset produced with 10000 VMT is expected to range from approximately 1000 MB (512x10 mode) to more than 3200 MB (2048x10 mode). For 10000 VMT within the right-of-way along the centerline trajectory, the data size ranges from approximately 500 MB to 2750 MB. It is notable that for the right-of-way the data generated from the original trajectory is more than the data size considering the road centerline for a vehicle travelling at 45 MPH.

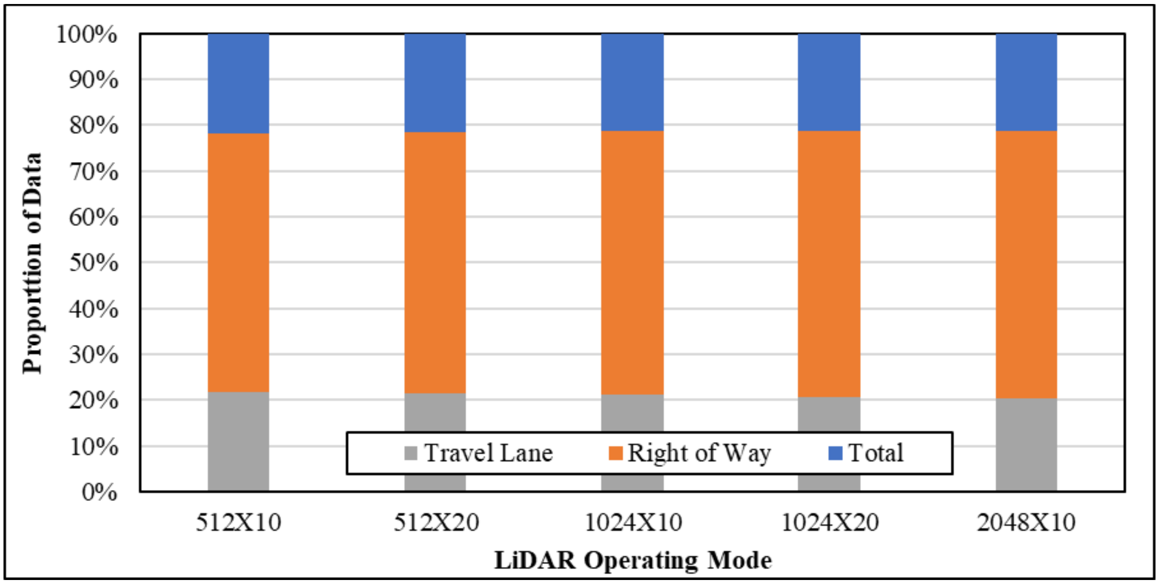
Figure 4.3 and Figure 4.4, based on the vehicle trajectory and centerline, respectively, represent the distribution of the collected data size in megabytes from the selected highway segments for each particular speed. The blue segment represents the data portion which is outside of the right-of-way. The orange segment represents the data portion outside of the travel lanes but within the right-of-way (using a buffer of 100 feet for the arterial and 80 feet for the collectors). The gray segment represents the data portion within the travel lanes (using a buffer of 60 ft for the arterials (45 and 55 miles/hr), 36 ft for the 40 miles/hr collector road, and 24 ft for the other collector road (25, 30, and 35 miles/hr).



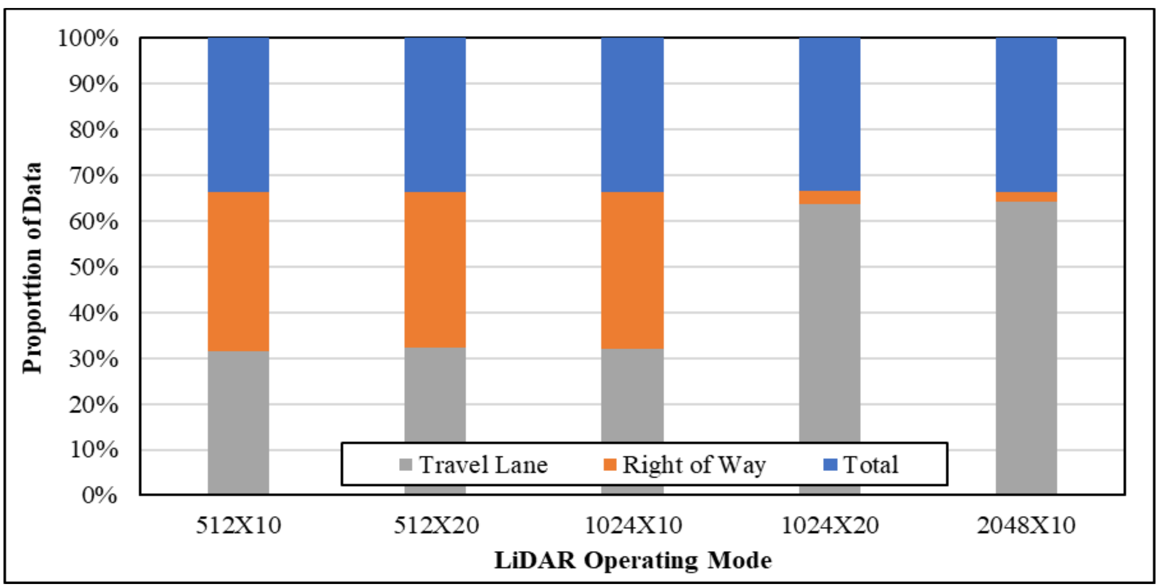
(a) 25 mph



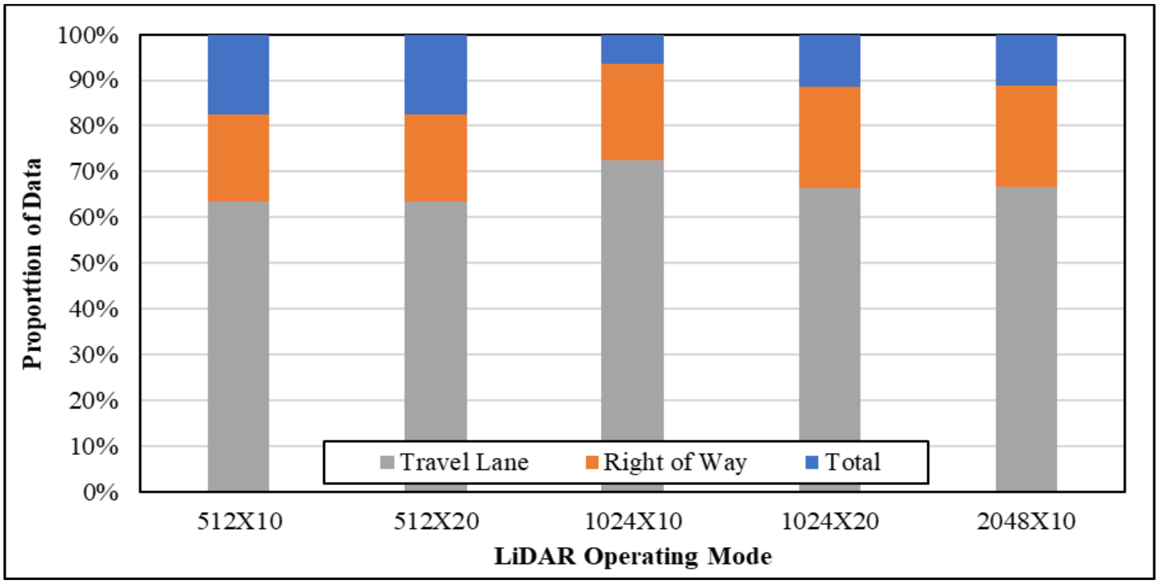
(b) 30 mph



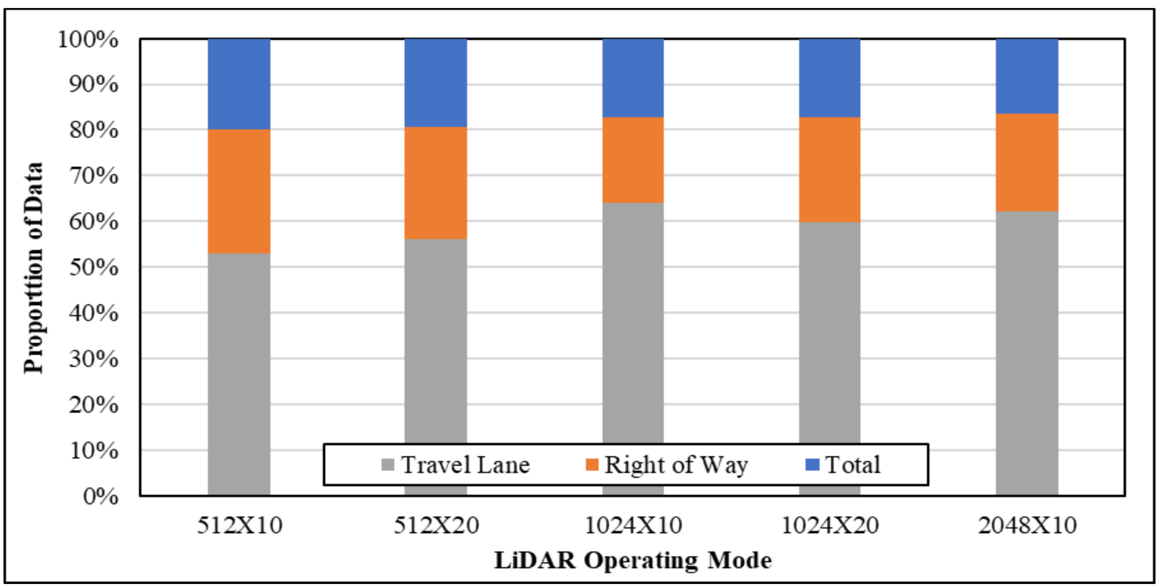
(c) 35 mph



(d) 40 mph

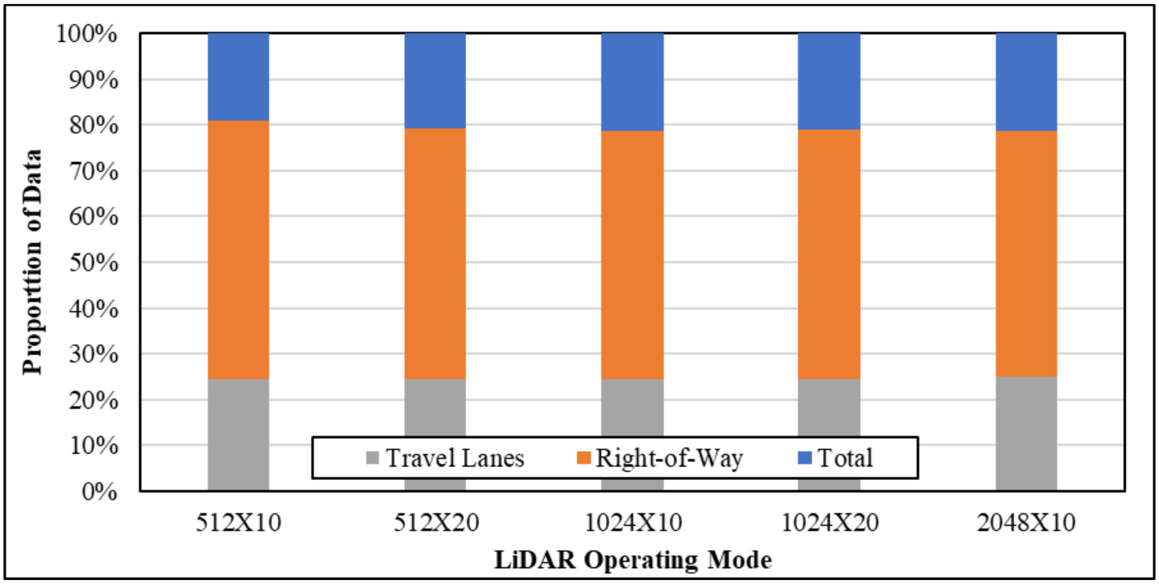


(e) 45 MPH

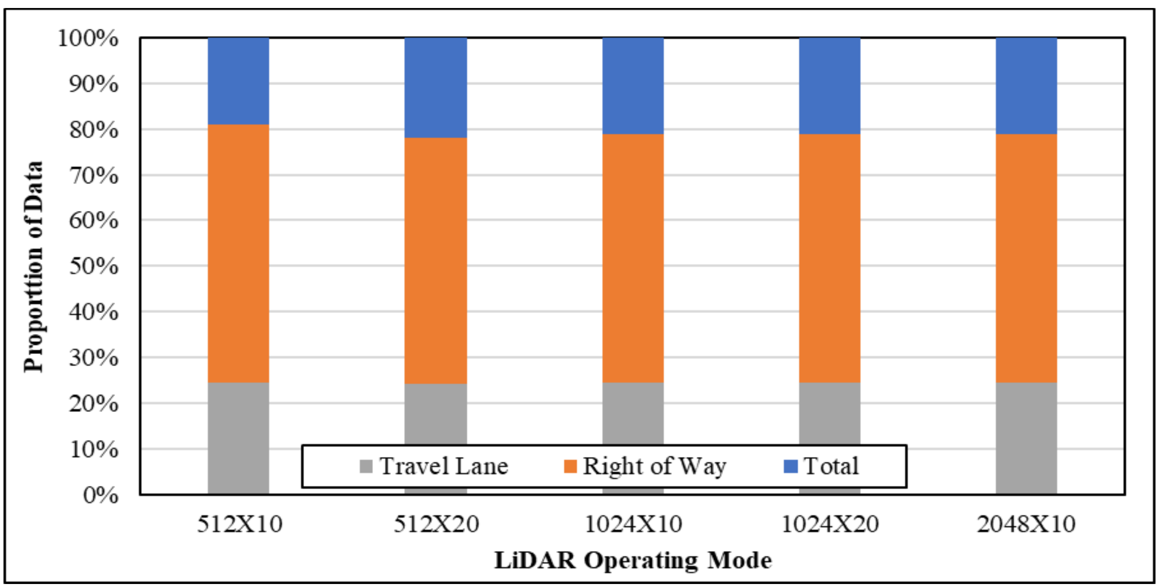


(f) 55 MPH

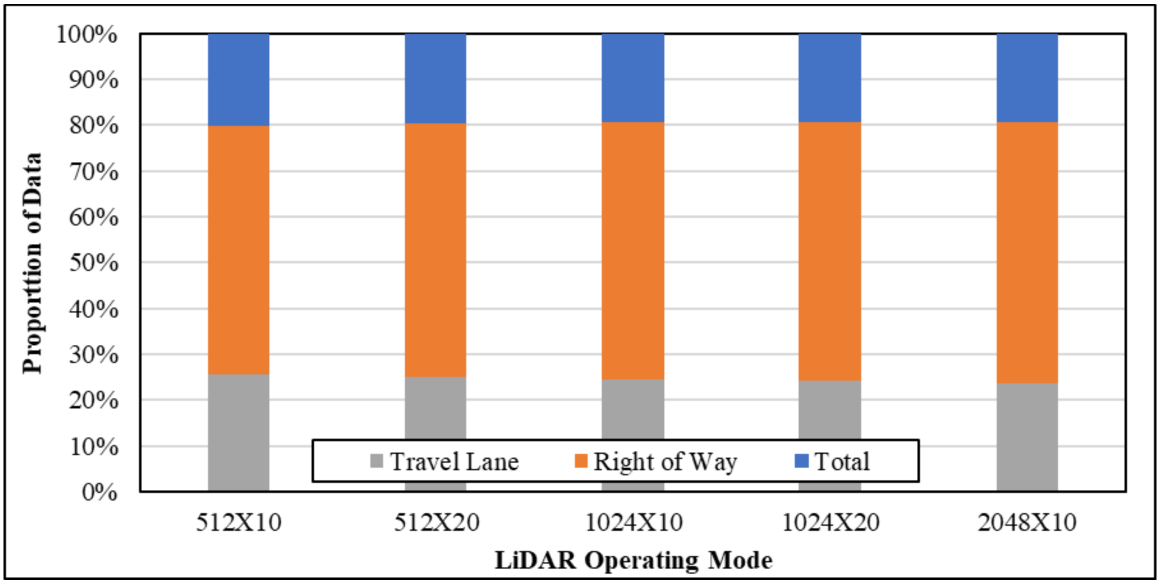
Figure 4.3 Data size distribution at different speeds and operating modes with buffers based on actual vehicle trajectory



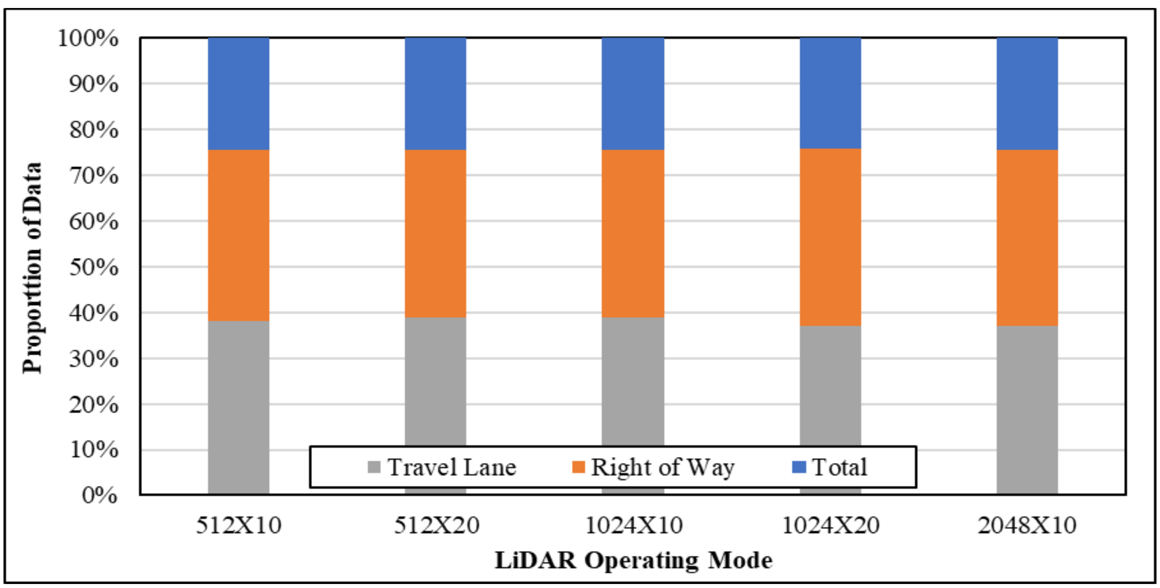
(a) 25 MPH



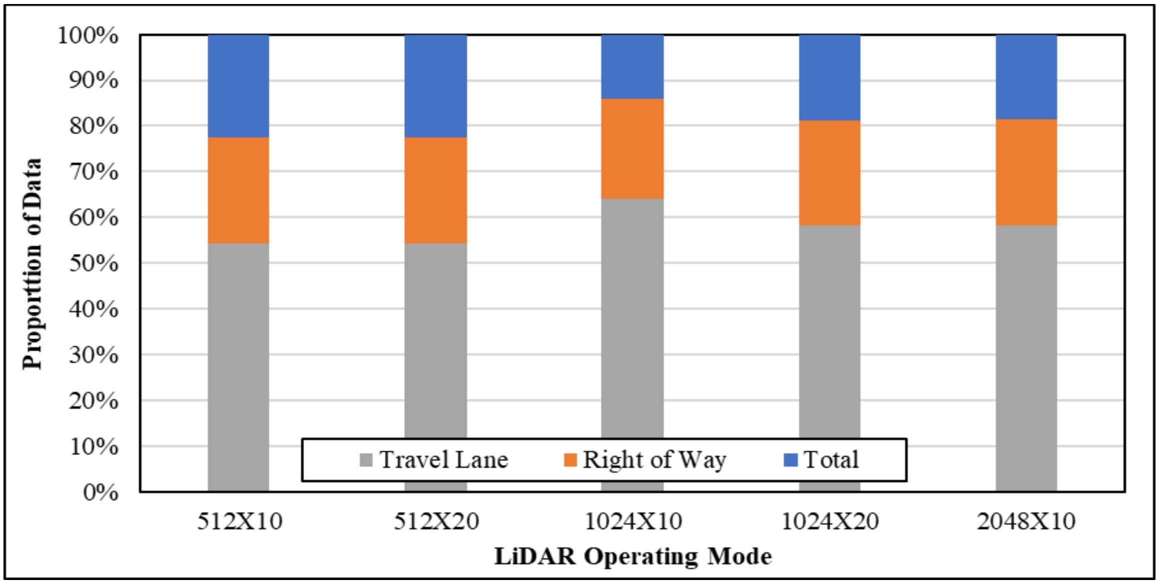
(b) 30 MPH



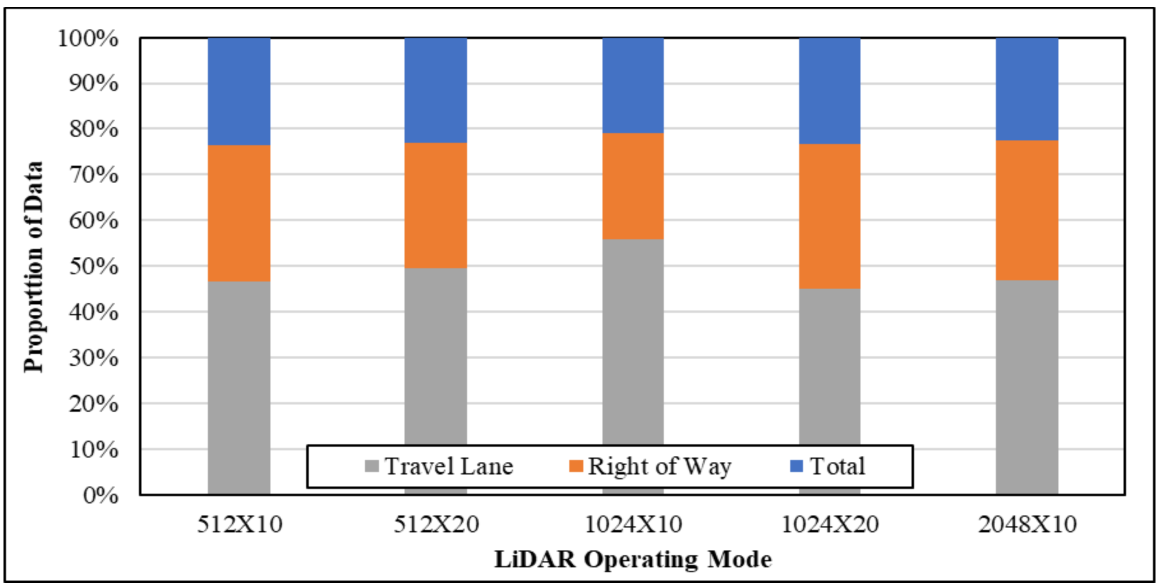
(c) 35 MPH



(d) 40 MPH



(e) 45 MPH

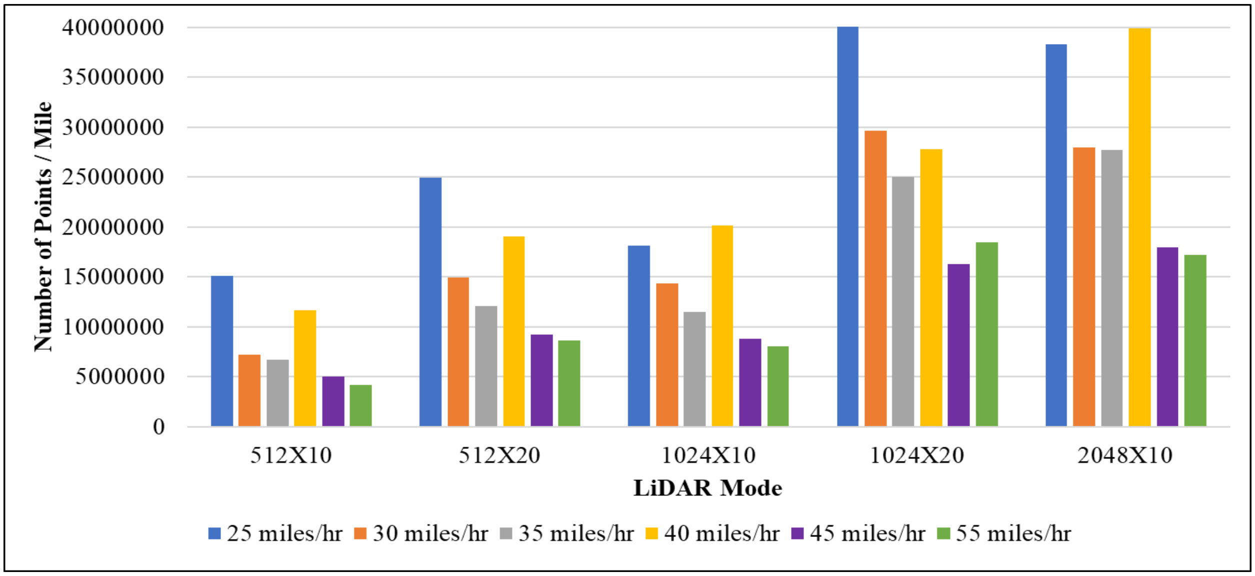


(f) 55 MPH

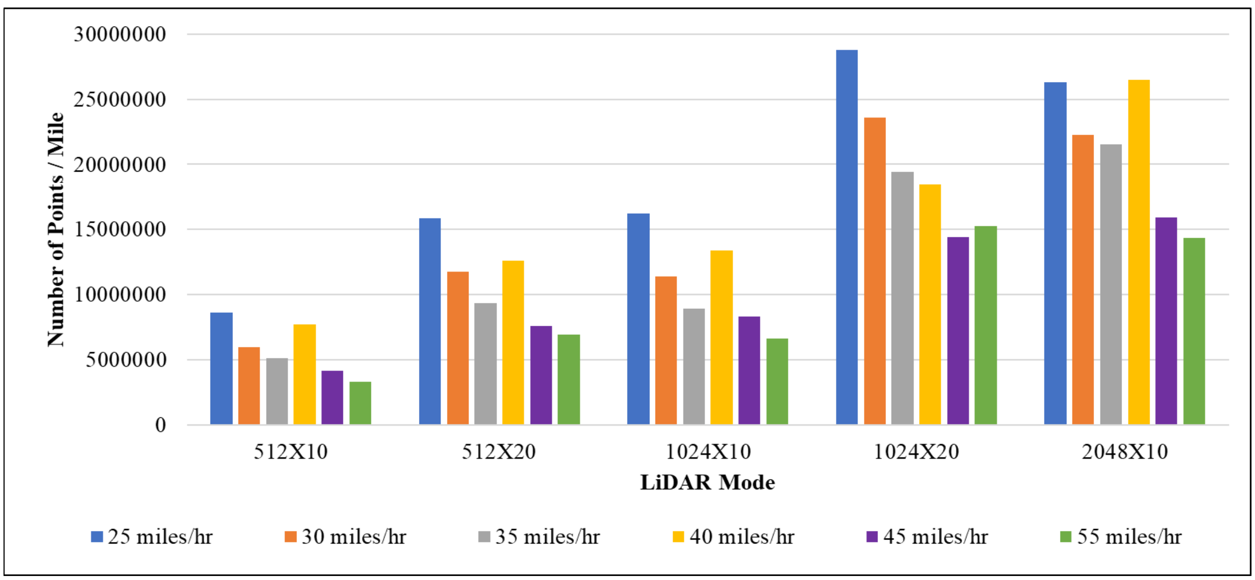
Figure 4.4 Data size distribution at different speeds and operating modes with buffers based on centerline

As expected, distribution of data points within the travel lanes and right-of-way are fairly consistent regardless of speed and operating mode except the 40 miles/hr (actual trajectory). Data size for the right-of-way is higher in 35 miles/hr compared to all the speeds for both the actual and the centerline trajectory. The proportion within the right-of-way on the collector roads had more variation due to the higher frequency of objects along the selected roads. The collector roads had greater variety and frequency of landscaping, signage, and on-street parking compared to the arterial.

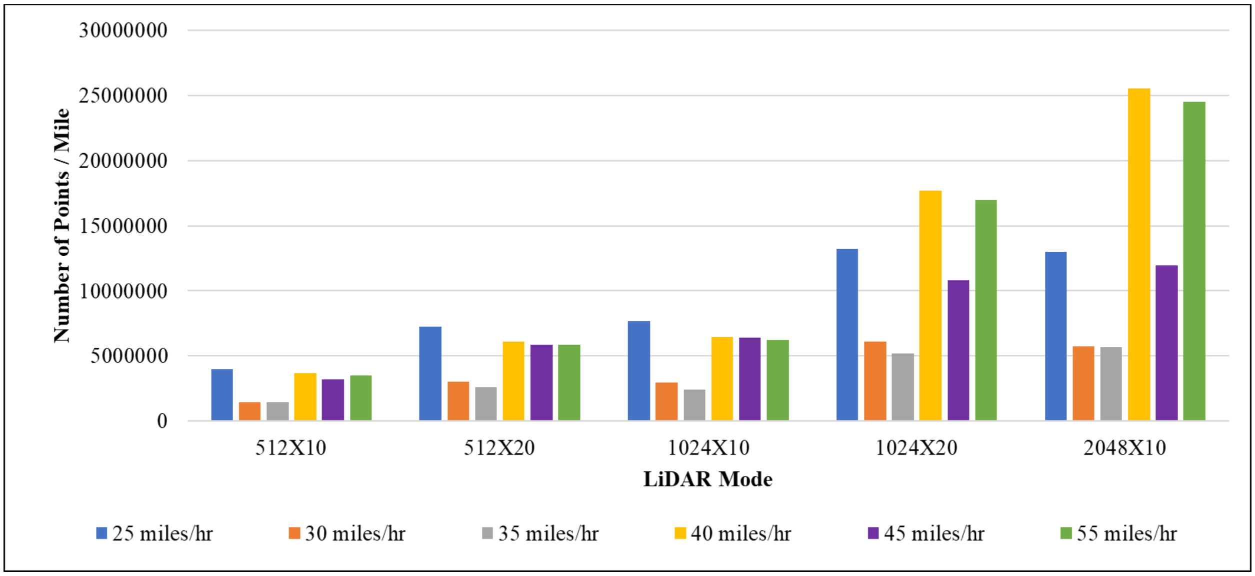
The number of points per mile measured in individual test runs (Figure 4.5) varied with the frequency of reflective objects within the range of the LiDAR unit. As expected, the number of points increased based on the operating mode of the LiDAR unit. However, the point density still varied within each LiDAR mode because of the dynamics of the mobile platform's speed and the environment from which the data were collected. The highest number of points were produced in the 2048X10 mode at 25 miles/hr, while the lowest number of points were produced in the 512X10 mode at 55 miles/hr. In general, it can be expected that the number of points in a point cloud data set mainly depends on the speed of the platform, the environment and the scanning capacity and speed of the sensor.



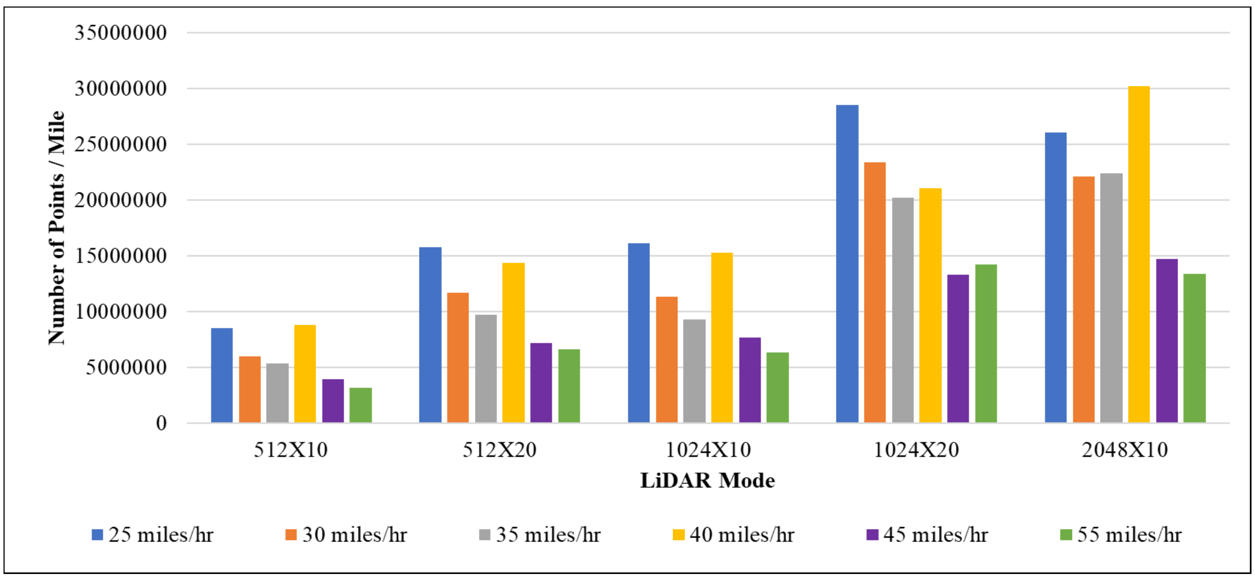
(a) Total



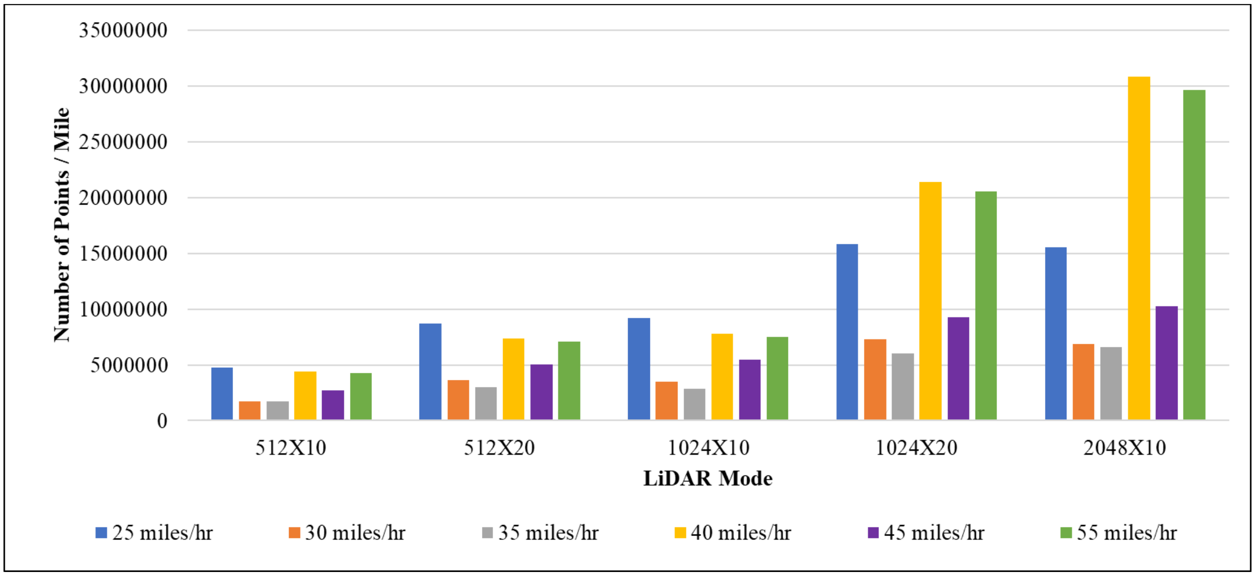
(b) Right-of-way buffer based on actual vehicle trajectory



(c) Travel lanes buffer based on actual vehicle trajectory



(d) Right-of-way buffer based on centerline



(e) Travel lanes buffer based on centerline

Figure 4.5 Number of points per mile

5. DISCUSSION

5.1 Introduction

Diverse applications of LiDAR data have changed the transportation field in recent decades. These applications are expected to continue to grow with the development of the technology and autonomous vehicles. Implication of the size of the data collected during this study is discussed here to provide an understanding about the possible data size that one can expect from a mobile platform from a similar study environment. Apart from this, it also provides a broader overview on the existing challenges for LiDAR data application in the transportation industry.

5.2 Mobile LiDAR Data Uses in Transportation

In this study, mobile LiDAR refers to the collection of 3D point data from a mobile platform. Mobile LiDAR combines multiple sensors synchronized on a mobile platform to generate detail-oriented georeferenced 3D point cloud data. In recent years, it has become a common and effective method for infrastructure mapping and monitoring among transportation agencies because of its affordability and efficiency. Mobile LiDAR is a primary data collection tool in the geospatial development industry. This technology can produce high resolution images of features at a standard speed with survey grade accuracy without disrupting the normal traffic flow on the highway. This advantage of mobile LiDAR significantly reduces the cost and time associated with data collection compared to other conventional methods. Mobile LIDAR systems provide a dense, geospatial dataset as a 3D virtual world that can be explored from a variety of viewpoints across a transportation agency. Although MLS reduces the duration of the data collection phase, it requires significant time and resources to extract meaningful information and even more time for developing 3D models (Sairam, et. al. 2016).

Mobile LiDAR has increased safety benefits given that all data collection tasks are performed inside the vehicle. It also mitigates impacts to traffic flow given that it does not require roadway closures. With proper processing, mobile LiDAR data can be used for various purposes such as traffic monitoring or crash analysis (Figure 5.1). Traffic monitoring is an effective way to assess and enhance traffic performance on a road network. Considerably few studies have been conducted on this issue with LiDAR. Most previous studies focused on highway asset management, object detection, and object classification. LiDAR has also gained popularity in terms of structural monitoring of assets such as culverts or embankments. For feature extraction from LiDAR data, identification of reflective objects (i.e., signs) has naturally received more attention than identification of non-reflective (and often non-standardized) objects. There has also been a lack of research in the area of mobile LiDAR assessment of highway design elements such as cross section, vertical and horizontal alignment, and clearance. A review report by NCHRP stated that there was a lack of knowledge in terms of extracting information from LiDAR (Olsen et al. 2013).

5.3 Challenges of Using LiDAR

Being the eye for an autonomous vehicle, LiDAR technology is an essential component of ADAS. However, there are some core challenges associated with the usage of this technology. High-performance spatial indexing is needed for quick data access while working with LiDAR since it is common to have projected data sizes of 1 TB or more (Liu et al. 2016). Olsen et al. (2013) stated that the complexity of LiDAR data and lacking availability of processing software presents challenges to the end users of the data. Often, studies use only a limited portion of the data as the end product.

Zhang et al. (2017) determined that volume of the data and the computational and technical requirements to process, manage, and store the point cloud can pose a great challenge. According to Nashashibi and Bargeton (2008), point cloud processing usually requires working between multiple software packages where information can be lost during import and export stages of the process. The technicality to operate and collect data with LiDAR is complex compared to traditional survey methods, as it often must be integrated and synchronized with other sensors like GPS and IMU. A mobile LiDAR platform has to move at a certain speed in an active highway environment while collecting data simultaneously. However, because of this speed, potential target objects are missed, and the resulting measurements are not collected. Lovas et al. (2004) stated that with airborne LiDAR, the ability to detect small features is affected by the altitude, orientation, and quality of the unit. For any platform, target objects may be obscured by other objects in the foreground. Spoof signals can cause trouble to the users by producing incorrect information, which may produce nonexistent objects in the LiDAR data processing algorithm. In the case of use by an autonomous vehicle, this error could cause a crash.

The perception and range of LiDAR is limited, which also poses a significant challenge to work in a dynamic environment. In addition, the atmospheric scattering and attenuation caused by weather phenomena, such as sun reflection, dust, rain, snow, fog, and clouds, affect the quality of data collected by the LiDAR sensor. Weather like fog, snow, or rain can degrade the performance of the LiDAR sensor by 25% (Asvadi et al. 2018). Shamayleh and Khattak (2003) stated that rain droplets may reduce the signal intensity reflected from a target, and back-scatter from the rain may cause a false positive detection. Emitted laser beams may provide echoes that do not correspond to real obstacles by illuminating snowflakes (Yang et al. 2012). Also, it is often unreasonable or infeasible to shut down traffic operations on a section of a highway while collecting data. Moving or impermanent objects, such as vehicles or pedestrians, affect the data by introducing what may be unwanted noise for many users. Since slower driving speeds result in denser point clouds, it's challenging to use a mobile platform on higher-speed roadways while balancing operator safety and data density. As the technology continues to grow, it is expected that some of these challenges will be mitigated by increased reliability and robustness of the sensors, decreased cost of those sensors, and increased usability of processing software and hardware.

5.4 Data Size Implications

This study shows data sizes that one can expect from a single-beam LiDAR unit on a mobile platform in the highway environment. For the selected roads, the collector road segments have two lanes while the arterial road segments have four lanes. The raw data size generated from the sensor ranged from a little over 0.5 GB to approximately 5.6 GB. The data size mainly increases with the mode of the LiDAR unit and the duration of data collection. After processing the raw data and transforming into LAS file format, the size of the data shrinks significantly. This represents an excellent demonstration about the LiDAR data management for the transport agencies. The regression models developed to estimate large-scale data provide a great scope for transportation agencies or policymakers to consider when addressing the unique challenges and uses associated with the upcoming autonomous vehicle fleet data management. This model is important for transportation agencies since it will help to understand the data generated from the sensor and will give a baseline to comprehend probable data size.

Among the selected vehicle speeds, the 25, 30, 35, and 40 mph data were collected from collector road segments. The 45 and 55 mph data were collected from the arterial segment. The buffer zone widths of the arterial roads for the right-of-way and travel lanes were larger compared to the collector road, per Cache County (2013) design standards. The buffer zone has a relative impact over the point cloud data size collected under different speeds since the objects of each side of the buffer zone varied. In terms of the data distribution, it was found that the right-of-way section of the roadway contained approximately 75–86% of the collected data for the collector and arterial roadways with centerline trajectory. When considering the buffers from the centerline trajectory, it can be noted that as the speed of the vehicle increased, the data size within the right-of-way area decreased slightly. For the 35 miles/hr, data size remained consistent for all the LiDAR modes. The data size in the travel lanes area increased with the vehicle speed for the centerline trajectory. Data sizes within the travel lanes section for all the modes in the 25 and 30 miles/hr runs were consistent. One possible reason for this may be that the highway environment might have remained the same while collecting data for these two different speeds. While considering the actual trajectory, out of the total data size approximately 40–94% of the total data size had fallen in the right of the way section. This greater level of variation might be a result of the asymmetric distribution of the objects falling within both sides of the LiDAR. For the travel lanes, the opposite scenario occurs, with the data size mostly increasing with the vehicle speed. One reason may be that the higher speed roads have more travel lanes compared to the lower speed roads. One can expect a data size of 20–65% from the travel lanes with actual trajectory.

However, for the actual trajectory, the data sizes for the 25 and 30 miles/hr in terms of the right of way are nearly the same. Yet there was a significant increase from 30 to 35 miles/hr in this case. There is an asymmetry in the data size for the 40 miles/hr in terms of the 1024X20 and the 2048X10 mode. The highest data size can be observed for the 35 miles/hr. The data for the right of way always increased with the speed except for the 40 miles. One reason can be because of the number of features available within the right of way section during that particular run. One possible option to verify the data size for the selected vehicle speed and the LiDAR unit would be to run multiple times and also to compare the data size with different LiDAR units. This may help to understand and verify the data size generated from this study.

6. CONCLUSIONS

6.1 Overview

This demonstrated the possible data size that transportation agencies can expect from a mobile LiDAR platform in an active highway environment. To make conservative projections about LiDAR data size certain variables like LiDAR mode, vehicle speed, duration of the data collection period was considered in this study. Apart from this, the remaining objectives of this work were to summarize the existing applications of this technology in the transportation field, different challenges, and possible future research to explore.

The first chapter of this study provided a basic introduction on how LiDAR works and other sensors including INS and GPS. It also covered the scope of this project and expected outcome with proper justification. Chapter 2 introduced the uses of LiDAR data for different transportation applications while highlighting algorithms used to extract information. Chapter 2 also covered the common limitations of the existing tools and techniques used for related applications. Chapter 3 provided a technical overview of the equipment and methodology used for data collection, including data processing and storage in the selected file format.

Chapter 4 discussed the development of the regression model to understand the large-scale data size for selected scenarios. Chapter 5 discussed the mobile LiDAR uses in transportation in a broader term and challenges associated with the data collection, processing, and storage. Implications of the projected data size for the transportation industry were also discussed.

6.2 Key Findings

Overall, it can be said that the size of the LiDAR data mainly depends on the duration of the runtime and the selected mode to collect the data. The following are the key findings of this research:

- The size of the data decreases significantly compared to the raw data after saving it in the ASPRS recommended (.las) file format. Processing, filtering, and conversion to (.las) file format shrunk the data size from over a gigabyte down to a few megabytes.
- The weather has a possible impact on the data size. Weather like rain, snow, and sunlight affect the laser beam reflected from the LiDAR sensor while the beam returns to the sensor after heating the target object. This is due to the reflective property of the light.
- In an active highway environment, the data size will vary, particularly in the right of ways and the travel lanes.
- For the selected roadway types (arterial and collector), the right-of-way and the travel lanes sections contain approximately 75%–86% and 25%–65%, respectively, of the collected data for the centerline trajectory.
- For actual vehicle trajectory, there was greater variability in the data size right of the way section. This is most likely due to asymmetric overlap of buffer zone with the roadway.
- In general, as vehicle speed increases, the data density, or data size per mile, will decrease. However, due to the limited data size within the travel lanes, the relationship between vehicle speed and data size per mile is not statistically significant in this study.

6.3 Future Research

LiDAR technology has been in continuous development for the past few decades. Constant technological progress combined with substantial reductions in data acquisition cost has made it possible to apply this technology more frequently and at a larger scale. Although several studies have explored the use of LiDAR for transportation applications, there are still opportunities for further research.

LiDAR can be used for surveillance and monitoring because of its high accuracy, resolution, and dense scanning capacity. However, in order to avoid occlusion in a dense traffic environment, multiple runs are necessary in potential future studies in this area.

Currently, highway asset inventory with LiDAR depends on the reflectivity and shape of the target objects. Future research may consider the detection, identification, and evaluation of damaged, old, or non-standard assets. Similarly, Gargoum and El Basyouny (2019) stated that there is a lack of research in the extraction of cross-sectional road elements and their attributes, such as vertical and lateral clearances. There is also an opportunity to explore the performance of LiDAR to detect objects in areas covered with vegetation. Previous studies on lane marking, curb, or pavement edge detection mainly focused on short and straight segments. Future research should explore its implication for network-level data collection. This technology could also be used as a part of the quality control process and as-built assessment for construction.

Automated extraction of features from LiDAR data to minimize the analysis time should be emphasized, since object detection is important for the development of a smart transportation system. Algorithm standardization could improve the differentiation between ground and non-ground objects in raw point clouds. In addition, standardization and availability of non-proprietary processing methods would ease the usability of the point cloud data. A potential study could compare the data quality and processing times of proprietary vs. freely available data processing platforms. Improvements in work zone layout, management, and quality control can be made by using 3D point cloud data given the amount of information that can be driven from this 3D data.

An important aspect for future research is to compare the performance of LiDAR with other sensing technologies. At present, there are various LiDAR sensors with a large variety of specifications and price points. A potential study could assess the capabilities of these sensors in terms of detecting distanced objects from a mobile platform. Additionally, comparisons should be made between solid state and spinning mirror LiDAR sensors. Considering the nature of the data acquired by this technology, there is potential for it to be used to create models for structures. The impact of different environmental conditions, such as lighting, precipitation, air quality, etc. on LiDAR data accuracy should also be assessed to understand the performance of this sensing technology under non-ideal conditions, particularly in the context of autonomous vehicle performance.

To reduce the severity and frequency of crashes, there is ongoing research and development of the use of LiDAR and other sensing technologies in AVs and CAVs. Unfortunately, much of this research is theoretical or proprietary. LiDAR has the potential to expedite traffic incident management and safety analysis if the technology can be simplified for use by individuals with limited training and if processing times can be reduced. With the 3D scanning capacity of the sensor, it is possible to detect moving objects. Therefore, it is possible to explore the detection of pedestrians, bicyclists, and other vehicles. LiDAR sensors could be integrated with other vehicle telematics technology to better understand traffic flow and safety implications of advanced vehicle technologies. For example, if upstream congestion could be detected and communicated, vehicle and/or driver responses and performance may be assessed in greater detail than ever.

6.4 Limitations and Challenges of this Study

One of the major limitations of this study was that only one run was performed to estimate the data size from each of the selected roads, speeds, and LiDAR modes. Since data were collected from the mobile platform, it is highly likely that the data size would vary from run to run for the same road segment. Additionally, the commercial GPS used for this study was not precise, leading to some positional distortion in the features surrounding the vehicle in the point cloud. It was challenging to calibrate and sync the timestamp of all the sensors correctly to extract the features surrounding the LiDAR with respect to the actual vehicle location. In particular, while collecting data during the winter, snowfall further impacted the accuracy of the GPS. It is expected that many features that were actively present in the study area might not have been captured by the sensor due to its configuration and the speed of the vehicle. Another challenge involved with collecting and storing real time LiDAR data was the configuration and storage capacity of the computer used for the study. Given the scope of this study, it was not possible to compare the data size generated from the selected LiDAR unit with other units.

This study used only one single beam LiDAR sensor unit and the lowest speed while collecting data was 25 miles/hr. Even at this speed, it was quite difficult to capture the features in a detailed manner. As the speed of the vehicle increased, the data density greatly decreased, which increased the difficulty of recognizing objects in the point cloud. However, cost limitations inhibited ability of the researchers to utilize more and high-quality LiDAR units that would be expected to be used on AVs. Another major challenge and limitation of this study was maintenance of a constant speed during the data collection period since it was an active highway. Constant speed was necessary for ease in data processing and analysis but is not representative of real-word vehicle/driver performance. In some runs, there was some data distortion for inconsistent vehicle speed.

REFERENCES

Asvadi, A., Garrote, L., Premebida, C., Peixoto, P., & Nunes, U. J. (2018). Multimodal vehicle detection: Fusing 3D-LIDAR and color camera data. *Pattern Recognition Letters*, 115, 20-29.

B. Yang, Y. Liu, F. Liang, and Z. Dong, "Using Mobile Laser Scanning Data for Features Extraction of High Accuracy Driving Maps," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 41, 2016.

B. Yang, L. Fang, Q. Li, and J. Li, "Automated extraction of road markings from mobile LiDAR point clouds," *Photogrammetric Engineering & Remote Sensing*, vol. 78, no. 4, pp. 331-338, 2012.

Cache County, (2013, October 22). Manual of Roadway Design & Construction Standards. Retrieved July 27, 2020, from https://www.google.com/url?sa=t%2C

C. Ai and Y. Tsai, "Automated Sidewalk Assessment Method for Americans with Disabilities Act Compliance Using Three-Dimensional Mobile Lidar," *Transportation Research Record*, vol. 2542, no. 1, pp. 25-32, 2016.

C. Ordóñez, C. Cabo, and E. Sanz-Ablanedo, "Automatic detection and classification of pole-like objects for urban cartography using mobile laser scanning data," *Sensors*, vol. 17, no. 7, p. 1465, 2017.

Cremean, L., & Murray, R. (n.d.). Model-based estimation of off-highway road geometry using single-axis LADAR and inertial sensing. *Proceedings 2006 IEEE International Conference on Robotics and Automation*, 2006. ICRA 2006, 1661-1666. doi:10.1109/robot.2006.1641945

Đapo, A., Babić, L., & Pribičević, B. (2011). Application of a 3D terrestrial laser scanner in a survey of a railway bridge, Sava Jakuševac. Paper presented at the *5th International Conference on Engineering Surveying*.

Dawkins, J. J. (2014). Model based off-road terrain profile estimation. *2014 American Control Conference*, 2792-2797. doi:10.1109/acc.2014.6859189

D. Barber, J. Mills, and S. Smith-Voysey, "Geometric validation of a ground-based mobile laser scanning system," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 1, pp. 128-141, 2008.

De Santos Berbel, C., Anta, J. A., Castro, M., & Iglesias, L. (2008). Sight distance for road safety analysis using GIS.

F. Ghallabi, F. Nashashibi, G. El-Haj-Shhade, and M.-A. Mittet, "LIDAR-Based Lane Marking Detection for Vehicle Positioning in an HD Map," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018: IEEE, pp. 2209-2214.

F. Nashashibi and A. Bargeton. (2008). Laser-based vehicles tracking and classification using occlusion reasoning and confidence estimation. *2008 IEEE Intelligent Vehicles Symposium*, Eindhoven, pp. 847-852.

Gargoum, S., & El-Basyouny, K. (2017). Automated extraction of road features using LiDAR data: A review of LiDAR applications in transportation. *2017 4th International Conference on Transportation Information and Safety (ICTIS)*, 563-574. doi:10.1109/ictis.2017.8047822

Gargoum, S., Karsten, L., & El-Basyouny, K. (2018). Network Level Clearance Assessment using LiDAR to Improve the Reliability and Efficiency of Issuing Over-Height Permits on Highways. *Transportation Research Record: Journal of the Transportation Research Board*, 036119811875868. doi:10.1177/0361198118758687

Gargoum, S. A., & El Basyouny, K. (2019). A literature synthesis of LiDAR applications in transportation: feature extraction and geometric assessments of highways. *GIScience, Remote Sensing*, 1-30.

Guan, H., Li, J., Zhong, L., Yongtao, Y., & Chapman, M. (2013). Process virtualization of largescale lidar data in a cloud computing environment. *Computers & Geosciences*, 60, 109-116.

Guan, H., Li, J., Yu, Y., Chapman, M., Wang, H., Wang, C., & Zhai, R. (2015). Iterative Tensor Voting for Pavement Crack Extraction Using Mobile Laser Scanning Data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3), 1527-1537. doi:10.1109/tgrs.2014.2344714

Hernández, J., & Marcotegui, B. (2009). Filtering of artifacts and pavement segmentation from mobile lidar data. Paper presented at the ISPRS Workshop Laser scanning 2009.

H. Guan, J. Li, Y. Yu, Z. Ji, and C. Wang, "Using mobile LiDAR data for rapidly updating road markings," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 24572466, 2015.

H. Guan, J. Li, S. Cao, and Y. Yu, "Use of mobile LiDAR in road information inventory: A review," *International Journal of Image and Data Fusion*, vol. 7, no. 3, pp. 219-242, 2016.

H. Guan et al., "Iterative tensor voting for pavement crack extraction using mobile laser scanning data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1527-1537, 2014.

H. Wang et al., "Automatic road extraction from mobile laser scanning data," in *2012 International Conference on Computer Vision in Remote Sensing*, 2012: IEEE, pp. 136-139.

H. Zheng, R. Wang, and S. Xu, "Recognizing street lighting poles from mobile LiDAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 407-420, 2016.

I. Puente, H. González-Jorge, B. Riveiro, and P. Arias, "Accuracy verification of the Lynx Mobile Mapper system," *Optics & Laser Technology*, vol. 45, pp. 578-586, 2013.

I. Tomljenovic and A. Rousell, "Influence of Point Cloud Density on the Results of Automated Object-Based building Extraction from ALS Data," 2014.

Jalayer, M., J. Gong, H. Zhou, and M. Grinter. (2015). Evaluation of remote sensing technologies for collecting roadside feature data to support highway safety manual implementation. J*ournal of Transportation Safety*, 7(4): 345–357.

J. Hernández and B. Marcotegui, "Filtering of artifacts and pavement segmentation from mobile lidar data," 2009.

J. Jeong and I. Lee, "Classification of LIDAR Data for Generating a High-Precision Roadway Map," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, vol. 3, pp. 251-254, 2016.

J. Landa and D. Prochazka, "Automatic road inventory using LiDAR," *Procedia Economics and Finance*, vol. 12, pp. 363-370, 2014.

Kemeny J, Turner K (2008) Ground based LIDAR. Rock slope Mapping and assessment. Technical report of the Central Federal Lands Highway Division U.S. Department of Transportation FHWA-CFL/TD-08-006 (www.iaeg.info)

Křemen, T., Štroner, M., & Třasák, P. (2014). Determination of Pavement Elevations by the 3D Scanning System and Its Verification. *Geoinformatics FCE CTU*, 12(0), 55-60. doi:10.14311/gi.12.9

Kumar, P., & Angelats, E. (2017). An Automated Road Roughness Detection from Mobile Laser Scanning Data. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W1, 91-96. doi:10.5194/isprs-archives-xlii-1-w1-91-2017

Lato, M., Hutchinson, J., Diederichs, M., Ball, D., & Harrap, R. (2009). Engineering monitoring of rockfall hazards along transportation corridors: using mobile terrestrial LiDAR. *Natural Hazards and Earth System Sciences*, 9(3), 935-946. doi:10.5194/nhess-9-935-2009

L. B. Cremean and R. M. Murray, "Model-based estimation of off-highway road geometry using single-axis ladar and inertial sensing," in Proceedings *2006 IEEE International Conference on Robotics and Automation*, 2006. ICRA 2006., 2006: IEEE, pp. 1661-1666.

Li, Y. H., Shinohara, T., Satoh, T., & Tachibana, K. (2016). Road signs detection and recognition utilizing images and 3d point cloud acquired by mobile mapping system. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1, 669-673. doi:10.5194/isprsarchives-xli-b1-669-2016

Li, Y., Hu, X., Guan, H., & Liu, P. (2016). An Efficient Method for Automatic Road Extraction Based on Multiple Features from LIDAR Data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41.

Liu Y., Monteiro S. T. and Saber E., 2016. Vehicle detection from aerial color imagery and airborne lidar data. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Prague, Czech, pp. 1384-1387.

Lovas T., Toth C. K. and Barsi A. (2004). Model-based vehicle detection from LiDAR data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Istanbul, Turkey, B4, pp. 134-138.

M. Castro, L. Iglesias, J. A. Sánchez, and L. Ambrosio, "Sight distance analysis of highways using GIS tools," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 9971005, 2011.

Mekker, M. M., Lin, Y., Elbahnasawy, M. K., Shamseldin, T. S., Li, H., Habib, A. F., & Bullock, D. M. (2018). Application of LiDAR and Connected Vehicle Data to Evaluate the Impact of Work Zone Geometry on Freeway Traffic Operations. *Transportation Research Record: Journal of the Transportation Research Board*, 036119811875805. doi:10.1177/0361198118758050

M. Kulawiak, M. Kulawiak, and Z. Lubniewski, "Integration, Processing and Dissemination of LiDAR Data in a 3D Web-GIS," *ISPRS International Journal of Geo-Information*, vol. 8, no. 3, p. 144, 2019.

Montgomery, W. D., Mudge, R., Groshen, E. L., Helper, S., MacDuffie, J. P., & Carson, C. (2018). America's workforce and the self-driving future: Realizing productivity gains and spurring economic growth.

M. Tupas, S. Lat, and R. Magturo, "Data Archiving and Distribution of LiDar and Derived Datasets in the Philippines," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 41, 2016.

Nashashibi, F., & Bargeton, A. (2008). Laser-based vehicles tracking and classification using occlusion reasoning and confidence estimation. *2008 IEEE Intelligent Vehicles Symposium*, 1-6. doi:10.1109/ivs.2008.4621244

NHTSA, U. (2016). Department of Transportation, National Highway Traffic Safety Administration. In.

Olsen, M. J., G. Roe, C. Glennie, F. Persi, M. Reedy, D. Hurwitz, K. Williams, H. Tuss, A. Squellati, and M. Knodler. 2013. "NCHRP 15-44 Guidelines for the Use of Mobile LiDAR in Transportation Applications, 2013."

Patlins, A., Kunicina, N., Zhiravecka, A., & Shukaeva, S. (2010). LIDAR sensing technology using in transport systems for tram motion control. *Elektronika ir elektrotechnika*, 101(5), 13-16.

P. Kumar and E. Angelats, "An Automated Road Roughness Detection from Mobile Laser Scanning Data," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, 2017.

P. Rodríguez-Gonzálvez, B. Jiménez Fernández-Palacios, Á. Muñoz-Nieto, P. Arias-Sanchez, and D. Gonzalez-Aguilera, "Mobile LiDAR system: New possibilities for the documentation and dissemination of large cultural heritage sites," *Remote Sensing*, vol. 9, no. 3, p. 189, 2017.

Q. Chen, "Airborne lidar data processing and information extraction," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 2, p. 109, 2007.

Sairam, N., Nagarajan, S., & Ornitz, S. (2016). Development of mobile mapping system for 3D road asset inventory. Sensors, 16(3), 367.

Santos-Berbel, C. D., Castro, M., Medina, S. L., & Paréns-González, M. (2014). Sight Distance Studies on Roads: Influence of Digital Elevation Models and Roadside Elements. *Procedia - Social and Behavioral Sciences*, 160, 449-458. doi: 10.1016/j.sbspro.2014.12.157

Shamayleh, H., & Khattak, A. (2003). Utilization of LiDAR technology for highway inventory. *Proceedings of the 2003 Mid-Continent Transportation Research Symposium*, Ames, IA.

S. Gargoum, K. El-Basyouny, J. Sabbagh, and K. Froese, "Automated highway sign extraction using LiDAR data," *Transportation Research Record*, vol. 2643, no. 1, pp. 1-8, 2017.

S. Gargoum, L. Karsten, and K. El-Basyouny, "Network level clearance assessment using LiDAR to improve the reliability and efficiency of issuing over-height permits on highways," *Transportation Research Record*, vol. 2672, no. 42, pp. 45-56, 2018.

Soilán, M., Riveiro, B., Martínez-Sánchez, J., & Arias, P. (2016). Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 92-101. doi: 10.1016/j.isprsjprs.2016.01.019

S. Pu, M. Rutzinger, G. Vosselman, and S. O. Elberink, "Recognizing basic structures from mobile laser scanning data for road inventory studies," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. S28-S39, 2011.

Toth, C. K., Grejner-Brzezinska, D. A., & Moafipoor, S. (2004). Precise vehicle topology and road surface modeling derived from airborne LIDAR data. Paper presented at the 60th Annual Meeting of ION.

T. Ogawa and K. Takagi, "Lane Recognition Using On-vehicle LiDar," in *2006 IEEE Intelligent Vehicles Symposium*, 2006: IEEE, pp. 540-545.

Tsai, Y. (., Ai, C., Wang, Z., & Pitts, E. (2013). Mobile Cross-Slope Measurement Method Using Lidar Technology. *Transportation Research Record: Journal of the Transportation Research Board*, 2367(1), 53-59. doi:10.3141/2367-06

V.-H. Cao, K. Chu, N.-A. Le-Khac, M. T. Kechadi, D. Laefer, and L. Truong-Hong, "Toward a new approach for massive LiDAR data processing," in *2015 2nd IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM)*, 2015: IEEE, pp. 135140

Wen, C., Li, J., Luo, H., Yu, Y., Cai, Z., Wang, H., & Wang, C. (2016). Spatial-Related Traffic Sign Inspection for Inventory Purposes Using Mobile Laser Scanning Data. *IEEE Transactions on Intelligent Transportation Systems*, 17(1), 27-37. doi:10.1109/tits.2015.2418214

W. Yao, S. Hinz, and U. Stilla, "Traffic monitoring from airborne LIDAR–Feasibility, simulation and analysis," in XXI Congress, proceedings. International archives of photogrammetry, remote sensing and spatial geoinformation sciences, Beijing, China, 2008, vol. 37, p. B3B.

W. Ouyang and B. Xu, "Pavement cracking measurements using 3D laser-scan images," *Measurement Science and Technology*, vol. 24, no. 10, p. 105204, 2013.

Wu, S., Wen, C., Luo, H., Chen, Y., Wang, C., & Li, J. (2015). Using mobile LiDAR point clouds for traffic sign detection and sign visibility estimation. 2015 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 565-568. doi:10.1109/igarss.2015.7325826

X. Chen, B. Kohlmeyer, M. Stroila, N. Alwar, R. Wang, and J. Bach, "Next generation map making: geo-referenced ground-level LIDAR point clouds for automatic retro-reflective road feature extraction," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2009*: ACM, pp. 488-491.

Yao, W., Hinz, S., & Stilla, U. (2008). Traffic monitoring from airborne LIDAR–Feasibility, Simulation and Analysis. Paper presented at the XXI Congress, Proceedings. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, Beijing, China.

Y.-C. J. Tsai and F. Li, “Critical assessment of detecting asphalt pavement cracks under different lighting and low intensity contrast conditions using emerging 3D laser technology,” *Journal of Transportation Engineering*, vol. 138, no. 5, pp. 649-656, 2012.

Y. Li, T. Shinohara, T. Satoha, and K. Tachibana, “Road Signs Detection and Recognition Utilizing Images and 3d Point Cloud Acquired by Mobile Mapping System,” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 1, 2016.

Y. Tsai, C. Ai, Z. Wang, and E. Pitts, “Mobile cross-slope measurement method using lidar technology,” *Transportation Research Record*, vol. 2367, no. 1, pp. 53-59, 2013.

Zhang, T., Vosselman, G., & Elberink, S. J. (2017). Vehicle recognition in aerial LiDAR point cloud based on dynamic time warping. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W4, 193-198.