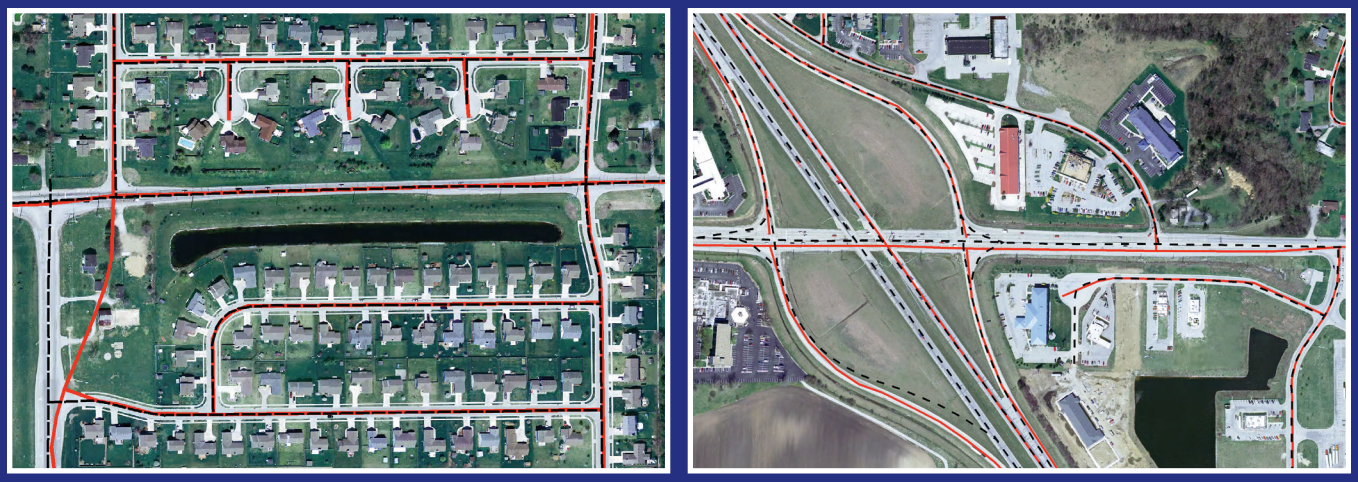
JOINT TRANSPORTATION
RESEARCH PROGRAM
INDIANA DEPARTMENT OF TRANSPORTATION
AND PURDUE UNIVERSITY

Performance Measure That Indicates Geometry Sufficiency of State Highways

Volume II—Clear Zones and Cross-Section Information Extraction



Jie Shan, Serkan Ural

SPR-3640 • Report Number: FHWA/IN/JTRP-2015/07 • DOI: 10.5703/1288284315529

RECOMMENDED CITATION

Shan, J., & Ural, S. (2015). *Performance measure that indicates geometry sufficiency of state highways: Volume II—Clear zones and cross-section information extraction* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2015/07). West Lafayette, IN: Purdue University. http://dx.doi.org/10.5703/1288284315529

AUTHORS

Jie Shan, PhD
Professor of Civil Engineering
Lyles School of Civil Engineering
Purdue University
(765) 494-2168
jshan@purdue.edu
Corresponding Author

Serkan Ural
Graduate Research Assistant
Lyles School of Civil Engineering
Purdue University

ACKNOWLEDGMENTS

The authors would like to thank the Study Advisory Committee members, John Weaver, Dwayne Harris, Kumar Dave, Bill Flora, and Gregory Katter, who helped throughout the entire project with their invaluable input and comments.

JOINT TRANSPORTATION RESEARCH PROGRAM

The Joint Transportation Research Program serves as a vehicle for INDOT collaboration with higher education institutions and industry in Indiana to facilitate innovation that results in continuous improvement in the planning, design, construction, operation, management and economic efficiency of the Indiana transportation infrastructure. https://engineering.purdue.edu/JTRP/index_html

Published reports of the Joint Transportation Research Program are available at: http://docs.lib.purdue.edu/jtrp/

NOTICE

COPYRIGHT

Print ISBN: 978-1-62260-338-1
ePUB ISBN: 978-1-62260-339-8

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/IN/JTRP-2015/07		

4. Title and Subtitle	5. Report Date
Performance Assessment Measure that Indicates Geometry Sufficiency of State Highways: Volume II—Clear Zones and Cross-Section Information Extraction	March 2015
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Jie Shan and Serkan Ural	FHWA/IN/JTRP-2015/07

9. Performing Organization Name and Address	10. Work Unit No.
Joint Transportation Research Program Purdue University 550 Stadium Mall Drive West Lafayette, IN 47907-2051	
	11. Contract or Grant No. SPR-3640

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes
Prepared in cooperation with the Indiana Department of Transportation and Federal Highway Administration.

16. Abstract

Evaluation method employed for the proposed corridor projects by Indiana Department of Transportation (INDOT) consider road geometry improvements by a generalized categorization. A new method which considers the change in geometry improvements requires additional information regarding cross section elements. Part of this information is readily available but some information like the embankment slopes and obstructions near traveled way needs to be acquired. This study investigates available data sources and methods to obtain cross-section and clear zone information in a feasible way for this purpose. We have employed color infrared (CIR) orthophotos, LiDAR point clouds, digital elevation and surface models for the extraction of the paved surface, average grade, embankment slopes, and obstructions near the traveled way like trees and man-made structures. We propose a framework which first performs a support vector machine (SVM) classification of the paved surface, then determines the medial axis and reconstructs the paved surface. Once the paved surface is obtained, the clear zones are defined and the features within the clear zones are extracted by the classification of LiDAR point clouds.

SVM classification of the paved surface from CIR orthophotos in the study area results with a classification accuracy over 90% which suggests the suitability of high resolution CIR images for the classification of paved surface via SVM. A total of 21.3 miles of relevant road network has been extracted. This corresponds to approximately 90% of the actual road network due to missing parts in the paved surface classification results and parts which were removed during cleaning, simplification and generalization process. Branches due to connecting driveways, adjacent parking lots, etc., were also extracted together with the main road alignment as by-product. This information may also be utilized if found necessary with further effort to filter out irrelevant pieces that do not correspond to any actual branches. Based on the extracted centerline and classification results, we have estimated the paved surface as observed on the orthophotos. Based on the estimated paved surface centerline and width, we have generated cross section lines and calculated the side slopes. We have extracted the buildings and trees within the clear-zones that are also defined based on the reconstruction of the paved surface. Among 86 objects detected as buildings, 14% were false positives due to confusion with bridges or trees which present planar structure.

17. Key Words	18. Distribution Statement
road extraction, LiDAR Point Cloud classification, building extraction	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	31	

Form DOT F 1700.7 (8-69)

EXECUTIVE SUMMARY

PERFORMANCE MEASURE THAT INDICATES GEOMETRY SUFFICIENCY OF STATE HIGHWAYS: VOLUME II—CLEAR ZONES AND CROSS-SECTION INFORMATION EXTRACTION

Introduction

Current evaluation methods of the Indiana Department of Transportation's (INDOT's) Roadway Asset Team for proposed corridor projects account for road geometry improvements by a much generalized categorization. A new method that considers the change in geometry improvements in a more detailed fashion is presented in Volume I of this report.

As part of this study, we investigated feasible ways of extracting additional elements from available datasets to support asset management in Indiana. To test the extent to which cross-section and clear zone information may be extracted in a reliable and practical way, we examined available data sources for their suitability. We combined the use of orthophotos, LiDAR point clouds, digital elevation, and surface models to identify remote sensing methods that are capable of extracting the required features efficiently and proposed a framework for determining the paved surface, average grade, and embankment slopes, extracting the obstructions near the traveled way such as trees and man-made structures.

Findings

Existing road network datasets, which we initially considered using as the reference for extracting information on road geometry and roadside features, did not have the required spatial accuracy to define the road extent and the clear zones. The datasets we recommend are orthophotos and the LiDAR datasets that are acquired in the scope of a statewide program, which are expected to be updated on a regular basis if not continuously.

Based on the estimated paved surface centerline and width, we were able to generate cross-section lines and calculate the slopes along these lines. Interpreting the slope values requires attention if a back slope is expected to start before two samples of elevation values are available on the foreslope. This situation is related to the slope calculations of very rapid falls followed by flat land. Slope calculations may be affected by the quickly changing slope values due to averaging.

The thinning algorithm does not result in intersections that coincide at a single node. If the intersections must be conserved topologically, an effort is required to edit them. However, this does not affect the cross-section information since the road extent and clear zones are defined based on the extracted paved surface extent. Similar to the options for whether to complete the missing parts of the alignment, intersections may or may not need to be topologically conserved as a single node. The decision in this case is whether the alignment extracted from aerial imagery will be required to serve as a replacement for a complete network in which the intersections are subject to further analysis, not just simply in aiding the estimation of the approximate paved surface width and reconstructing the surface extent.

The LiDAR dataset provides the capability to extract features within the clear zones at the roadside in a limited fashion, helping to extract the buildings and trees within the clear zones. It is not possible to extract vertically aligned features such as fences, walls, posts, signs, and so forth due to the limitations of the acquisition technique. Mobile or terrestrial LiDAR acquisition would be more suitable for extracting such features. Also, extracting individual trees out of a group of trees that are close to each other was not possible using the available dataset. In cases where there is prior knowledge of tree types in the area, individual trees may be estimated based on assumptions regarding the type and structure of those tree types.

Implementation

We have employed and pipelined different software as tools for the implementation of the proposed framework in the study area. Our main concern has been the availability of the algorithms that we have proposed and their convenience for performing quick tests with the existing datasets. Hence, we have not aimed at establishing an operational-level software integration. In summary, we have employed ESRI ArcGIS software for the preprocessing of the orthophotos and DEMs, LAStools for the preprocessing of LiDAR point clouds, Monteverdi interface of Orfeo Toolbox for the classification of the paved road surface from the orthophotos, ArcGIS for the vector-based cleaning and generalization of the classification results, Matlab and Wolfram Mathematica for the morphological cleaning, thinning, pruning, boundary extraction, and generalization, ArcGIS for the reconstruction of the paved road surface as well as generating the cross-section lines using Python scripting and also calculating the slopes, LAStools for LiDAR ground filtering and point cloud classification, and R programming language libraries for the delineation of classified points. This pipelining of software is not the only available and certainly not the most efficient way of performing the tasks of the proposed framework. The methods and algorithms mentioned in their respective sections of the report may be combined with a more efficient integration of software tools. Several tools that we preferred to use based on convenience may be eliminated and the tasks performed by these tools may be switched to other tools already used to perform other tasks. As an example, Matlab may be used for implementing SVM classification of the images instead of Orfeo Toolbox, as well as for handling all morphological operations. It will require some additional effort to combine these tasks under one tool that will work at an operational level.

As mentioned previously, the proposed framework is not a fully automated process. Human involvement is required at several steps, most importantly while collecting training samples for classification, which is performed once and then applied throughout the entire study area. Collected samples must represent all occurrences of the pavement types in the study area.

Other instances of human involvement include setting the parameters for determining the size of small components, simplification level, buffer sizes for removing the branches from the boundary raster or for calculating the pavement width, ground filtering, point cloud classification, and delineation of the classification results. Once these are determined adequately based on the properties of the study area, they may be used reliably.

Finally, human involvement may be required for completing discontinuities as necessary and for editing intersections if they are required to be topologically conserved as a single node.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Current evaluation methods of the Indiana Department of Transportation's (INDOT's) Roadway Asset Team for proposed corridor projects account for road geometry improvements by a much generalized categorization. A new method that considers the change in geometry improvements in a more detailed fashion may be proposed. Such a method would need to include the examination of all cross-section elements like lane width, median, shoulder, clear zones, etc. The Indiana inventory database stores some of these elements but some information like embankment slopes and height, ditch dimensions, and obstructions near the traveled way are not included in this database.

In this report, we investigate the available data sources which may be employed for extracting the required cross-section and clear zone information. We propose a framework for determining the paved surface, average grade, embankment slopes, extracting the obstructions near the traveled way like trees and man-made structures.

We employ orthophotos, LiDAR point clouds, digital elevation and surface models and investigate remote sensing methods that are capable of extracting the required features in a feasible way.

This report is organized in four chapters. Chapter 2 which follows this introduction chapter explains available datasets and their suitability for the purposes of this study. Chapter 3 provides the details of our proposed framework for extracting clear zones and cross-section information using remote sensing techniques. Chapter 4 lays down our overall conclusion remarks on our proposed framework based on our observations and experience following its implementation.

2. AVAILABLE DATASETS

We have investigated the potential data sources which may be used in obtaining and/or extracting clear zones and cross-section information for Indiana road network. We initially focused on what data may be acquired that are already processed at a level which would be easily integrated into a GIS (Geographic Information System) environment to be further analyzed. Data sources that are already or potentially available state-wide were given priority. Existing vector-based data subject to investigation for availability, reliability, and effects on performance include the road networks, right of way (ROW), and property boundaries. Remote sensing datasets considered for analysis are real color (RGB) and color infrared (CIR) aerial orthophotograpy, digital elevation models (DEM) and digital surface models (DSM), and light detection and ranging (LiDAR) datasets. Each dataset will be discussed in further detail.

2.1 Existing Vector-Based Datasets

2.1.1 Road Network

Road network is available via INDOT and U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing System (TIGER). We have obtained the TIGER road dataset of 2011 for Tippecanoe County from U.S. Census Bureau. Two datasets comply with each other in majority. However, differences are also observed. We have observed discrepancies such as one part of the road not existing in one of the datasets, misalignment of the road lines representing the same road segments, and discrepancy in the level of detail. Figure 2.1 shows two examples of discrepancies between the two datasets. These discrepancies may be due to lack of up to date road lines, erroneous road line geometries, or the difference between the internally consistent structures of two datasets on how the road lines are recorded and maintained. The geometric accuracy of the road center lines is important for the acquisition of the features within road proximity since spatial accuracy is crucial for the integration of other datasets with the road information.

Considering the administration of INDOT over the road network available in their geographic information system (GIS) including in-house update and modification

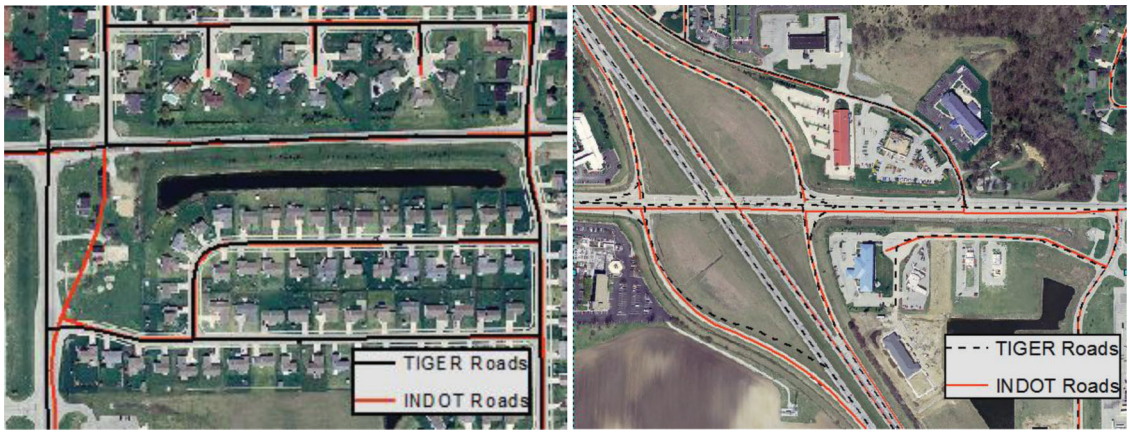


Figure 2.1 Discrepancies between INDOT and TIGER road datasets from part of the study area.

opportunities as well as the institutional reliance on the dataset, we have opted for performing our analysis based on the road network provided by INDOT. This option provides more confidence over the reliability of the data for continuous operational purposes than relying on external sources of road network which would require complying with other standards and being limited by their update and maintenance schedule. Also, determining existing inaccuracies in the road network is easier for this dataset which has been operational and part of institutional memory in contrast to datasets from other sources.

In the course of our study, we have noticed several issues about the INDOT road network which complicate the integration and analysis of the datasets for the extraction of clear zones and cross-section information. The first issue is the geometric accuracy of the road lines. The road network is frequently observed to be not aligned with the center of the paved road when referenced to orthophotos with reliable horizontal accuracy reported as their production quality. This may be observed in Figure 2.2.

One of the main objectives of this study is to determine the methods for extracting the ground features which are within the clear zone. In order to extract the ground features using remote sensing methods, one needs to explicitly define the clear zone and confine the area to perform the feature extraction. Simplest way of determining the area that the clear zone covers is to apply a clear-zone offset to the center of the paved area considering the width and the type of the road. Since the available road network does not provide such confidence in identifying the center of the paved road, additional effort is required to first determine the extent of the paved area.

The second issue is related to the divided roads. In case of divided roads, we have often observed that, there is only one single line aligned with that divided part of the road. We were not able to determine a representation standard such as the single line always being on one side towards a particular direction. Road lines not following a standard way of representing the divided roads also affects the identification of a clear zone since the clear zone will be different depending on which side of the road the road line is.

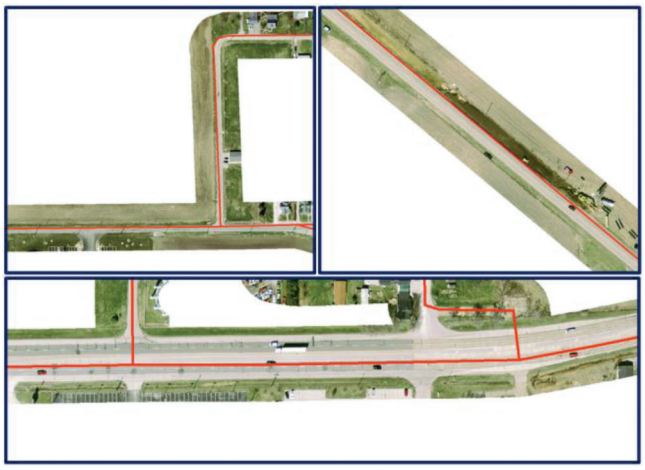


Figure 2.2 Examples of discrepancies between the INDOT road network layer and the orthophotos.

Either of the assumptions, that the road line is the actual center of the paved road, and that the single line represents the specific side of the divided road, are not valid. Hence, it would be unreliable to apply a predetermined clear zone buffer to all road lines based on these false assumptions. The extent of the paved road is to be determined first.

2.1.2 Building Footprints

Locally at county or city level, different departments collect, analyze and maintain data regarding various features. Tippecanoe County is one of the counties in Indiana maintaining a GIS with data acquired from different sources. They provided us the building footprints that from their GIS. The dataset includes the outlines of the buildings. However, the dataset is incomplete and not up to date. Completeness is of importance in order to be able to extract all existing ground features. At the time of our request, we have been informed that new building footprints were to be produced from November 2010/March 2011 orthophotos. However, such building footprint data are not necessarily available consistently for all Indiana.

2.1.3 Right of Way (ROW)

Information regarding the right of way (ROW) or the edge of pavement is also of significance for the determination of the strip in which the features are to be extracted. Tippecanoe County GIS provided the ROW information obtained from the County Highway Department. The dataset is far from complete since it is a work still in progress and only includes roads maintained by the county. County GIS office also has an edge of pavement GIS layer in their database in which some areas are not included. In addition, such datasets may not necessarily be available across the state similar to the case with the building outlines. Figure 2.3 shows a sample section from the ROW dataset together with INDOT road network.

2.1.4 Property Boundaries

Even though the legal aspect of land registration is beyond the scope of this study, we have investigated the land parcel data that are publicly available in order to evaluate a possible attempt to extract the right of way (ROW) for the roads. In case of consistent registration of the land parcels, one may establish a method to determine the ROW. The assumption in this case is that the parcels either registered as roads or the negative space left in between may be candidates for a continuous ROW layer. We relied on our observations on the available datasets before determining any

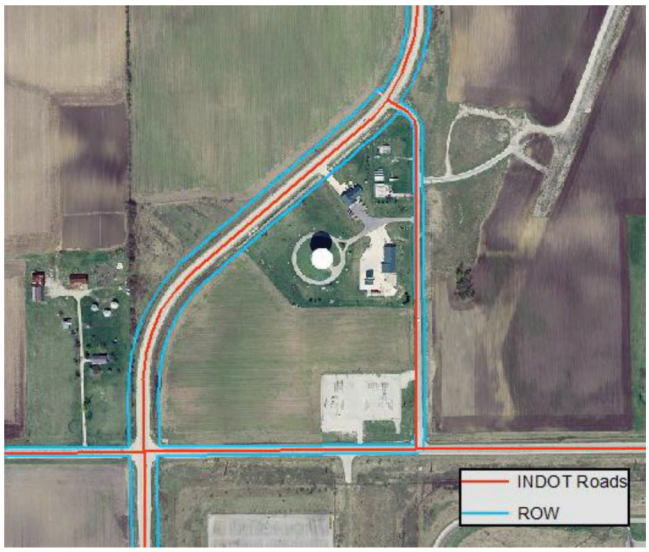


Figure 2.3 Sample from the ROW dataset of Tippecanoe County Highway Department acquired from Tippecanoe County GIS together with INDOT road network.

requirement for further investigation of land registration procedures.

The IndianaMap Data Sharing Initiative between Indiana Geospatial Information Council (IGIC), Indiana Department of Homeland Security (IDHS), Indiana Office of Technology (IOT), Indiana Geological Survey (IGS), and participating Indiana counties, provides a collection of land parcels maintained by county agencies as a map service. The dataset includes land parcels from 79 counties in Indiana and it is current through April 27, 2013 at the time of compilation of this report. We have observed that the available parcel database does not provide the required information and detail consistently throughout entire Indiana to allow the establishment of a ROW layer by compiling the information available in the dataset. Figure 2.4 shows part of the study area depicting the inconsistency of the property boundaries with the actual roads at one instance of an intersection of two roads. We have found no further value in pursuing a process to analyze parcel boundaries in an attempt to generate a ROW layer.

2.2 Remote Sensing Data for Ground Feature Extraction

Remote sensing is a powerful tool for extracting information on ground features. Remote sensing data may include aerial orthophotos, digital elevation models (DEM) and surface models (DSM) derived either from aerial photography or from LiDAR (Light Detection and Ranging) data, as well as the raw point clouds from LiDAR systems. Remote sensing is widely used for extraction of ground features which may include man-made or natural structures. There are numerous algorithms developed for remote sensing data processing to extract required information that are used in a wide range of applications. Theories, methods and algorithms from different disciplines including signal and image processing, pattern recognition, machine learning, computer vision, statistics, etc., are used in different applications which require to process remote sensing imagery. Remote sensing research area maintains a powerful set of mature tools and ongoing research to achieve better methods and algorithms.

In the last two decades, LiDAR is also increasingly utilized for remote sensing as an emerging technology. Laser scanners are active sensing systems using laser beams to measure the range between the target and the sensor. There are airborne, terrestrial and mobile platforms employed for LiDAR data acquisition with different system properties. These systems provide a collection of point location data referred as point clouds. These point clouds are then processed to extract information and derive other commonly used products like DEM and DSM. Ground features may also be acquired by directly processing the point clouds.

We have investigated the availability and adequacy of the remote sensing data for the purposes of the project including aerial imagery and elevation datasets. The Indiana Spatial Data Portal (ISDP) provides statewide datasets including aerial photos and elevation data.

Figure 2.4 Sample from the parcel boundaries dataset together with INDOT road network.

2.2.1 2010 National Agriculture Imagery Program (NAIP) Dataset

The latest statewide complete acquisition of aerial imagery on ISDP in the initial stage of the project was from the National Agriculture Imagery Program (NAIP) acquired in the summer of 2010. NAIP acquires orthoimagery during the agricultural season. The imagery contains four bands in the spectral ranges of red, green, blue, and near infrared with 1 m ground sampling distance (GSD). The reported horizontal accuracy for the ortho-rectified imagery with respect to true ground is 6 m. Two concerns arise regarding the use of this dataset for the purposes of the project. The first concern is related to the acquisition of the imagery in the summer, which is leaf-on season. It is convenient for NAIP to acquire imagery in that period of the year since the intention is to extract information regarding agriculture. However, this is not the case for our project since one of the objectives is to acquire the best information on manmade structures without the obstruction of high vegetation. Figure 2.5 presents an example of the occlusion caused by the trees over a building in the NAIP image. The difference in the spatial resolution of two datasets is also visible in the figure. Second concern is related to the horizontal accuracy of the ortho-rectified imagery. An accuracy of 6 m may result in extracting misleading information regarding the feature locations. We are not only concerned about extracting ground features. Their relative location with respect to the roads is also of importance. Hence, we believe that imagery with higher horizontal accuracy would be more suitable for the purposes of the project.

2.2.2 2005 Statewide Orthophotography Project Dataset

Most recent statewide ortho-imagery on ISDP other than NAIP datasets during the initial stage of the project was from 2005 Statewide Orthophotography Project coordinated by Indiana Geographic Information Council (IGIC) in partnership with the State of Indiana and local government participation. Color orthophotos are available with minimum 1-foot GSD statewide and with 6-inch resolution for 13 counties. Reported accuracy is less than 5′ for 1-foot data and less than 2.5′ for 6-inch data with 95% confidence level based on accuracy assessment complying with National Standard for Spatial Data Accuracy (NSSDA). As a more specific example, the accuracy reported for data collected over Tippecanoe County is 1.914′ with 95% confidence. 4-band color infrared (CIR) orthophotos are also provided statewide at 1 m GSD to meet 5 m horizontal accuracy at 95% confidence level. The data were collected in March and April 2005 at leaf-off conditions. Figure 2.6 shows samples of the three-band true color (RGB) and four-band CIR orthophotos from Tippecanoe County.

Apart from digital orthophotos, derived products are also available statewide as a result of 2005 Statewide Elevation Project. These products include Digital Surface Models (DSM) and Digital Elevation Models (DEM) extracted from the orthophotos of 2005 statewide Ortophotography Project using an auto-correlated pixel matching system. The raster DEM and DSM files are provided with 5′ GSD. The DEM files reportedly support the generation of orthophotos with accuracy better than 5′ for 1-foot data and less than 2.5′ for 6-inch data with 95% confidence level. Figure 2.7 shows samples of the DEM and normalized DSM (nDSM) calculated by differencing the DSM and the DEM from Tippecanoe County.

The orthophotos of 2005 Statewide Orthophotography project provide high resolution imagery with high accuracy. Resolution and accuracy at this level would allow the extraction of ground features from these orthophotos. Derived DEM and DSM products may also aid in feature extraction process. However, one main concern needs to be mentioned for the use of these datasets. Since the data acquisition is from the year 2005, any changes since 2005 would not be reflected in the feature extraction results.

2.2.3 2011–2013 Indiana Orthophotography (RGBI), LiDAR, and Elevation Project Dataset

With the availability of new orthophoto and LiDAR acquisitions planned as a part of Indiana Statewide and

Figure 2.5 Example of occlusion caused by trees over a building. The building is clearly identifiable in the true-color orthophoto with 6-inch spatial resolution from 2005 Statewide Orthophotography Project taken during leaf-off season (left) while it is mostly obstructed by trees in the 2010 NAIP image (right).

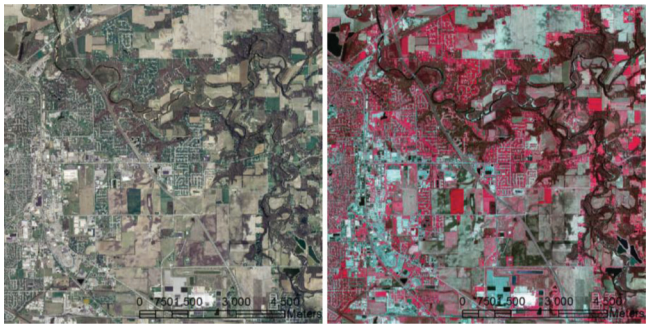


Figure 2.6 Samples of RGB (left), CIR (right) orthophotos from Tippecanoe County acquired in scope of 2005 Statewide Orthophotography Project.

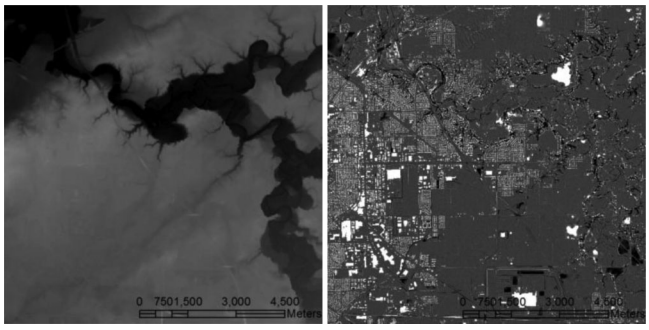


Figure 2.7 Samples of DEM (left), and nDSM (DSM-DEM) from Tippecanoe County acquired in scope of 2005 Statewide Orthophotography Project.

Imagery Program administered through Indiana Office of Technology – State Geographic Information Officer (www.igic.org), an up to date dataset of orthophotos, DEMs and LiDAR point clouds is available as of the time this report is compiled. The program was planned as a three year cycle, repeating program for the acquisition of base orthophoto and LiDAR products for the state of Indiana. Three acquisition areas were determined in the State of Indiana as center, east, and west for which the acquisitions have been carried out in 2011, 2012, and 2013 respectively.

The base product deliverables include four-band CIR orthophotos, LiDAR point clouds, and hydro-flattened DEM obtained from the LiDAR data. Orthophotos are available on ISDP at 1-foot, or 6-inch resolution. U.S. Geological Survey (USGS) compliant LiDAR data are available at either 1.0 or 1.5 m nominal pulse spacing (NPS). DEMs obtained from the LiDAR data with hydro-flattening process are of 5′ spatial resolution. Figure 2.8 shows the spatial resolution of orthophotos and NPS of LiDAR datasets by county.

One of the main advantages of this dataset over the 2005 Statewide Orthophotography Project Dataset is that the infrared (IR) band is also available at the same resolution of 1-foot as the RGB orthophotos rather than 1 m GSD. IR band increases the ability to distinguish vegetation in comparison with using the RGB bands only. Figure 2.9 shows part of the dataset with RGB and NIR-R-G band combinations at the same spatial resolution. Sample from the LiDAR point cloud dataset is presented in Figure 2.10.

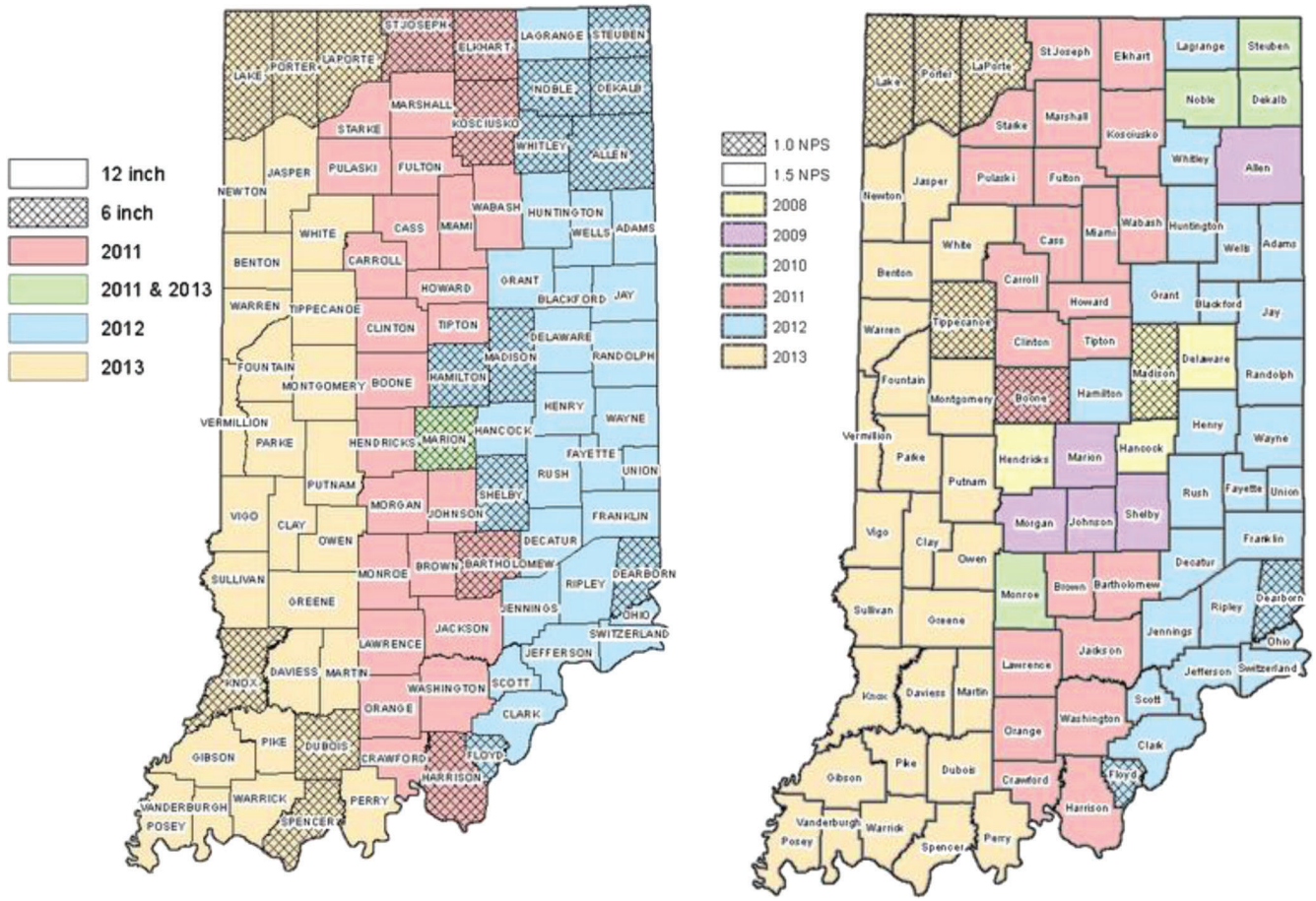


Figure 2.8 Left: 2011–2013 IndianaMap Orthophotograpy source date and spatial resolution by county; Right: 2011–2013 IndianaMap LiDAR source dates and NPS. (Source: http://gis.iu.edu/files/images/dataset/2011-2013_rgbi_date_resolution.jpg.)

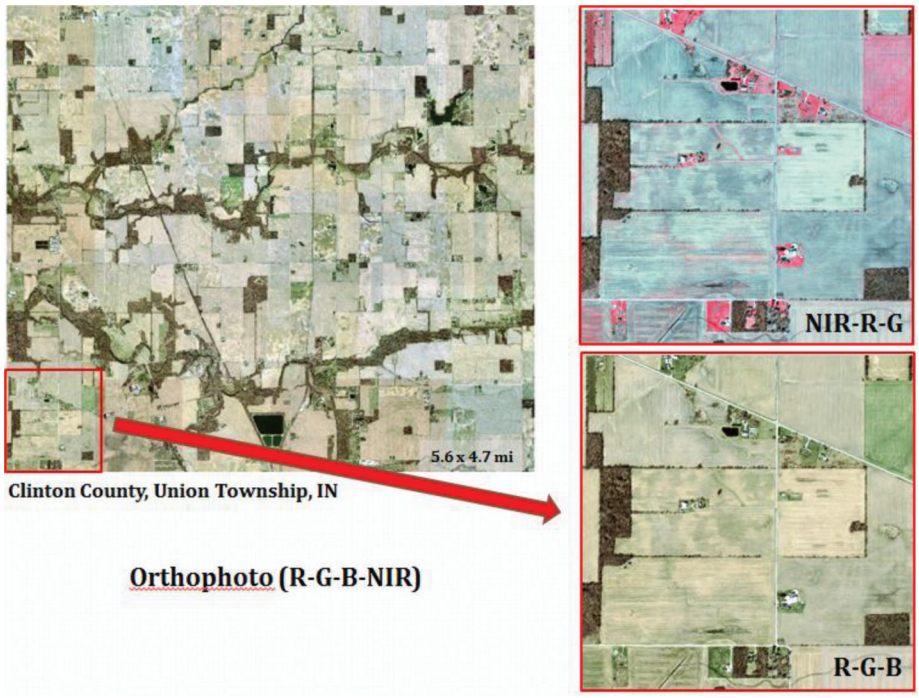


Figure 2.9 Part of IndianaMap Orthophotograpy dataset with NIR-R-G and R-G-B band combinations at same 1-foot spatial resolution.

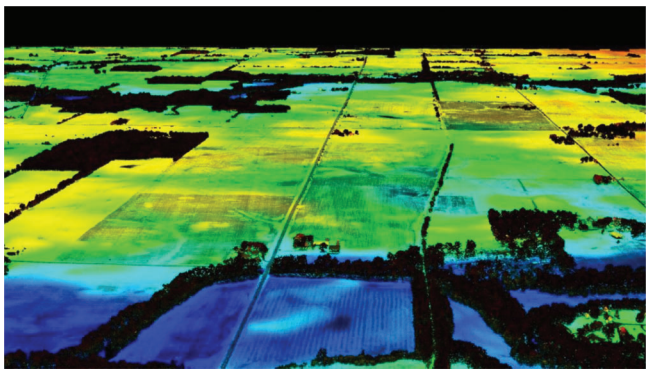


Figure 2.10 Sample from 2011–2013 IndianaMap LiDAR point cloud dataset.

3. EXTRACTING CLEAR ZONES AND CROSS-SECTION INFORMATION

In order to extract ground features via remote sensing methods, clear zones must be first identified with respect to the road network. Ideally, it is possible to establish a workflow to extract clear zones and cross-section information by first identifying the clear zones via spatially accurate road network and information on the road properties like lane width and the number of lanes. The main summary of the workflow in this case would be as shown in Figure 3.1(a). However, available road network dataset doesn't provide planimetric accuracy consistently to allow such identification as also mentioned in Section 1.1.1. Additional step as shown in Figure 3.1(b) is required to first extract the road surface extent via remote sensing which will be used for acquiring the clear zones area where the ground features will be extracted.

We have set up a framework for determining the paved surface extent, identifying the clear zones based on the paved surface, extracting the features within the clear zones, and acquiring cross-section information within the clear zones as well as the average grade along the road. We have employed available remote sensing datasets using the methods and tools established via research in computer vision, machine learning, geospatial analysis and computational geometry areas. The framework consists of multiple integral processes. First,

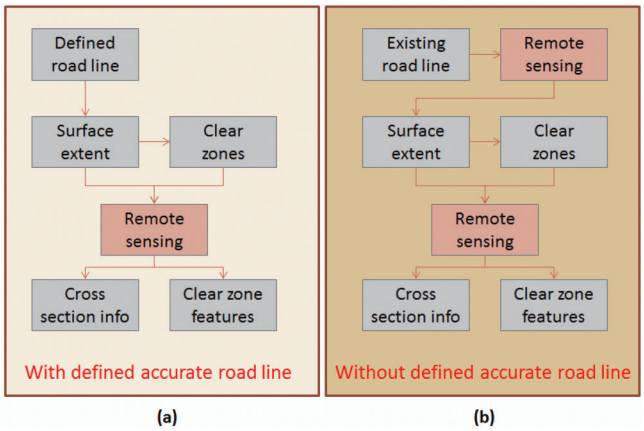


Figure 3.1 Main summary workflow for extracting clear zones and cross-section information with (a) and without (b) existing accurate road lines.

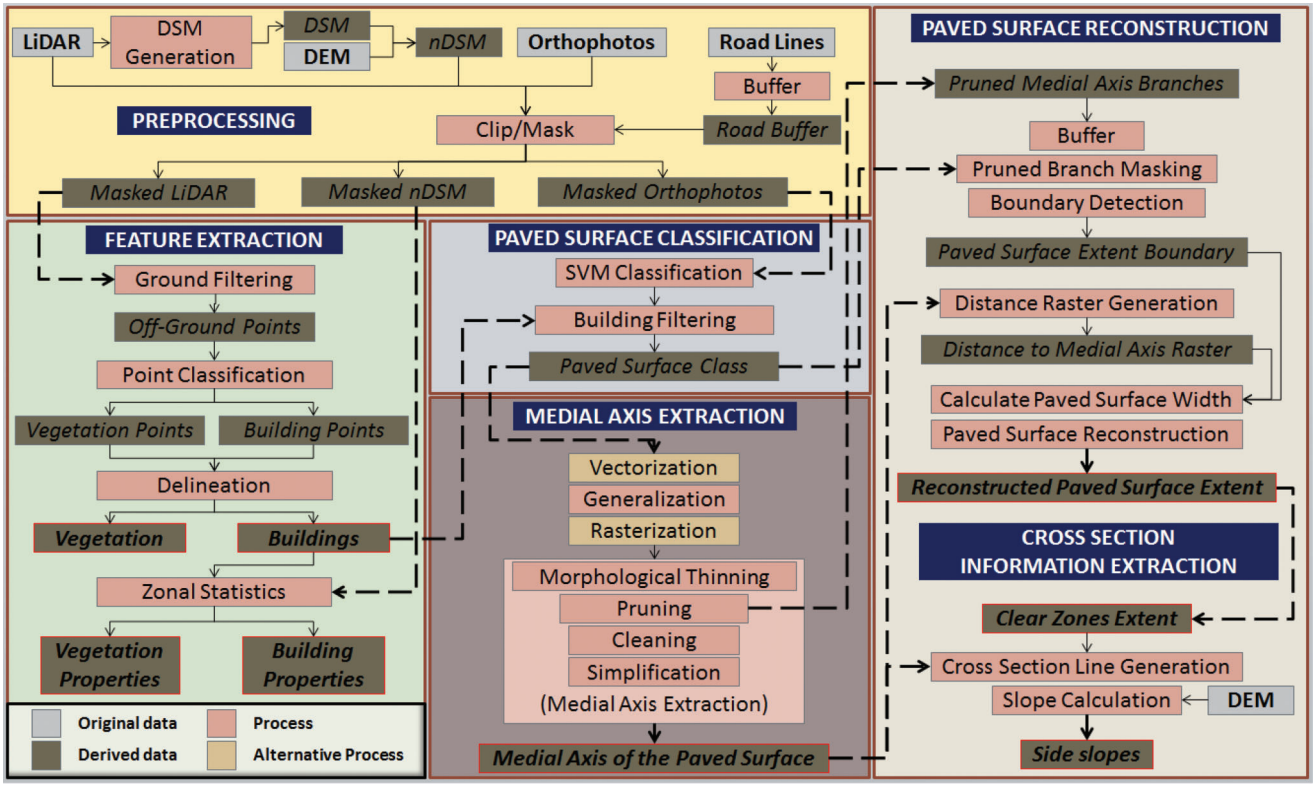


Figure 3.2 Overall process flowchart for the established framework.

there is a preprocessing step for preparing the datasets to be analyzed. Then these preprocessed datasets are passed on to process modules which carry on feature extraction, paved surface classification, medial axis extraction, paved surface reconstruction, and cross-section information extraction. Figure 3.2 shows the overall process flowchart for the established framework. Tasks share the original datasets and also the derived datasets of the other tasks. Relationships between the processes and their dependencies are indicated in the chart.

3.1 Preprocessing

We have selected a study area of approximately five by two miles in the Union township, Clinton County, IN for implementing the proposed framework. Since the most recent datasets are available from the 2011–2013 Indiana Orthophotography (RGBI), LiDAR and Elevation Project, we have used the CIR orthophotos, LiDAR point clouds and the LiDAR derived DEMs from that acquisition. We have also generated a DSM using the LiDAR point cloud of the study area.

The difference of the DSM and the DEM provides an nDSM which is basically the height above ground. The DEM is already provided as part of Indiana Orthophotography, LiDAR and Elevation project. We have used the LiDAR point clouds to generate a DSM at the same spatial resolution of the provided DEM and calculated the nDSM. The DSM is generated by assigning the maximum elevation of the points falling into the corresponding grid. Figure 3.3 shows the LiDAR point cloud, DEM, DSM and nDSM of part of the dataset.

We used the INDOT road network to generate a buffer around the road lines to use as a mask for limiting the amount of data involved in the process. Since most of the algorithms' computational costs are

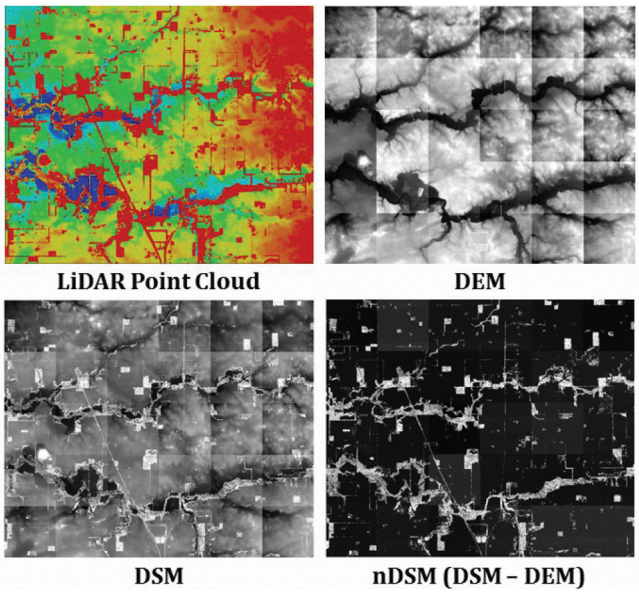


Figure 3.3 Lidar point cloud, DEM, calculated DSM and nDSM of Clinton County, Union Township, IN.

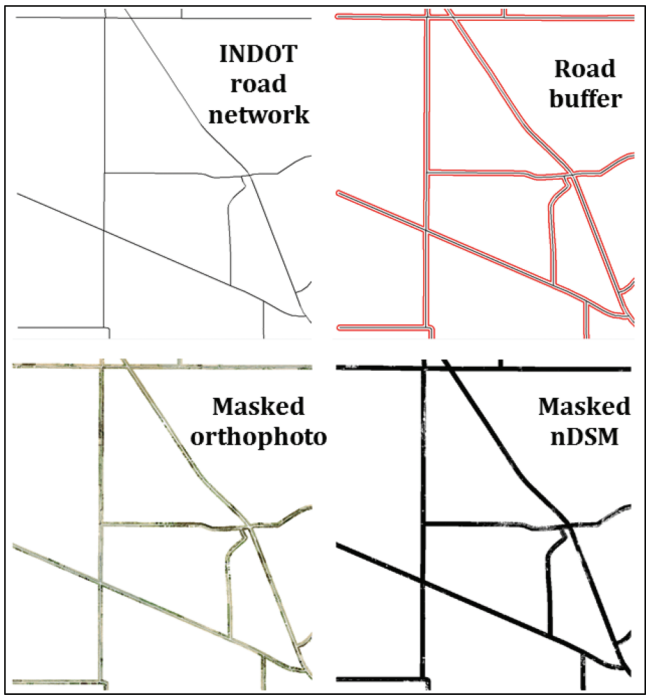


Figure 3.4 Sample of masked orthophotos and nDSM using a buffer around INDOT road network.

dependent on the size of the input, it is important to confine the application of the algorithms to the part of the dataset that is relevant. This is achieved by this masking step. Similarly, LiDAR point clouds are also clipped to contain points that are within this buffer. Figure 3.4 shows the masked orthophotos and nDSM for part of the study area.

The road network is available in ESRI shapefile format and the raster files are in tiff format. ESRI ArcGIS software is used to prepare the raster datasets to be available for further processing. LiDAR point clouds are available in LAS format, which is a public file format for the exchange of 3-D point cloud data (American Society of Photogrammetry and Remote Sensing (ASPRS), 2008). We have used the trial version of LAStools software to prepare the point cloud datasets. LAStools is a collection of tools for processing LAS files including functionality for conversion, filtering, clipping, tiling, triangulating, rasterization and several others for point cloud datasets. It allows scripting for batch processing and establishing process pipelines (Isenberg, 2013).

3.2 Classification of the Paved Surface

Once the irrelevant areas of the dataset are masked with a buffer around the existing road network, classification of the road surface may be carried out for the relevant areas. There are many classification algorithms available in literature for the classification of remote sensing imagery. Images may be classified pixel by pixel or considering groups of pixels as objects. Pixel-based classification methods consider each pixel individually in the feature space, regardless of their spatial configuration, and assign one class to each pixel independent from other pixels. Object-based classification on the other hand, considers the similarity of neighboring pixels' values to aggregate them as image objects first and then assigns classes to each image object instead of individual pixels. We have tested algorithms from both of these approaches for paved road surface classification.

3.2.1 Support Vector Machine (SVM) Classification

We have applied Support Vector Machine (SVM) classifiers as part of a pixel-based classification approach for the classification of the paved surface. SVM classification creates a maximum-margin hyper-plane in a transformed input space and splits the classes by maximizing the distance to the nearest clean split samples. This search for the optimal separating hyper-plane is performed after the original training data are transformed into a higher dimension. There is always a separating hyper-plane if the new dimension is sufficiently high and the transformation is appropriate. This hyper-plane is found with the aid of the "support vectors" and margins generated by the support vectors (Han, Kamber, & Pei, 2012). Details on SVMs may be found in Theodoridis and Koutroumbas (2009) or Han et al. (2012).

We collect training samples for the road and non-road pixels to train the SVM classifier. The training data include samples that are representative of all different types of road surfaces that we have observed that exist in the dataset. Due to different surface types, different maintenance patches and color changes from wear and tear, it is important to determine the existing road types and collect training samples accordingly. Figure 3.5 shows samples of several such visible spectral differences of road surfaces from the orthophotos.

Figure 3.5 Examples of visible spectral differences of road surfaces from the orthophotos.

Once the training samples are collected, we train the SVM classifier using the Monteverdi software interface for the Orfeo Toolbox (OTB) library of image processing algorithms. OTB is an open source library distributed under a free software license CeCILL which contains image processing algorithms including SVM classification of images. It is developed by the French Space Agency (CNES) based on the medical image processing library ITK to support the use of images from Pleiades (PHR) and Cosmo-Skymed (CSK) satellites (CNES, 2013).

We use a C-Support Vector Classification type of SVM with a linear kernel and set the parameter c = 1. After training the SVM classifier, we classify the study area tile by tile using the same trained model for each tile. The classification results are then validated using randomly selected test samples. The confusion matrix for part of the dataset tested is provided in Table 3.1. An overall accuracy of 98% is achieved for road/non-road binary classification. The ratio of misclassification of non-road pixels as roads is higher than the misclassification of road pixels as non-road. Figure 3.6 shows a sample of the binary classification results.

TABLE 3.1
Confusion matrix for road/non-road classification of CIR orthophotos

Confusion Matrix			
	Road	**Non-road**	**Sum**
Road	**13417 (97.76%)**	307 (1.23%)	13724
Non-road	340 (2.48%)	**24579 (98.64%)**	24919
Sum	13757	24886	
Accuracy			
	Kappa	0.9635	
	Overall Accuracy	0.9833	

3.2.2 Building Filtering

A very common misclassification which occurs while extracting the paved road surface by classifying CIR orthophotos is the classification of some of the buildings that are contagious to the roads, as roads, due to their spectral similarity. Additional information is required to avoid such misclassification. Ideally, having up to date building outlines would be sufficient to exclude the buildings. However, such up to date and complete building databases are not commonly available. Hence, one needs to determine the building outlines to use them as a mask for acquiring just the paved surfaces out of the classification result. This may be achieved by several different methods.

Using the DEM and DSM as an additional source of information is one remedy to distinguish the buildings and roads with similar spectral properties. In such cases, even though the spectral responses from a road pixel and a building pixel are similar, their heights above ground will be different. The nDSM is used to filter out such pixels that are classified as roads from the classification results. It may be used to aid filtering out the buildings in two alternative ways; either by using them individually as a mask to remove elevated pixels or introduced in the classification process as an additional feature. Following the prior approach, we have converted the nDSM into a binary raster by applying a height threshold. This binary raster included all pixels that are above certain height as 1 and the rest as 0. We also had the classification results as a binary raster such that the roads were encoded as 1 and non-roads as 0. Subtracting the binary raster of elevated objects from the binary road classification result and leaves the pixels that are on the roads in the classification raster. Figure 3.7 depicts this process for part of the study area. The binary raster in Figure 3.7(b) includes all objects that are above certain height in the nDSM while Figure 3.7(a) includes only the pixels that are classified as roads. Figure 3.7(c) shows both the road pixels and the pixels that are initially classified as roads but that are above the height threshold. Once the high objects are removed, the road class can be vectorized to a road polygon GIS layer.

One issue that needs to be addressed when removing high objects happens in case the trees that appear as high objects in the nDSM coincide with the roads.

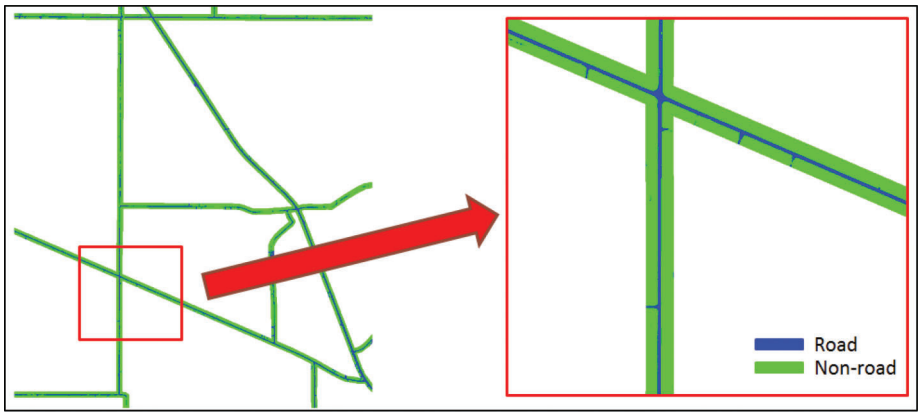


Figure 3.6 Sample from the road surface classification result using SVM.

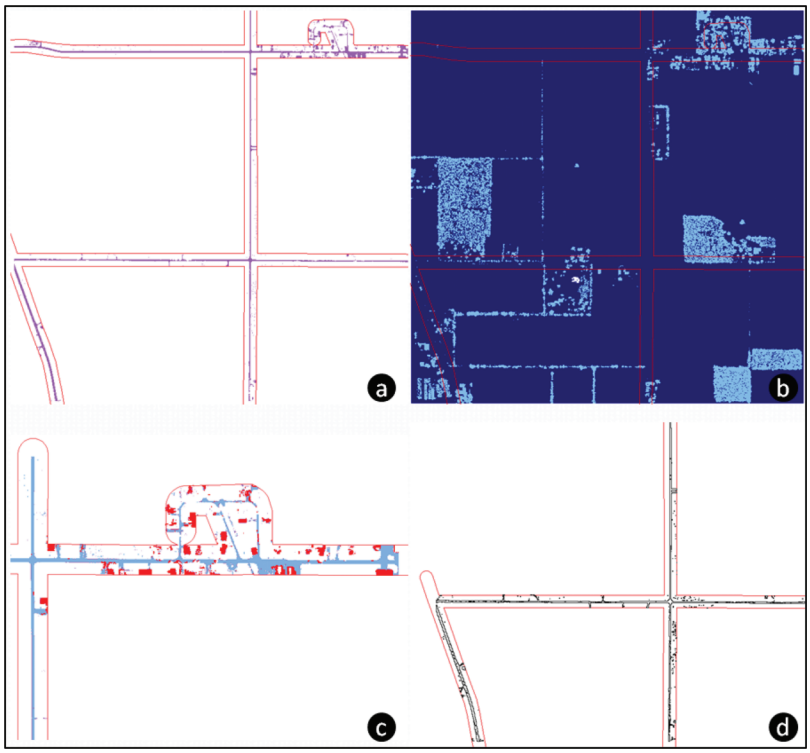


Figure 3.7 Removing high objects from road classification results and vectorization: (a) road classification result, (b) binary nDSM of pixels above threshold value, (c) roads and high objects classified as roads, (d) vectorized road polygons.

Using the nDSM directly as a filter to remove high objects may cause problems if the trees are hanging over the road and the road is already correctly classified from the orthophotos. In such cases, filtering process will remove the information that is already correctly acquired during the classification process. In order to avoid such problems, an additional step is required to adjust the nDSM which will be used as a filter. We generate an NDVI (Normalized Difference Vegetation Index) and use it to mask the nDSM so that only high objects that are not vegetation remain to be used to filter the classification results. However, some artifacts are introduced when the NDVI is applied in this fashion for classifying the objects with significant height above a threshold as buildings and vegetation. Orthophotos used for generating the NDVI are high resolution images taken during leaf-off season. NDVI reflects the high resolution nature of the images and provides a non-homogenous index within the area that the tree covers on the ground. As a result of this, some of the pixels within the footprint of a tree have calculated indices which indicate non-vegetation. This translates into patches of high objects that can't be excluded as vegetation from the lower resolution nDSM even though their heights are due to the trees existing on the ground.

In order to overcome this artifact, we preferred the use of building outlines obtained as a result of the feature extraction process using 3-D LiDAR point cloud. Having these building outlines, buildings are easily excluded from the paved surface classification results. In case there is no LiDAR point clouds available, one may prefer to follow an object-based classification approach including the trees as an additional class to enforce smoothness within the tree patches. Since the LiDAR point cloud is available for our study we have not evaluated the outcome from such approach.

3.3 Extraction of the Medial Axis of the Paved Surface

Classification results provide an irregularly shaped noisy raster sampling of the paved surface. Two of the major tasks in this study are extracting the features within the clear zones and acquiring cross-section information. The road extent needs to be defined in order to achieve these goals. It is not possible to obtain such a definition of the road extent with the irregular nature of the classification results as they are. We employ a series of processes to reconstruct the road based on the raster classification results. First, we apply a cleaning and generalization procedure followed by a series of morphological operations to extract the medial axis of the paved surface. Based on this medial axis as the centerline and the boundary of the classified paved surface, reconstruction of the paved surface extent is achieved.

3.3.1 Vector-Based Cleaning and Generalization of Classification Results

Classification results of the paved road surface do not directly provide a topologically consistent, complete, accurate geometric model of the road extent.

Initial classification results need to go through multiple-step post-classification cleaning and generalization process to be able to acquire the intended information from the classification of the paved road surface. Once the surface class is obtained, one may follow either a vector-based or a raster-based approach for the generalization and regularization of the extracted road surface as indicated on the workflow in Figure 3.2. We have investigated and will explain implementation details for both approaches.

Vector-based generalization and cleaning of the classification results consist of several steps including the vectorization of the raster classification results, repairing the geometry of the vectorized polygons, eliminating multi-part polygons, merging small polygons with larger ones, removing isolated polygons, and simplification. Figure 3.8 shows an overall flowchart of this process.

The polygons that are generated as a result of the vectorization of the classification raster may have issues regarding their geometries and topologies in the GIS. The step after the vectorization is to ensure that the road polygon layer includes no such issues. We use the tools provided in ESRI ArcGIS software to eliminate such problems by checking and fixing the geometries of the polygons. These problems are due to the requirements of having data consistency in a GIS. Depending on the system and software used to keep and maintain the GIS, such requirements may change. In the ESRI ArcGIS software, some of the problems that one may face include short segments, null geometry, incorrect ring ordering, incorrect segment orientation, self-intersections, unclosed rings, empty parts, duplicate vertices, and discontinuous parts (ESRI, 2013). Once the GIS layer geometry problems are repaired we separate the multi-part polygons to be treated as single part objects.

Pixel-based classification generates many small polygons since each pixel is classified individually. Merging these small polygons with the polygons that are of the same class in their adjacency ensures the integrity of the road polygon layer and helps to reduce the issues that may occur while further analyzing these polygons. One reason for the existence of such small polygons larger than several pixels is due to vehicles traveling on the roads. Any geometric parameters which may be queried for analysis purposes from such fragmented dataset would not be reliable to represent the actual properties of the scene and may be misleading. Apart from such small polygons with spatial neighborhood to larger polygons of the road class, there are also small polygons that are misclassified as roads which have no spatial connection to the main road polygon. We determine a threshold area for such small polygons that are isolated from the main road polygons and remove them from the classification results. Figure 3.9 shows some sample results of this merging and removing process.

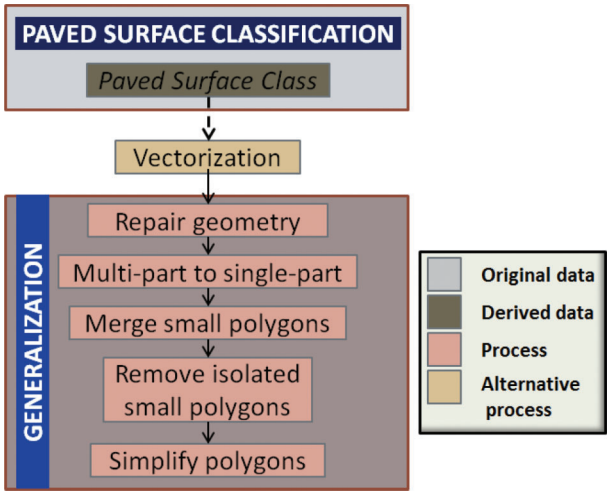


Figure 3.8 Flowchart for the vector-based cleaning and generalization of road classification results.

After removing and merging the small polygons, the next step involves generalization of the polygons since the classification results have irregular boundaries due to noise in the classification process. This is achieved using the Douglas-Peucker algorithm (Douglas & Peucker, 1973) which is a commonly used algorithm for simplification. The algorithm reduces the number of vertices based on a user defined offset tolerance. The vertices of the simplified polygon include the points from a subset of the original vertices.

We performed all post-classification vectorization, cleaning and generalization tasks using the ArcGIS model builder environment to ensure standardized workflow that is applicable to a tiled dataset. The model is provided in Figure 3.10 while the extracted road surface for part of the dataset is represented in Figure 3.11.

Vectorized, cleaned, and generalized road surface classification results provide the areas covered by roads as well as parking lots, driveways, etc., since they have similar or identical spectral response as the paved road surface. Also, trees hanging over the roads cause gaps and discontinuities in the road polygons. One may talk about two different types of deviations of opposite characteristics from the ideal road polygon; inclusive and exclusive. These deviations need to be accounted for by further processing.

Figure 3.9 Parts of the vectorized road classification results before (a) and after (b) the removal of small polygons with no connection to the main road polygon as well as before (c) and after (d) merging the small hollow polygons to the main road polygon.

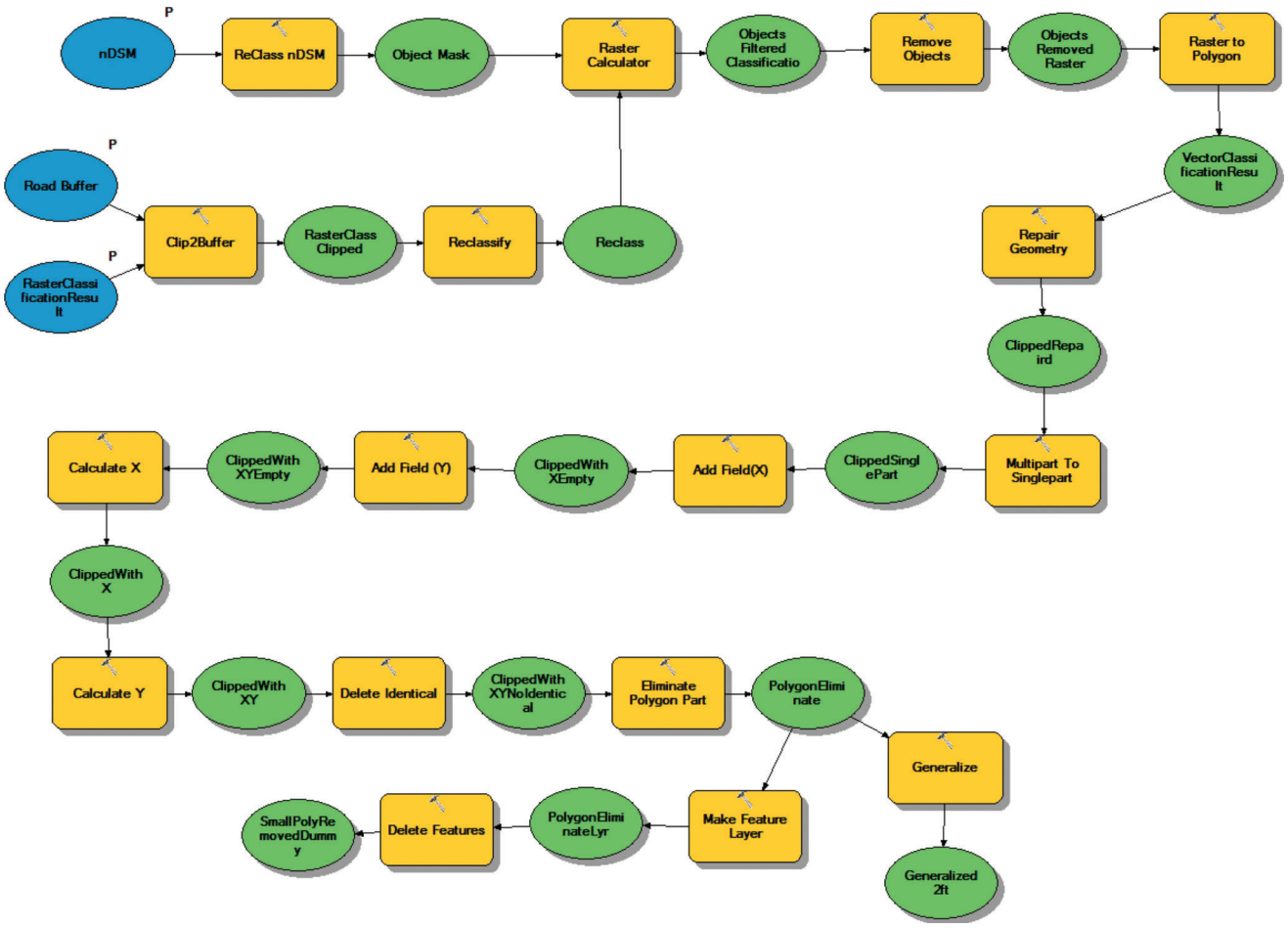


Figure 3.10 Geoprocessing model for vectorization, cleaning and generalization of road classification.

3.3.2 Raster-Based Cleaning and Generalization of Classification Results by Morphological Operations

Alternatively, cleaning and generalization operations may be performed directly on the raster surface classification results using morphological operations. We favor this option over vector-based cleaning process which we have just described due to the flexibility and strength of mathematical morphology which is the basis for these operations.

Mathematical morphology is a technique which studies form, shape and structure. Initially, it emerged in 1960s to study random sets in mining applications for working on properties of ores (Shih, 2009). Its use quickly expanded in the area of image processing first to deal with 2D binary images, followed by applications on gray-level and multispectral images (Najman & Talbot, 2010). In image processing, morphological operations provides means for simplification of images by preserving the main characteristics of shape and form and reducing irrelevant deviations from the overall structure of the shapes (Haralick, Sternberg, & Zhuang, 1987). This leads to the possibility to deal with image features like edges, corners, holes, wedges, etc. There are many efficient algorithms in literature which tackle such issues using mathematical morphology. Mathematical morphology is used to perform tasks like filtering, thinning, pruning, image enhancement, restoration, segmentation, defect identification, object recognition, etc. It employs the principles of various disciplines such as set theory, lattice algebra, discrete and continuous geometry, topology, function, measure theory and more (Najman & Talbot, 2010; Shish, 2009). It serves as a natural approach for machine vision recognition problems since identification of the objects are directly related with their shape (Haralick et al., 1987).

Morphological analysis extracts knowledge based on the response of generally non-linear transformations Najman & Talbot, 2010. Mathematical morphological transformations, when applied to a set of any dimension, results with two sets; transformed, i.e., selected points and not selected ones (Haralick et al., 1987). Transformations are defined over a lattice which is defined as a set with an ordering relationship with reflexive, anti-symmetric and transitive properties, and complete. Also, a structuring element which is a kernel image with a designed shape is required to perform a morphological operation. Morphological operations modify the original input image which is also known

Figure 3.11 INDOT road network (left), orthophoto (middle), road pavement classification result (right) for part of the Indiana Orthophotography Project dataset.

as the active image, by probing it with structuring elements of varying sizes and shapes (Najman & Talbot, 2010; Shih, 2009). The basic morphological operators are dilation and erosion operators. Dilation operator results in the filling, expansion, or growth of the active image while the erosion operator has a shrinking effect. These basic operators may be combined to establish more complex operators like opening and closing operators or hit-or-miss transformation. More complex algorithms may also be designed for tasks like boundary extraction, region filling, extraction of connected components, convex hull, thinning, thickening, skeletonization, pruning and edge detection (Haralick et al., 1987; Shih, 2009). Further details on morphological operations may be found in the cited references.

Using morphological operations is very well suited for the post-processing tasks that are introduced following the classification of the paved surface. For example, when we consider the exclusive problem of completing the missing parts of the paved surface classification, parts that are missing as holes are completely surrounded by the road polygon or by foreground (road) class in the raster case. In order to complete the missing holes, morphological reconstruction (Soille, 1999) may be performed where the holes (background pixels in a binary classification) that can't be reached from the edge of the image are filled. We employed the libraries provided by Matlab software to carry out this and other morphological operations. Figure 3.12 shows the processing steps involved for the cleaning and generalization of classification results by morphological operations.

Figure 3.13 shows an example of filling such holes followed by a morphological opening operation. In this case, missing parts are observed in the initial classification result due to color change on the road surface as well as a vehicle at the intersection. These imperfections are easily handled by morphological operations. It is

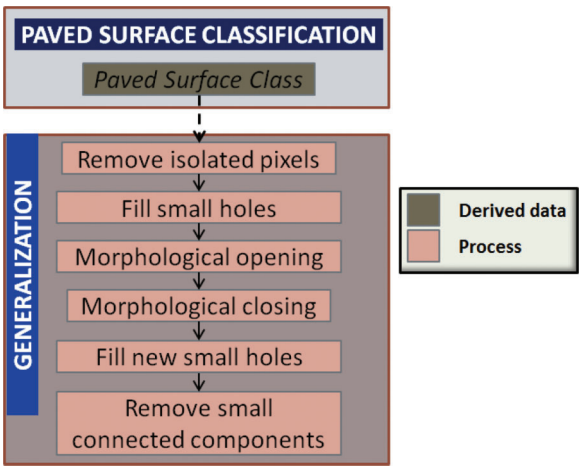


Figure 3.12 Flowchart for the raster-based cleaning and generalization of road classification results.

Figure 3.13 An example of before (left) and after (middle) filling the missing parts of paved surface classification using morphological reconstruction followed by a morphological opening (right).

also observed that the morphological opening operation removes most of the irrelevancies observed as noise in the classification result.

3.3.3 Medial Axis Extraction by Thinning and Pruning

Deviations in the paved surface classification results from the roads still exist at this stage due to parking lots and driveways connected to the roads. Instead of attempting to detect and compensate for these deviations by dealing with the polygonal structure of the classification results, we propose to first extract the medial axis of the paved surface. Such a reduction allows more flexibility to apply further methods to extract the road structure. After thinning, it is possible to employ morphological operators to prune the excessive irregular branches that are not part of the main roads, estimate the surface width from the surface boundary, and finally reconstruct the surface based on this medial axis and the estimated road width.

Various terms are used synonymous to "thinning" in literature like "medial axis transformation", or "skeletonization" even though there are subtle differences between these terms. The term "thinning" may be considered for use in case of reducing the generally elongated patterns into line like structures while "medial axis transformation" would be more appropriate for the determination of centers of maximal blocks (Lam, Lee, & Suen, 1992). There are many algorithms developed for thinning. We employ the Matlab implementation of the algorithm in Lam et al. 1992. Figure 3.14 shows an example of thinning of the cleaned and generalized paved surface classification result.

The skeleton obtained after thinning provides a simpler representation of the paved surface. Apart from the main skeleton of the actual road surface, branches corresponding to any extension to the roads or due to any non-uniformity are generated as well. Mathematical morphology provides tools also for pruning these unwanted branches or spurs. There are various approaches for pruning. In principle, branches may be located by checking their endpoints which have only one neighbor and moving from there towards the

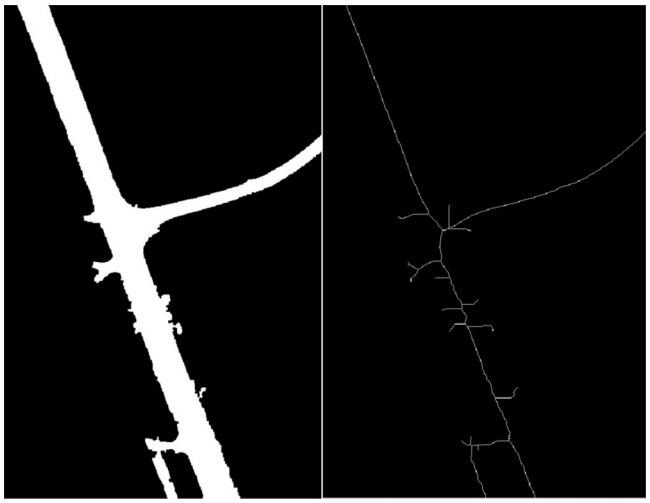


Figure 3.14 An example of morphological thinning in part of the study area.

intersections which have at least three neighbors (Shih, 2009). We have employed the implementation that is available in the software Wolfram Mathematica for pruning. Figure 3.15 shows the result of pruning operation on part of the paved surface skeleton in the study area.

After pruning the centerline, we perform a connected component analysis and remove groups of connected components that are less than certain number of pixels. This step ensures that small pieces that are classified as paved surfaces but are irrelevant are cleaned. Once the medial axis is extracted as the centerline of the paved surface, a simplification process is applied to obtain a smooth estimate of this noisy centerline. This step is performed after the pruned centerline is converted to vector format. We apply the Douglas-Peucker simplification algorithm (Douglas & Peuker, 1973), this time for the polyline feature instead of the polygon as in the vector-based cleaning and generalization process. One particular important benefit of this step is that it

Figure 3.15 Morphological pruning example for unwanted branches in part of the study area.

realigns parts of the centerline which significantly diverged from the actual axis of the paved surface area during thinning process due to the existence of the branches. This may be observed in Figure 3.16. Pruned centerline on the top right side of the figure bulges towards the driveway which connects to the main road because of the removed part of the paved surface classification result. It is aligned with the actual road after the simplification step. Cleaned and generalized centerline for the whole study area is also presented on the left side of Figure 3.16.

3.4 Reconstruction of the Paved Surface

Once the centerline is available, it may be used as a reference to reconstruct the paved surface upon which the clear zones may be established. We propose to achieve this by first, extracting the boundary of the paved surface from the classification results, then, finding the distance of each boundary pixel to the centerline, and after that estimating an average road width for each centerline segment of 50-foot intervals for piecewise reconstruction of the paved surface.

As mentioned previously, mathematical morphology provides algorithms for the purpose of boundary extraction. In a binary image, boundary pixels may be identified as having at least one background pixel in its 8-connected neighborhood. Boundary extraction may be performed first by an erosion operation over the active image with an isotropic structuring element. Then the set difference of the active image and its erosion is taken (Shih, 2009).

Figure 3.16 Pruned centerline after cleaning and generalization for the whole study area (left), and examples of pruned centerline before (upper right) and after the cleaning and generalization process (lower right).

We extract the boundary of the paved surface with which we plan to calculate the estimated width along the centerline. This boundary raster, however, still includes the irrelevant branches since the branches have only been removed over the medial axis previously. In order to remove the branches from the boundary raster as well, we generate large enough buffers around branch lines which are previously calculated as the difference of the original medial axis and the pruned medial axis. Branches are removed by using these buffers as a mask over the boundary raster. Figure 3.17 shows examples of this branch removal process.

Next, we generate another raster for which the distance of each pixel to the nearest centerline point is calculated. By multiplying this distance raster with the binary boundary raster, we obtain a boundary distance raster in which each boundary pixel holds its distance to the centerline. Values on the boundary pixels now correspond to the width of the paved surface at that point for the side that the boundary pixel is at. In order to reconstruct the paved surface, we generate large enough buffers around each segment of the centerline that will include all boundary pixels for that segment. Then we calculate zonal statistics over each buffer with the "distance to centerline" values of the boundary pixels. The average distance value for each centerline segment buffer is then assigned to the corresponding centerline segment. This allows the reconstruction of the paved surface extent based on the average width of each centerline segment. Figure 3.18 shows examples of reconstructed paved surface and removed branches overlaid with orthophotos.

The mask buffers may sometimes be smaller or larger than they are required to mask the branches. Such situations affect the result of paved surface width estimation. The mask being slightly larger than the actual branching section of the paved surface does not have as significant effect as the opposite case. If the buffer is larger, more boundary pixels are removed. Such a situation would not affect the paved surface width estimation as long as there are enough pixels within that segment to infer the paved surface width. In the opposite case, remaining patches of branch boundary pixels would affect the estimated width since there will be additional pixels which will shift the average distance within the segment to be larger. This situation is observed in some segments throughout the dataset as an expansion in the reconstructed surface extent.

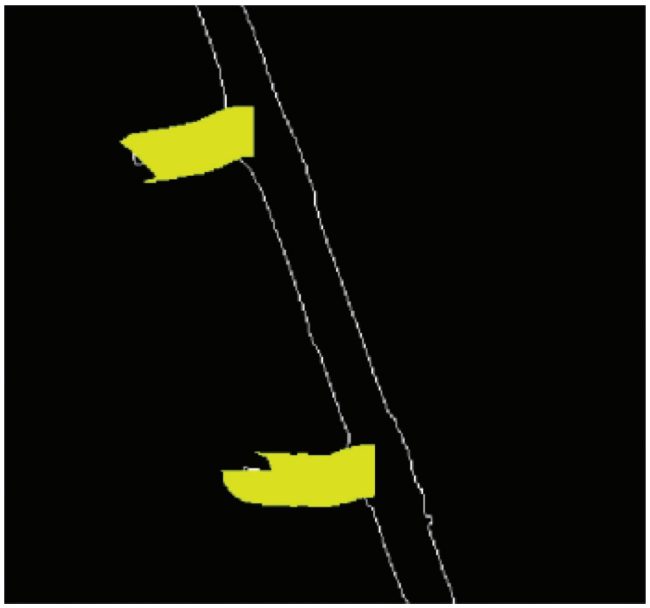


Figure 3.17 Examples of the boundary raster and buffers generated around irrelevant branches used to mask the boundary raster.

3.5 Extraction of the Cross-Section Information

Reconstruction of the paved surface extent provides the opportunity to define the clear zones based on this extent. We have established 10-20-30-foot buffers around the reconstructed paved surface. Cross-section information is to be acquired within these clear zones. We have generated cross-section lines corresponding to the center of each centerline segment starting from the extent of the reconstructed paved surface to the extent of the clear zones. Using the DEM, we were able to calculate the slopes along these cross-section lines. Figure 3.19 presents samples from the generated clear zones and the cross sections.

3.6 Feature Extraction

With defined clear zones, the study area is now confined to the regions from which the features will be extracted. We employ 3-D airborne LiDAR point clouds to extract the features within clear zones. Since there is no high point density mobile or terrestrial LiDAR datasets available, detecting the small features which cannot be captured by the airborne LiDAR within the clear zones is not possible. However, buildings and trees are detectable using the airborne LiDAR dataset.

3.6.1 Light Detection and Ranging (LiDAR)

Laser scanners are active sensing systems using laser beams to measure the range between the target and the sensor. Several useful technical properties of the lasers like the ability to realize high energy pulses in short intervals and the use of short wavelengths allow the use of relatively small apertures for very precise ranging. In addition to these properties, the availability of high pulse repetition rates, allow the use of the LiDAR systems for the production of range images (Wehr & Lohr, 1999).

There are airborne, terrestrial and mobile platforms employed for LiDAR data acquisition with a variety of system configurations which keep improving constantly. Along with pulse-based systems, there are also continuous wave systems which utilize phase measurements for ranging. The common product of

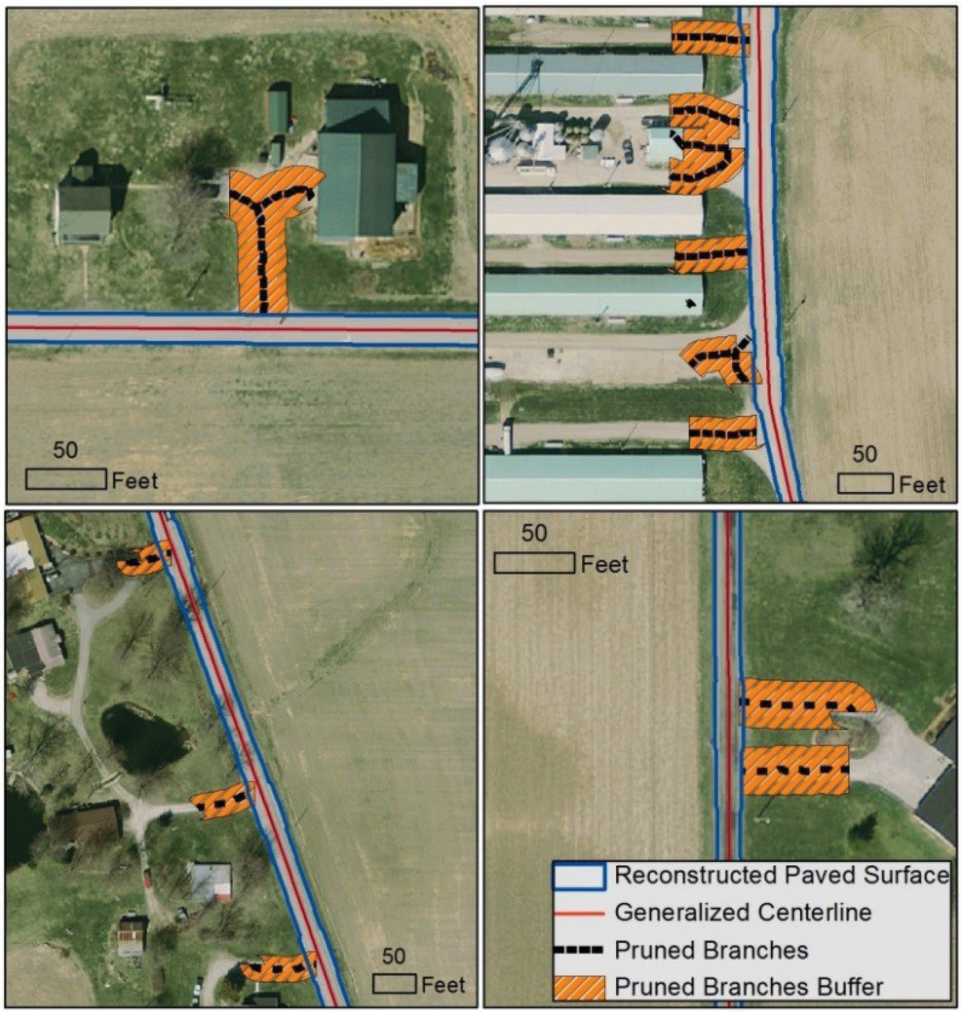


Figure 3.18 Examples of reconstructed paved surface and excluded branches.

all these systems is a collection of point location data referred as point clouds which may then be used to derive various products. Along with the point locations, LiDAR systems may also provide the intensity values or multiple returns in a single echo (Ackermann, 1999; Baltsavias, 1999; Vosselman, 2009;Wehr & Lohr, 1999).

LiDAR is increasingly utilized for remote sensing of the earth as an emerging technology in the last two decades. It is considered as one of the most important data acquisition technologies introduced for geospatial data acquisition lately (Petrie & Toth, 2009). It is extensively and routinely used today in topographic mapping as a direct 3-D data collection technique.

Most common airborne LiDAR sensors use directed laser pulses emitted at specific spatial and time intervals to calculate the ranges to the footprints of these reflected pulses followed by the calculation of their 3-D locations. These calculated locations provide 3-D sampled representations of the terrain and the objects on the terrain with varied densities depending on the operational parameters of LiDAR data acquisition. However, simpler representations are required for practical purposes for many applications with respect to more efficient analysis, operability and data size. Extracting meaningful information of ground features using this type of LiDAR sensor data requires labeling and segmentation of these unstructured sets of points which are often called "point clouds."

3.6.2 LiDAR Point Cloud Processing

Distinctive nature of unstructured LiDAR point clouds from the structured raster properties of remote sensing images prevents direct implementation of most remote sensing methods and algorithms that are successfully applied to remote sensing images. Since the emergence of the use of LiDAR scanning for topographic mapping, methods and algorithms specific for LiDAR point cloud processing to extract information on Earth's topography as well as the natural and man-made features have been studied. Some of these algorithms transform the point clouds into a structured 2-D range image which allows the implementation of raster-based methods to extract information from

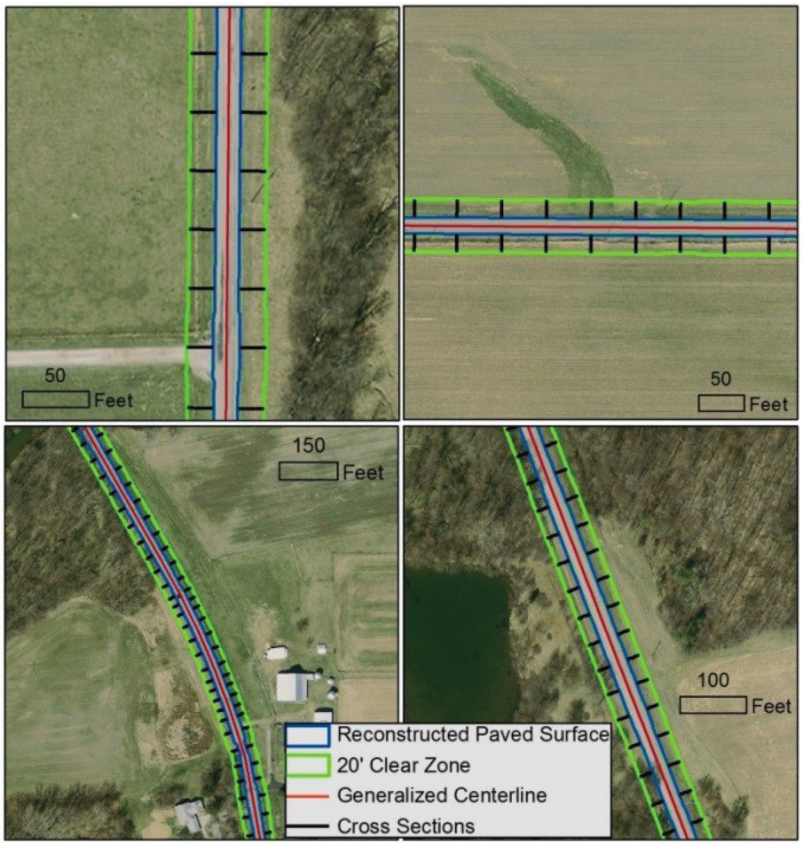


Figure 3.19 Samples from the 20-foot clear zone and the cross sections along the road.

LiDAR data. This approach has its benefits since it allows the use of many robust methods developed for raster processing over the decades. It also has disadvantages since this transformation from the point cloud to the range images may result with loss of information due to generalization which may only be preserved when the properties inherent to the unstructured nature of 3-D point clouds are exploited.

Modeling the terrain has traditionally been a major objective and motivation of initial topographic laser scanning applications. Classification of points as "ground" and "non-ground," is commonly called "ground filtering". There are various filtering algorithms in the literature for which the details may be found in Kraus and Pfeifer 1998, Vosselman 2000, Sithole and Vosselman 2004, Shan and Sampath 2005, and Pfeifer and Mandlburger 2008. Urban airborne lidar point clouds are analyzed by various approaches and frameworks in the literature. Most of these variations include ground filtering as the initial process either just by itself or together with building detection which involves determining the building points. Once the ground points are filtered, remaining points are processed for extracting planar surface patches for reconstructing man-made structures like buildings which are separated from vegetation and other non-surface objects.

3.6.3 Building and Vegetation Extraction from 3-D LiDAR Point Clouds

We have used the trial version of LAStools software for ground filtering as well as classification of buildings and vegetation from the point cloud datasets. The main workflow we have employed for feature extraction as presented in Figure 3.2 follows the conventional approach of initial ground filtering. Figure 3.20 presents sample results of the ground filtering process from the study area.

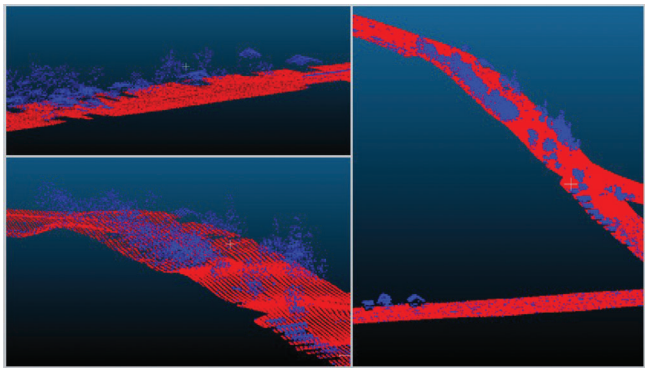


Figure 3.20 Ground filtering results from parts of the study area. Red color indicates ground points while the blue color represents off-ground points.

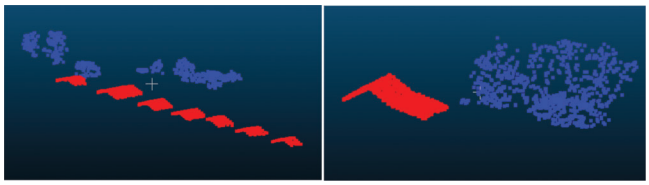


Figure 3.21 Samples of building and high vegetation point classification results. Blue points are classified as vegetation while red points are buildings.

Having ground and off-ground points classified, each point in the point cloud may be assigned an estimated height since an approximate ground may be inferred within the local proximity of each point. This allows the determination of buildings and vegetation above some certain height threshold. Building extraction requires the determination of points that are on a planar surface. Figure 3.21 shows sample building and tree point classification results from the study area.

Once the points are classified, delineation provides the outlines of the buildings and trees. Trees are delineated as either individual trees or as groups of trees. Since the LiDAR dataset is from an airborne system, the points that are on the trees are mostly from the crowns and rarely from the trunks. This situation limits the opportunity of extracting individual trees to estimate their locations. Since we are dealing with tree crowns, they are very likely to connect with each other and make determining the actual proximity of the tree trunk to the road very challenging. At this stage, one may try to estimate the individual trees by relying on a-priori knowledge about the tree type and structure. We have not attempted any effort to extract individual trees within the scope of this project. Figure 3.22 shows samples from the delineation of classified building and tree points. We have employed the *alphahull* library of R programming language to perform the delineation.

Having the polygon outlines of the features, we were able to overlay them with the nDSM to calculate zonal statistics for each building and tree group. We have also calculated the distance of each building to the paved surface edge of the closest centerline segment. By associating the properties of extracted feature outlines with the clear zones, the sizes of and the distances to the extracted features becomes available upon query. Aggregation of the properties per centerline segment is also available after defining the required properties to be aggregated for each centerline segment.

A total of 86 objects were classified as buildings within the study area. Among these, 74 of them were actual buildings while 12 were either roads or vegetation. The reason for misclassification of buildings as roads is due to

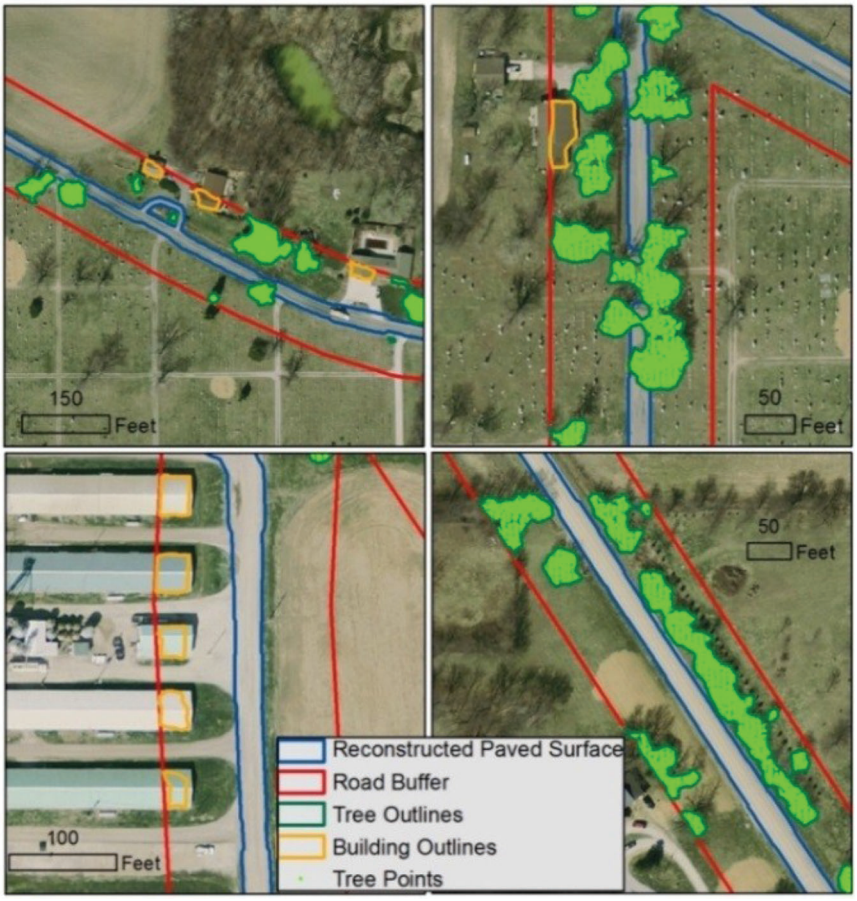


Figure 3.22 Samples of building and high vegetation classification delineation.

the elevation difference between the object and its immediate ground. Among nine misclassifications in the study area, seven of the objects misclassified as buildings were actually bridges and two of them were part of the road with very steep slopes. Figure 3.23 shows several of these misclassifications overlaid on orthophoto.

There are also trees that are misclassified as buildings. This occurs when the points on the trees show planar structure. This can be avoided by employing the NDVI and post-classifying any outline for which the majority of the NDVI indices within the polygon are in the range of vegetation as buildings.

3.6.4 Feature Extraction with Object-Based Image Analysis (OBIA)

Before the availability of the 2011–2013 Indiana Orthophotography (RGBI), LiDAR and Elevation Project Dataset, we have acquired sample orthophotos together with the DEM and DSM from the 2005 Statewide Orthophotography Project to test different methods for the extraction of ground features using these datasets.

We have defined two classes; roads and buildings, and established an object-based hierarchical classification model for delineating the extent of the road pavement, and the buildings. In pixel-based image classification methods, each pixel in the image is classified individually based on their spectral responses. Object-based classification considers objects as classification units instead of individual pixels. The image is first segmented into homogeneous components based on the spectral properties of the pixels. This segmentation step results in a hierarchical network of image objects which resemble real ground features (Blaschke & Strobl, 2001). Classification is performed following this segmentation step using the image segments as classification units. Once image segments are obtained, it

Figure 3.23 Examples of bridges and parts of the roads with steep slope that are classified as buildings.

Figure 3.24 Sample building and road extent extraction results overlaid with INDOT road centerlines and the aerial image.

is possible to extract higher level information from these segments based on their spectral, spatial, contextual properties and their size (Baatz & Schäpe, 2000). The second step is the classification of the image segments based on these cues using various classification methods. Obtained thematic classes are more homogenous. Even though there may be variations among the pixels of one segment, the pixels within an object are classified as one class.

We have established a classification scheme using geospatial data processing software ERDAS Imagine. The software provides a module named "Objective" which allows setting up object-based classification models and perform classification based on the constructed model. We used the RGB orthophoto together with the nDSM for segmentation. We followed a one class at a time approach and collected training samples representing the classes of interest and trained a multi-bayesian network for calculating the probabilities of each pixel belonging to the relevant class. We have also included a texture layer derived from the orthophotos while calculating the probabilities for the building class. This allowed a better elimination of high vegetation from the building class due to difference in their texture. We have applied probability filters to extract the segments that are above a probability threshold of belonging to a class. Once all segments that belong to the class of interest are extracted, we have merged the segments of the same class based on their spatial configuration. This was

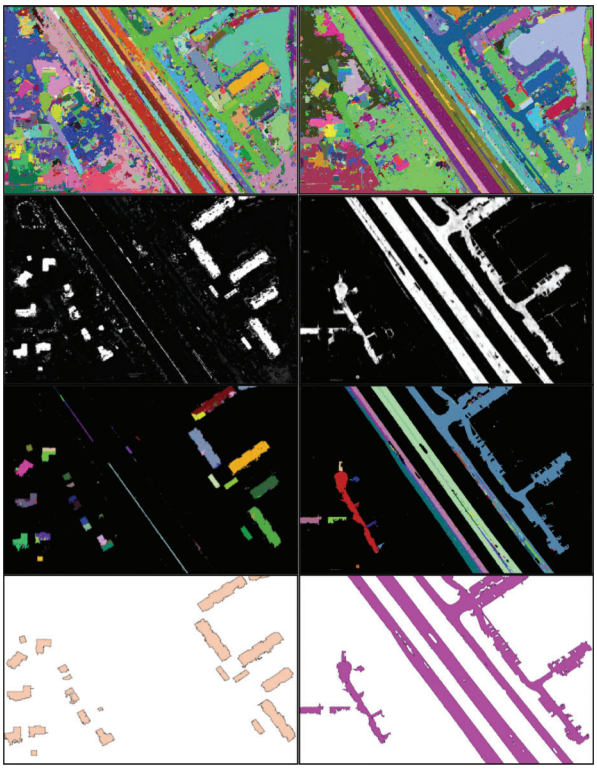


Figure 3.25 Intermediate steps for the extraction of buildings (left column) and roads (right column) with object-based classification. From top row to bottom: segmentation, pixel probabilities, filtered segments, final objects.

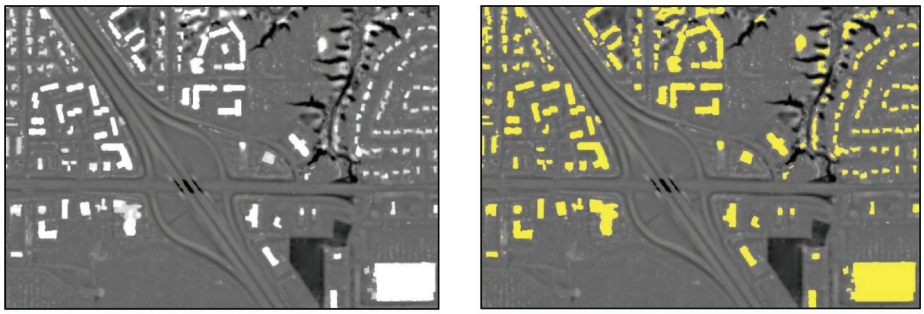


Figure 3.26 Initial building detection from nDSM.

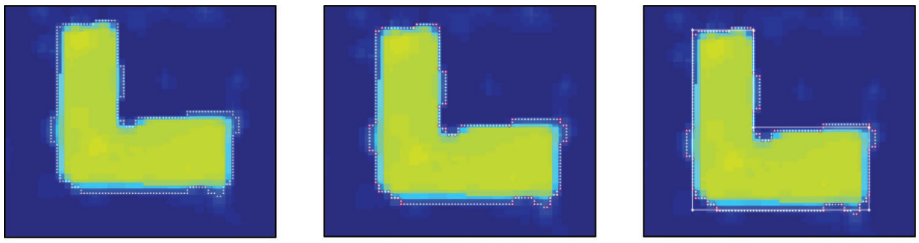


Figure 3.27 Building footprint extraction.

followed by the vectorization of the object boundaries and application of filters which account for the geometric properties of the features of interest. Figure 3.24 shows the final extracted results overlaid with INDOT road centerlines and orthophoto. Figure 3.25 shows intermediate results throughout the processing steps and extracted ground features for a small example area.

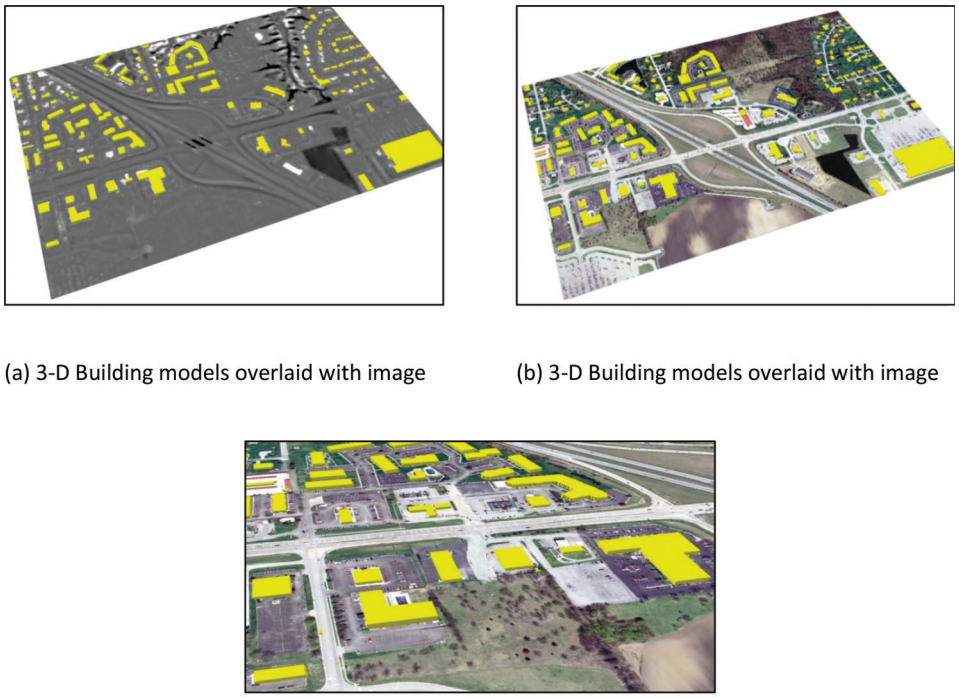
3.6.5 Feature Extraction Using DEM and DSM Only

As an alternative approach, we attempted to extract features by using the elevation products only. Initial building footprints are extracted from the normalized digital surface model (nDSM) shown in Figure 3.26(a), which is derived by subtracting the DEM from the DSM. We extracted cells whose elevation is higher than 7 ft. (~2.0m). The result shown in Figure 3.26(b) includes building as well as vegetation. In order to further remove cells on vegetation, additional information such as normalized difference vegetation index (NDVI) or texture information need to be further analyzed. However, the resolution of the given DSM is not sufficient to discriminate two features in this test. Furthermore, spectral characteristic and the lack of NIR band of the given image also lead to difficulty discriminating these two features.

In order to remove small or isolated groups of cells, we applied the threshold of minimum area. The remaining cells are considered to represent buildings and their boundaries are determined. For each detected building, its initial boundary (i.e., initial footprint) is determined by tracing the outermost pixels. Potential corner points are then determined based on the Douglas-Peucker (1973) algorithm. In any two consecutive corner points, we found the longest line segment by applying RANSAC algorithm. The remaining line segments are then adjusted with constrains such as parallelism or perpendicularity if corresponding segments are within the predefined tolerance. Figure 3.27 illustrates this process and determined building footprints are overlaid with image data in Figure 3.28.

We also calculated the mean elevation of each building, thereby creating 3-D building models in the test area, which is shown in Figure 3.29.

Figure 3.28 Building footprints.



(a) 3-D Building models overlaid with image

(b) 3-D Building models overlaid with image

(c) Close-up view building models overlaid with image

Figure 3.29 3-D Building models determined from DSM.

4. CONCLUSION

We have investigated the data requirements, available data sources and methods for roadside clear zones and cross-section information extraction for Indiana road network. We have selected a study area representative of non-urban structure since it was initially considered as higher priority.

4.1 Data Availability and Suitability

Existing road network datasets, which we have initially considered using as the reference for extracting information on road geometry and roadside features, did not have the required spatial accuracy to define the road extent and the clear zones. Hence, we have proposed and implemented a framework for extracting the paved surface by classifying the orthophotos, determining the medial axis, estimating the width and then reconstructing a regularized paved surface in order to define the clear zones.

The datasets that we have investigated provide continuity by means of statewide spatial coverage. The orthophotos and the LiDAR datasets are acquired in scope of a statewide program which is expected to be continuously updated if not on a regular basis.

High-resolution CIR orthophotos provide an advantage for classification due to the availability of the IR band at the same spatial resolution as the RGB. SVM classification of the paved surface from orthophotos results with a high classification accuracy which may be considered as an indication of the suitability of high resolution CIR images for the purpose.

LiDAR dataset is not of high point density with its nominal pulse spacing of 1.0-1.5 m. One of the main concerns of such statewide LiDAR acquisition programs is to provide the basic requirement for generating a reliable DEM. Having 5' spatial resolution, the DEM supports the calculation of average slope along a specific direction. It was possible to employ the DEM for calculating average grade of the alignment and side slopes.

4.2 Alignment and Road Elements

We have generated an alignment of the road by classifying the paved surface from the CIR orthophotos and post-processing the classification results to obtain an estimated centerline. SVM Classification provides good classification results with no issues regarding computational time. On the other hand, object-based classification, which also provides good classification results, is computationally costly. Initial segmentation of the images takes several hours even for small tiles of high resolution images. For the road vs. non-road binary classification of our project, SVM classification provides the required computational efficiency since it is a pixel-based classification method.

We have extracted a total of 29.6 miles of road network from the orthophotos. This length is reduced to 21.3 miles after cleaning the irrelevant branches connected to the main roads and small misclassified pieces. The total length of the road network in the given dataset was 23.6 miles. Approximately 10% difference is due to the missing parts in the paved surface

classification results and the additional parts which were removed during cleaning, simplification and generalization process. This difference may be reduced by establishing an expert database of training samples which includes all possible surface types that are required to be extracted in the area. Collecting training samples is the most important part of the framework which requires expert interaction with no comparable alternative with similar outcome. Accounting for each and every type of actual surface in entire Indiana may not be feasible since it will be required to be maintained regularly. However, if the samples are collected on project basis, a near-complete training database may be established which would have an impact on the classification accuracy. Such an effort would reduce only some of the discontinuities since the major amount of gaps in the classification results are due to the trees that are overhanging the roads.

Depending on the main objective, there are two options regarding discontinuities. First option is based on the assumption that the road alignment is completely required for the whole network. This requirement would arise due to any need for querying the road network geometry for specific parameters; e.g., "What are the exact locations where the pavement width is reduced some certain amount?" Unless complete and continuous road network is available, the result for such a query will be incomplete. In order to achieve the completeness requirement, it is possible to obtain the missing parts by using the existing INDOT road network as a reference and updating the sections that are spatially inaccurate utilizing the design principles. It would require additional effort by an operator to complete these missing parts which constituted approximately 10% of the road network in case of the study area.

Second option is to leave the discontinuities as they are. Even though the locations and geometries of the features that are extracted within the roadside clear zones are explicitly available, it is very likely that these features will be used in an aggregate nature per a specified segment length when they are considered as input for a safety measure. Also, the cross sections are sampled at an interval. Assuming that the local properties of the roads do not change rapidly and there is a close enough section which is correctly extracted to compensate for the missing part, such gaps may as well be left as they are unless they are densely localized in one section of the road network.

Classification of the paved surface resulted in the extraction of the main roads as well as branches due to connecting driveways, adjacent parking lots, etc. This information may be ignored as irrelevant or may be utilized by considering them as by-product. However, when considered as the former, they can't be used directly. Further effort is required to filter out irrelevant pieces that do not correspond to any actual branches.

Based on the extracted centerline and the classification of the paved surface, we were able to estimate the paved surface width for each extracted road segment. This estimated width is the width of the surface as observed on the orthophotos. It has not been possible to determine individual lanes or shoulders since the variation of the spectral responses from the available CIR imagery do not allow distinguishing such detail.

4.3 Cross Sections

Based on the estimated paved surface centerline and width, we were able to generate cross-section lines and calculate the slopes along these lines. Even though not expected to happen frequently, it must be noted that calculated slopes may become unreliable in two situations. The first situation happens in case there is a back slope immediately after the foreslope where the immediacy is related to the 5′ spatial resolution of the DEM. An elevation value is available at each 5′ by 5′ raster cell. Interpreting the slope values requires attention if a back slope is expected to start before two samples of elevation values are available on the foreslope.

The second situation is related to the slope calculations of very rapid falls followed by flat land. Slope calculations may be affected by the quickly changing slope values due to averaging. However, we have yet not encountered such situations within the test area.

Average grade along the centerline is also calculated without any issues. In that case, there is no concern for unreliable calculations since the average grade along the centerline is smoothly varying by design.

4.4 Intersections

The thinning algorithm does not result in intersections that coincide at a single node. If the intersections must be conserved topologically, an effort is required to edit them. However, this does not affect the cross-section information since the road extent and clear zones are defined based on the extracted paved surface extent. Similar to the options for whether to complete the missing parts of the alignment, intersections may or may not need to be topologically conserved as a single node. The decision in this case is whether the alignment extracted from aerial imagery will be required to serve as a replacement for a complete network in which the intersections are subject to further analysis, not just simply in aiding the estimation of the approximate paved surface width and reconstructing the surface extent.

4.5 Roadside Information

In consideration of the objectives of our project, the LiDAR dataset provides the capability to extract features within the clear zones at the roadside in a limited fashion, helping to extract the buildings and trees within the clear zones. It is not possible to extract vertically aligned features such as fences, walls, posts, signs, and so forth. due to the limitations of the acquisition technique. Mobile or terrestrial LiDAR

acquisition would be more suitable for extracting such features. Also, extracting individual trees out of a group of trees that are close to each other was not possible using the available dataset. In cases where there is prior knowledge of tree types in the area, individual trees may be estimated based on assumptions regarding the type and structure of those tree types

False positives occur mainly in two cases during the classification of the LiDAR point cloud as buildings and trees. The first group of false positives happens due to the bridges or parts of the roads with steep side slopes. Such parts of the roads may be classified as buildings due to their structural similarity, i.e., high planar surface with respect to the local ground estimation. The second group occurs when the points on parts of the trees form a planar structure similar to the building roofs. This may be filtered out using the NDVI generated from the CIR orthophotos since the spectral response from the trees are different from the response of the buildings.

4.6 Implementation

We have employed and pipelined different software as tools for the implementation of the proposed framework in the study area. Our main concern has been the availability of the algorithms that we have proposed and their convenience for performing quick tests with the existing datasets. Hence, we have not aimed at establishing an operational-level software integration. In summary, we have employed ESRI ArcGIS software for the preprocessing of the orthophotos and DEMs, LAStools for the preprocessing of LiDAR point clouds, Monteverdi interface of Orfeo Toolbox for the classification of the paved road surface from the orthophotos, ArcGIS for the vector-based cleaning and generalization of the classification results, Matlab and Wolfram Mathematica for the morphological cleaning, thinning, pruning, boundary extraction and generalization, ArcGIS for the reconstruction of the paved road surface as well as generating the cross-section lines using Python scripting and also calculating the slopes, LAStools for LiDAR ground filtering and point cloud classification, and R programming language libraries for the delineation of classified points. This pipelining of software is not the only available and certainly not the most efficient way of performing the tasks of the proposed framework. The methods and algorithms mentioned in their respective sections of the report may be combined with a more efficient integration of software tools. Several tools that we preferred to use based on convenience may be eliminated and the tasks performed by these tools may be switched to other tools already used to perform other tasks. As an example, Matlab may be used for implementing SVM classification of the images instead of Orfeo Toolbox, as well as for handling all morphological operations. It will require some additional effort to combine these tasks under one tool that will work at an operational level.

As mentioned previously, the proposed framework is not a fully automated process. Human involvement is required at several steps, most importantly while collecting training samples for classification, which is performed once and then applied throughout the entire study area. Collected samples must represent all occurrences of the pavement types in the study area. Other instances of human involvement include setting the parameters for determining the size of small components, simplification level, buffer sizes for removing the branches from the boundary raster or for calculating the pavement width, ground filtering, point cloud classification, and delineation of the classification results. Once these are determined adequately based on the properties of the study area, they may be used reliably. Finally, human involvement may be required for completing discontinuities as necessary and for editing intersections if they are required to be topologically conserved as a single node.

REFERENCES

Ackermann, F. (1999). Airborne laser scanning—present status and future expectations. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(2–3), 64–67. http://dx.doi.org/10.1016/S0924-2716(99)00009-X

American Society of Photogrammetry and Remote Sensing. (2008). *LAS specification*. Retrieved from http://www.asprs.org/a/society/committees/standards/asprs_las_format_v12.pdf

Baatz, M., & Schäpe, A. (2000). Multiresolution Segmentation: An optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII. Beitrage zum AGIT-Symposium Salzburg 2000* (pp. 12–23). Karlsruhe: Herbert Wichmann Verlag.

Baltsavias, E. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(2-3), 83–94. http://dx.doi.org/10.1016/S0924-2716(99)00014-3

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004

Blaschke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS Zeitschrift für Geoinformationssysteme*, *6*, 12–17.

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent digitized line or its caricature. *Cartographica*, *10*(2), 112–122.

ESRI. (2013). ArcGIS Resource Center, Desktop 10. Available from http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Check_Geometry/001700000034000000/

French Space Agency (CNES). (2013). Orfeo toolbox [Software]. Available from http://www.orfeo-toolbox.org/otb/about-otb.html

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.).Waltham, MA: Elsevier Inc.

Isenberg, M. (2013). LAStools [Software]. Available from http://www.cs.unc.edu/~isenburg/lastools/download/lastools.zip

Haralick, R. M., Sternberg, S. R., & Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *9*(4), 532–550.

Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*(4), 193–203. http://dx.doi.org/10.1016/S0924-2716(98)00009-4

Lam, L., Lee, S., & Suen, C. (1992). Thinning methodologies: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*(9), 869–885.

Najman, L. & Talbot, H. (2010). Introduction to mathematical morphology. In L. Najman & H. Talbot (eds.), Mathematical morphology*: From theory to applications* (pp. 3–33). ISTE: London, UK.

Petrie, G., & Toth, C. K. (2009). Introduction to laser ranging, profiling, and scanning. In J. Shan & C. K. Toth (eds.), *Topographic laser ranging and scanning* (p. 590). Boca Raton, FL: CRC Press.

Pfeifer, N., & Mandlburger, G. (2008). LiDAR data filtering and DTM generation. In J. Shan & C. Toth. (Eds.), Topographic laser ranging and scanning*: Principles and processing* (pp. 307–334). Boca Raton, FL: CRC Press.

Shan, J., & Sampath, A. (2005). Urban DEM generation from raw LiDAR data: A labeling algorithm and its performance. *Photogrammetric Engineering & Remote Sensing*, *2051*(February), 217–226.

Shih, F. Y. (2009). *Image processing and mathematical morphology: Fundamentals and applications*. Boca Raton, FL: CRC Press.

Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *59*(1–2), 85–101.

Soille, P. (1999). Morphological image analysis: Principles and applications. Berlin, Heidelberg: Springer-Verlag.

Theodoridis S., & Koutroumbas, K., (2009). *Pattern recognition* (4th ed.). Boston, MA: Academic Press.

Vosselman, G., (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *33* (B3/2; PART 3), 935–942.

Vosselman, G. (2009). Advanced point cloud processing. In D. Fritsch (Ed.), *Photogrammetric week '09* (pp. 137–146). Heidelberg, Germany: Wichmann.

Wehr, A., & Lohr, U. (1999). Airborne laser scanning–an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(2–3), 68–82. http://dx.doi.org/10.1016/S0924-2716(99)00011-8

About the Joint Transportation Research Program (JTRP)

On March 11, 1937, the Indiana Legislature passed an act which authorized the Indiana State Highway Commission to cooperate with and assist Purdue University in developing the best methods of improving and maintaining the highways of the state and the respective counties thereof. That collaborative effort was called the Joint Highway Research Project (JHRP). In 1997 the collaborative venture was renamed as the Joint Transportation Research Program (JTRP) to reflect the state and national efforts to integrate the management and operation of various transportation modes.

The first studies of JHRP were concerned with Test Road No. 1—evaluation of the weathering characteristics of stabilized materials. After World War II, the JHRP program grew substantially and was regularly producing technical reports. Over 1,500 technical reports are now available, published as part of the JHRP and subsequently JTRP collaborative venture between Purdue University and what is now the Indiana Department of Transportation.

Free online access to all reports is provided through a unique collaboration between JTRP and Purdue Libraries. These are available at: http://docs.lib.purdue.edu/jtrp

Further information about JTRP and its current research program is available at: http://www.purdue.edu/jtrp

About This Report

An open access version of this publication is available online. This can be most easily located using the Digital Object Identifier (doi) listed below. Pre-2011 publications that include color illustrations are available online in color but are printed only in grayscale.

The recommended citation for this publication is:

Shan, J., & Ural, S. (2015). *Performance measure that indicates geometry sufficiency of state highways: Volume II—Clear zones and cross-section information extraction* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2015/07). West Lafayette, IN: Purdue University. http://dx.doi.org/10.5703/1288284315529