GEORGIA DOT RESEARCH PROJECT 14-22 FINAL REPORT EXPLORATION OF USING GDOT'S EXISTING VIDEOLOG IMAGES AND PAVEMENT SURFACE IMAGING DATA TO SUPPORT STATEWIDE MAINTENANCE PRACTICES **Georgia Department of Transportation** OFFICE OF RESEARCH **15 KENNDY DRIVE FOREST PARK, GA 30297-2534**

1.Report No.: GA-16-1422	2. Government Accession No.:	3. Recipient's Catalog No.:		
4. Title and Subtitle: Exploration of Using GDOT's Existing Videolog Images and Pavement Surface Imaging Data to Support Statewide Maintenance Practices		5. Report Date: April 2016		
		6. Performing Organization Code:		
7. Author(s): Zhaohua War	ng; Yichang (James) Tsai	8. Performing Organ. Report No.:		
 Performing Organization Name and Address: Georgia Institute of Technology 760 Spring Street NW, Suite 230 Atlanta, GA 30308-0695 		10. Work Unit No.:		
		11. Contract or Grant No.: RP 14-22		
 12. Sponsoring Agency Name and Address: Georgia Department of Transportation Office of Research 15 Kennedy Drive Forest Park, GA 30297-2534 		 Type of Report and Period Covered: Final; December 2014 – June 2016 		
		14. Sponsoring Agency Code:		
15. Supplementary Notes: Prepared in cooperation with the U.S. Department of Transportation. Federal Highway Administration.				
16. Abstract: Videolog and pavement imaging data is a valuable asset that has supported the Georgia Department of Transportation (GDOT) and enable it to fulfill the requirements of its Highway Performance Monitoring System (HPMS). To maximize the return on investment, GDOT is seeking to utilize the existing videolog and pavement imaging data for extracting roadway asset data that is indispensable for supporting the statewide asset management and maintenance programs. For this purpose, this research project explored the utilization of GDOT's existing videolog and pavement imaging data for extracting guardrails, rumble strips, and traffic signs. Image-processing-based algorithms were developed, which were tested using both				

GDOT's videolog images and the data collected by using the Georgia Tech Sensing Vehicle (GTSV). In the meantime, through the comprehensive investigation and analysis of GDOT's existing data and the exploration results, suggestions were made for data specification and data quality control that can be added to GDOT's existing requirements for future outsourcing contracts. Therefore, the data could be better utilized to support the statewide asset management and maintenance in the future.

17. Key Words: Videolog; 3D Pavement Imaging; Guardrail;	18. Distribution Statement: No Restriction		
19. Security Classification (of this report): Unclassified	20. Security classification (of this page): Unclassified	21. Number of Pages: 101	22. Price:

Contract Research

GDOT Research Project No. 14-22 Final Report

EXPLORATION OF USING GDOT'S EXISTING VIDEOLOG IMAGES AND PAVEMENT SURFACE IMAGING DATA TO SUPPORT STATEWIDE MAINTENANCE PRACTICES

By

Zhaohua Wang, Ph.D., P.E. Yichang (James) Tsai, Ph.D., P.E.

Georgia Institute of Technology

Contract with Georgia Department of Transportation In cooperation with U.S. Department of Transportation Federal Highway Administration April 2016

The contents of this report reflect the views of the author(s) who is (are) responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Georgia Department of Transportation or of the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Acknowledgements

The work described in this final report was supported by the Georgia Department of Transportation (GDOT) research project 14-22. We would like to thank Mr. Timothy Dale Brantley, Ms. Ernay Robinson, and Mr. Larry Barnes from the Office of Maintenance; Ms. Jane Smith (retired), Mr. Paul Tanner, and Mr. Terry Peak from the Office of Transportation Data for their strong support and heavy involvement. We would like to thank the research team, Dr. Chengbo Ai, Ms. Josephine Simon, Ms. Claudia Gargiulo, and Mr. Yifei Fan for their diligent work. We would like to thank Ms. Gretel Sims for her assistance in managing this project.

Table of Contents

Acknowledgements	
Table of Figures	viii
Table of Tables	xii
Executive Summary	xiii
Chapter 1 Introduction	1
1. Research Background and Research Need	1
2. Research Objectives	1
3. Report Organization	2
Chapter 2 GDOT's Videolog and Pavement Imaging Data	3
1. General Information	3
2. Data Utilization	4
2.1 Videolog Images	4
2.2 Pavement Images	6
3. Consideration of Data Reuse	8
4. Selected Test Sections	10
References	11
Chapter 3 Quality Assessment on Videolog images	12
1. Data Analysis	12
2. Data Quality Evaluation Scheme	13
3. Image-Based Quality Assessment	14
3.1 Blurriness	14
3.2 Haziness	18
3.3 Blockiness	22
3.4 Blurriness and Noise	24
v	

3.5	Entropy	26
4. Co	emposition-based Quality Assessment	30
4.1	Camera Configuration	30
4.2	Environmental Deficiencies	31
5. Oł	oject-Based Quality Assessment	32
5.1	Traffic Sign	33
5.2	Guardrail	33
5.3	Rumble Strip	35
6. Su	mmary	36
Refere	nces	36
Chapter 4	Exploration of Guardrail Detection Using Videolog Images	38
1. Int	roduction	38
2. Li	terature Review	39
3. Pr	oposed Algorithm on Guardrail Detection	41
3.1	Preprocessing	41
3.2	Guardrail OM Sign Detection	42
3.3	Guardrail Body Detection	44
3.4	Continuous Integration of Detection	47
4. Ca	se Study	48
5. Su	mmary	50
Refere	nces	50
Chapter 5	Exploration of Rumble Strip Detection Using Videolog Images	51
1. Int	roduction	51
2. Pr	evious Work	53
3. Pr	oposed Algorithms	53

3.1 Preprocessing	54
3.2 Warping	55
3.3 Three Methods for Rumble Strips Detection	56
4. Case Study	61
5. Summary	65
References	65
Chapter 6 Exploration of Sign Detection Using Videolog Images	66
1. Previous Work	66
2. Case Study	70
3. Summary	73
References	73
Chapter 7 Consideration of Data Specification and Quality Checking	74
1. Videolog Images	74
1.1 Image Composition	74
1.2 Image Quality	79
2. 3D Pavement Imaging Data	80
2.1 Data Specification	81
2.2 Data Quality Validation	82
3. Data Format	84
References	85
Chapter 8 Conclusions and Recommendations	86

Table of Figures

Figure 2.1: Stitched Videolog Images	4
Figure 2.2: Image Capturing Interval	5
Figure 2.3: Identifying Location and Measuring Width	5
Figure 2.4: Pavement Imaging Data (Left: Intensity; Right: 3D Elevation)	6
Figure 2.5: Automatic Crack Detection	7
Figure 2.6: Profile, IRI, and Rutting Measurements	8
Figure 2.7: Georgia Tech Sensing Vehicle (GTSV)	10
Figure 3.1: Data Viewer for GDOT Videolog	13
Figure 3.2: The GUI of the Developed Image Quality Assessment Tool	13
Figure 3.3: Examples of Blurry Images	15
Figure 3.4: Flowchart of No-reference Blurriness Estimation (Crete et al., 2007)	16
Figure 3.5: Translation Curve of Estimated Blurriness with respect to Human Perception	18
Figure 3.6: Estimated Blurriness of GDOT and GTSV data on I-20 EB (Carroll County)	19
Figure 3.7: A Hazed Image due to Incorrect Camera Color Parameter	19
Figure 3.8: Estimated Haziness using GDOT and GTSV data on I-20 EB (Carroll County)	21
Figure 3.9: Blocking Effect due to Different Levels of Compression	22
Figure 3.10: Translation Curve of Estimated Blockiness with respect to Human Perception	24
Figure 3.11: Estimated Blockiness using GDOT and GTSV Data on I-20 EB (Carroll County)	25
Figure 3.12: Framework of Proposed Model (Choi et al., 2009)	25

Figure 3.13: Blurriness & Noise Comparison between GDOT and GTSV Data on I-20 EB	
(Carroll County)	. 26
Figure 3.14: Examples of Blurriness and Noise Metric	. 27
Figure 3.15: Images at Different Entropy Level	. 29
Figure 3.16: Entropy Comparison between GDOT Data and GTSV on I-20 EB (Carroll Count	ty)
	. 29
Figure 3.17: Comparison between GDOT and GTSV Data	. 30
Figure 3.18: Examples of Lens Debris	. 32
Figure 3.19: Object-based Assessment – Legibility of Traffic Signs	. 34
Figure 3.20: Object-based Assessment – Legibility of Guardrail	. 35
Figure 3.21: Object-based Assessment – Legibility of Rumble Strips	. 36
Figure 4.1 Guardrail Starting Point and Appearance	. 38
Figure 4.2 Flow Chart of Guardrails Detection Algorithm	. 42
Figure 4.3 Color Segmentation Result	. 43
Figure 4.4 (a) Gradient Magnitude of the reference image (in Figure 4.3 (a)) and (b) extracted	1
ROI	. 46
Figure 4.5 Interference of Passing Vehicle	. 49
Figure 5.1 Continuous rumble strips example	. 51
Figure 5.2 Discontinuous rumble strips example	. 52
Figure 5.3 Rumble Strip Shapes	. 54
Figure 5.4 Operation Flow of Rumble Strip Detection	. 54
Figure 5.5 Initial images for Rumble Strip Detection	. 55

Figure 5.6 Preprocessed Images for Rumble Strip Detection	55
Figure 5.7 Warping Region Definition and Extraction	56
Figure 5.8 Edge Detection and Hough Transform Output	57
Figure 5.9 Output of Method #1	58
Figure 5.10 Output of Method #2	59
Figure 5.11 Grayscale Variation along Cross-section Line of Rumble Strips	60
Figure 5.12 Detected Rumble Strips	61
Figure 5.13 Impact of Image Capturing Location on Rumble Strip Detection	62
Figure 5.14 Impact of Image Quality on Rumble Strip Detection	63
Figure 5.15 Unfavorable Locations of Discontinuous Rumble Strips in Image	64
Figure 5.16 Unfavorable Lighting Conditions for Rumble Strips Detection	64
Figure 5.17 Noise Issue in Rumble Strip Detection	65
Figure 6.1 Generalized Traffic Sign Model for Traffic Sign Detection (Tsai et al., 2009)	67
Figure 6.2 Flowchart of Traffic Sign Detection Algorithm (Tsai et al., 2009)	69
Figure 6.3 Framework of the Enhanced Traffic Sign Inventory (Tsai & Wang, 2013)	70
Figure 6.4 False Positive Caused by Passing Vehicle	72
Figure 6.5 Improper Vertical View Angle	73
Figure 7.1 Camera View Angles	76
Figure 7.2 Impact of Different Vertical View Angles	76
Figure 7.3 Roadside Feature Out of Range	77
Figure 7.4 Impact of Sign Pixel Size on Detection Rate (Tsai, et al., Unpublished)	77

Figure 7.5	A Sign Captured in Consecutive Images	78
Figure 7.6	False Detection Caused by Lane Location	79
Figure 7.7	3D Laser System (Tsai & Wang, 2013)	81
Figure 7.8	Calibration/Validation Board Developed by Tsai, et al.	83
Figure 7.9	Validation on Rut Depth and Cracks	84

Table of Tables

Table 2.1: Covera	age of GDOT Videolog and Pavement Imaging Data (GD	OT, 2012) 3
Table 2.2: GDOT	Required Road Characteristics (GDOT, 2012)	
Table 2.3: Selecte	ed Test Sites on I-20	
Table 4.1 Results	s of Guardrail Detection Algorithm	
Table 5.1 Test D	atasets for Rumble Strips Detection	
Table 5.2 Test R	esults for Rumble Strip Detection	
Table 6.1 Test D	atasets for Sign Detection	
Table 6.2 Test R	esults for of Traffic Sign Detection	

Executive Summary

Videolog and pavement imaging data is a valuable asset that has supported the Georgia Department of Transportation (GDOT) and enabled it to fulfill the requirements of Highway Performance Monitoring System (HPMS). To maximize the return on investment, GDOT is seeking to utilize the existing videolog and pavement imaging data for extracting roadway asset data that is indispensable for supporting the statewide asset management and maintenance programs. For this purpose, this research project explored the utilization of GDOT's existing videolog and pavement imaging data for extracting guardrails, rumble strips, and traffic signs. Image-processing-based algorithms were developed and were tested using both GDOT's videolog images and the data collected by using the Georgia Tech Sensing Vehicle (GTSV). In the meantime, through the comprehensive investigation and analysis of GDOT's existing data and the exploration results, suggestions were made for data specification and data quality control that can be added to GDOT's existing requirements for future outsourcing contracts. Therefore, the data could be better utilized to support the statewide asset management and maintenance in the future.

Chapter 1 Introduction

1. Research Background and Research Need

Roadway asset inventory and pavement condition data are essential to support the statewide asset management and maintenance programs in the Georgia Department of Transportation (GDOT). Due to stringent budgets and insufficient staffing, GDOT and many other state DOTs have adopted the semi-automatic method for conducting roadway asset inventory and pavement condition surveys in which roadway videolog images and pavement surface imaging data are used to automatically or manually extract roadway characteristics, pavement rutting, cracking, etc. In a four-year, \$10 million (master contract) project, GDOT's Office of Transportation Data (OTD) has collected roadway videolog images, pavement surface imaging data, International Roughness Index (IRI) data, and GPS data on the Federal-Aid highway system covering 51,640 centerline miles for which it is responsible. The data has been used to extract roadway characteristics, many of which can be derived to fulfill the requirements of the Highway Performance Monitoring System (HPMS). Nevertheless, the data has not yet been used in other offices, such as the Office of Maintenance (OM). Thus, there is a need to explore the opportunity of adding the valuable data to GDOT's roadway maintenance and asset management programs by utilizing the videolog and pavement imaging data already collected in OTD. In extensive discussions with both OM and OTD, the Georgia Tech research team has identified a research need to explore the utilization of GDOT's existing videolog images and pavement surface imaging data to support statewide maintenance practices.

2. Research Objectives

The objective of this research is to explore the utilization of GDOT's existing roadway videolog and pavement surface imaging data to support statewide maintenance activities. Automatic methods will be developed and tested to 1) explore an automatic, quantitative, and full-scale method for conducting data quality checking; 2) explore the means for extracting guardrails; 3) explore the means for extracting rumble strips; 4) validate a cost-effective means for conducting sign inventory; and 5) suggest factors that need to be considered in data specification. The following are the major tasks:

- Review GDOT's existing roadway videolog and pavement imaging data.
- Determine test cases and select test roadway sections.
- Assess videolog image quality using quantitative indicators.
- Explore the utilization of existing roadway videolog images for guardrail detection.
- Explore the utilization of existing roadway videolog images for rumble strip detection.
- Explore the utilization of existing roadway videolog images for sign inventory.
- Suggest the factors that need to be considered in data quality checking and data specification.
- Summarize research findings and develop a final report.

3. Report Organization

This report is organized into eight chapters. Chapter 1 introduces the research background, need, research objective, and major tasks. Chapter 2 briefly reviews DOT's existing videolog and pavement imaging data. Chapter 3 presents the quality assessment of videolog images. Chapter 4 presents the exploration of guardrail detection using videolog images. Chapter 5 presents the exploration of rumble strip detection using videolog images. Chapter 6 presents the traffic sign inventory using videolog images. Chapter 7 discusses the data specification and quality checking. Chapter 8 summarizes the project, presents conclusions, and offers recommendations for future research and implementation.

Chapter 2 GDOT's Videolog and Pavement Imaging Data

1. General Information

Since 2013, GDOT OTD has collected statewide videolog and pavement imaging data through a four-year contract with an external vendor. The goal was to cover all Federal Aid (FA) eligible roadways within the state of Georgia. As shown in Figure 2.1, there were about 51,640 centerline miles of highways covered (GDOT, 2012).

Roadway	Miles	Data Collection	Subtotal Miles	Centerline Miles
State Routes	18,000 miles	Collected Bi-directionally	18,000 miles	36,000 miles
Non-State Routes	12,000 miles	Collected Uni-directionally	12,000 miles	12,000 miles
Non-State Routes	1,000 miles	Collected Bi-directionally	1,000 miles	2,000 miles
Additionally	Approx. 2500	Ramps and Collector/Distributors		640 miles
Approximate Total Centerline Miles			51,640 miles	

Table 2.1: Coverage of GDOT Videolog and Pavement Imaging Data (GDOT, 2012)

The collected videolog and pavement imaging data are mainly used to extract road characteristics (GDOT, 2014) and fulfill the sample data collection for HPMS. Table 2.2 list the major road characteristics that GDOT requires (GDOT, 2012).

Names of Intersecting	Signalization at	Road names of routes	Through Lanes
Roads	Intersections	being collected	
Medians	Direction	Guardrails	Noise Walls
Storm Drains	Sidewalks	Rumble strips	Striping
Signage	Curbs	Gutters	Mileposts
Railroad Crossings	Bike Lanes	Pavement Transitions	Ride/Share Sites
Weigh Stations	Rest Areas	Welcome Centers	Roadside Parks
Ramps	Gores	Tunnels	Bridges
Causeways	Pedestrian Walkways	Ramps	Access Control
Toll Lanes	HOT Lanes	HOV Lanes	Speed Limit
Route Number	Route Signing	Type Signal	Signal present
Stop Sign Present	At Grade Intersection	Shoulder Left	Shoulder Right
Peak Parking	Widening – Obstacles	Widening – Potential	Curves
Terrain Type	Grades	Percent Passing Sight	

 Table 2.2: GDOT Required Road Characteristics (GDOT, 2012)

Other than road characteristics, GDOT also expected to collect pavement condition data such as IRI, rutting, ridge to valley depth (RVD), digital pavement images, pavement distress data for asphalt and concrete pavements, geometry, and cross slopes. Nevertheless, constrained by the total amount of the master contract (\$10 million), GDOT is unable to acquire all the above data.

2. Data Utilization

Currently, GDOT OTD has hosted all the videolog images in a data server. However, the downward pavement imaging data is hosted by the vendor. A dedicated data viewer software provided by the vendor can be used to explore the data. The following subsections briefly review all the major functions and the provided information.

2.1 Videolog Images

Videolog images capture the front view of roadways. At each location, three images are captured, which can be stitched together using software, as shown in Figure 2.1. Each image has a resolution of 2,456Figure 2.1: Stitched Videolog Images×1,200. It can be seen that the left and right shoulders and the entire pavement are covered. Thus, the images can be used to extract all roadway and roadside features, such as signs, guardrails, and pavement geometry.



Figure 2.1: Stitched Videolog Images

Using the software, the image capturing time/frame and GPS locations can be viewed. As shown in Figure 2.2, the times/frames (with red underscores) associated with four consecutive videolog images show that videolog images (three images at a location) are captured every 10 or 11 frames. Based on NTSC video format, 30 frames are captured in one second. Thus, images were captured about every 1/3 second. From Figure 2.2, it can be estimated that the average distance between two image-capturing locations is about 26 ft. (8 meters). So, the vehicle speed

could be estimated by the distance the one traveled, 26 ft. in 1/3 second, i.e., 53 mph. If the vehicle speed changes, the distance interval also changes.



Figure 2.2: Image Capturing Interval

The software also provides functions to measure pavement width and identify point locations (milepoint or latitude/longitude) as shown in Figure 2.3. It should be noted that the measurements closer to the camera are more accurate. This function is very useful for locating point features, such as signs, starting and ending points of guardrails, etc.



Figure 2.3: Identifying Location and Measuring Width

The videolog images use the Joint Photographic Experts Group (JPEG) format and can be easily read using other software, and, thus, can be used to automatically extract some roadway assets, such as signs, guardrails, and rumble strips. However, it was found that the location information (milepoint and latitude/longitude) could not be exported for each image. The only way is to use the above software and manually check the location information for each image. For large-scale datasets, this method is very cumbersome and almost infeasible. Thus, as suggested in Chapter 7, the open access to image locations is desirable.

2.2 Pavement Images

GDOT also collected 3D pavement imaging data. As shown in Figure 2.4, the left image is the laser intensity image; the right one is the 3D elevation data. The technology of 3D laser triangulation was adopted to acquire the pavement surface elevations. The advantages of using 3D laser data are twofold:

- Pavement cracks are easier to detect with the use of 3D elevations. Shadows and oil marks cannot be shown on 3D elevation data;
- Other elevation-related distresses, such as rutting, raveling, IRI, and macrotexture, can be detected and/or measured using the same set of data.



Figure 2.4: Pavement Imaging Data (Left: Intensity; Right: 3D Elevation)

The software provides function to automatically detect pavement cracks. Figure 2.5 shows an example. However, the detection results are not satisfactory. The detected cracks in Figure 2.5

are actually all false positives, i.e. falsely detected cracks. The two straight cracks are most likely flat-tire scratches, and all other short cracks are noise. For practical application, tremendous manual effort is needed to remove all these false detections.



Figure 2.5: Automatic Crack Detection

Other than crack detection, the software also provides various profile information, IRI, and rutting measurements. This information is useful for investigating project-level or segment-level pavement conditions using an interactive tool. However, it is cumbersome for network-level data management.

Unlike videolog images, pavement imaging data is stored in a proprietary format. That means it can only be viewed using proprietary data viewer software. It is infeasible for GDOT to extract pavement distresses (e.g. cracks) using software or services from other vendors, which will dramatically limit the use of the raw data. With the growing interest from state highway agencies, more and more vendors are manufacturing hardware and providing data services, but they normally develop and use proprietary software and technology to process, display, and report the collected data. Using proprietary formats for storing 3D pavement imaging data creates the challenge of meeting transportation agencies' different data requirements and makes it difficult to unify data analysis, reporting, sharing, and evaluation outside the vendor community. Thus, the Federal Highway Administration (FHWA) has initiated a project to develop a common and interchangeable data format for pavement imaging data. It is reasonable

for GDOT to request an open format of pavement imaging data in future outsourcing contracts with external vendors. Thus, GDOT will have better flexibility when choosing vendors for roadway feature and pavement distress extraction.



Figure 2.6: Profile, IRI, and Rutting Measurements

3. Consideration of Data Reuse

GDOT has invested about \$7.5 million thus far in the past four years to acquire videolog and pavement imaging data and have the vendor extract various road characteristics and pavement condition data through ten task orders. Currently, the raw data is mainly used by the contracted vendor to extract road characteristics and pavement condition data. If the data can be reused by

and benefit other offices in GDOT to support different businesses, the return on investment would be maximize.

Since transportation assets change over time, the investment on videolog and pavement imaging data is recurring. For example, the service life of an asphalt pavement is about 10 years, during which its condition changes continuously. To determine the right treatment at the right time, it is critical to monitor pavement conditions at a sufficiently small time interval, e.g. 1 year. The same situation exists for other asset, such as traffic signs, guardrails, etc. Therefore, cyclic data (e.g. videolog and pavement imaging data) collection is needed. It would be very valuable if the raw data were well prepared for reuse, which could be achieved by establishing a more rigorous data specification in future outsourcing contracts.

In GDOT's master contract, many data items were listed. However, not all of them can be acquired because of the constrained amount of the contract. It would be more economical if the existing raw data could be reused. One way for data to be reused is to utilize existing raw data for extracting other transportation assets. This could be done internally by GDOT or externally by vendors. Currently, a state highway agency has to stick to the contracted data service provider for other derived data services (e.g., road characteristics extraction) due to the use of proprietary data formats. This situation makes it difficult for contract negotiation and hinders broader adoption of videolog and pavement imaging data for transportation asset data collection. Consequently, a more open data format is urged so that highway agencies can use videolog and pavement imaging data to save labor and improve data accuracy. If a more open data format is implemented, more vendors with stronger capabilities in data processing could get involved in promoting the use of sensing technology, which would reduce the cost for highway agencies. The project initiated by FHWA ("Standard Data Format for 2D/3D Pavement Image Data") is endeavor to do this.

In addition to data format, data quality is an important factor that will determine how effectively data can be reused. The quality of videolog and pavement imaging data can be assessed by evaluating how well it can be used for feature extraction, such as traffic signs, guardrails, etc. A specification is needed for data quality requirements in the outsourcing contract.

9

In this research project, the videolog and pavement imaging data will be explored to assess the applicability of using existing data for extracting various roadway assets. The assessment will be conducted by evaluating automatic detection of guardrails, rumble strips, and traffic signs. After that, a specification for data format and data quality will be recommended for GDOT. It is hoped the specification could be applied in future outsourcing of data collections.

4. Selected Test Sections

With the help from the OM and OTD in GDOT, six test sections were selected on I-20 in three counties: Cobb, Douglas, and Carroll County. In each county, test sections were selected in two directions, eastbound and westbound. There were 54.2 lane miles. Videolog and pavement imaging data were, also, exported for the selected test sections.

County/County No	Route No	Route Suffix	Milepoint From	Milepoint To	Notes
Cobb/067	0402	00	0	3.65	I-20 EB; AC
Cobb/067	0402	00	3.65	0	I-20 WB; AC
Douglas/097	0402	00	8.9	18.87	I-20 EB; AC
Douglas/097	0402	00	18.87	8.9	I-20 WB; AC
Carroll/045	0402	00	2.68	16.16	I-20 EB; PCC
Carroll/045	0402	00	16.16	2.68	I-20 WB; PCC

Table 2.3: Selected Test Sites on I-20

For comparison purposes, the videolog images on the selected test sections were also collected by using the Georgia Tech Sensing Vehicle (GTSV) (see Figure 2.7).



Figure 2.7: Georgia Tech Sensing Vehicle (GTSV).

References

- GDOT (2012). Request for Proposals to Provide Videolog Digital Imaging Services for Georgia Department of Transportation. RFQ-RFP 484-031212
- GDOT (2014). Road Characteristics Extraction Guide. Georgia Department of Transportation Office of Transportation Data.

Chapter 3 Quality Assessment on Videolog images

In this chapter, the quality of videolog images is discussed in order to establish a fundamental understanding of the characteristics in GDOT's existing videolog images (for brevity, it is called "GDOT data" hereafter in this chapter). Based on the fundamental understanding of the image characteristics, the potential utilization of the data can be recommended to support statewide maintenance practices of critical roadway assets, e.g., traffic sign, guardrails, rumble strips, etc. Several quantitative indicators are recommended for inclusion in image quality specifications.

1. Data Analysis

The GDOT data is in the format of a JPEG image sequence and organized by a proprietary data viewer software. Figure 3.1 shows the Graphic User Interface (GUI) of the software. The image sequence is referenced to a proprietary Global Positioning System (GPS) trajectory file, where, at each triggering location, three images were captured and associated with unique GPS coordinates, as shown in the bottom left of the GUI in Figure 3.1. The image trigger interval is possibly based on time, about every 1/3 second, as discussed in Chapter 2. The three images covering three camera views, including front left, front center, and front right, are stitched together to create a panoramic view, as shown in the upper part of the GUI in Figure 3.1. The image sequence containing a series of folders containing JPEG files is the only data that can be accessed externally, whereas the detailed GPS coordinates for each image is not exportable and cannot be externally accessed. Each image file has a resolution of 2,456×1,200 and a depth of 24 bits using conventional JPEG compression technique; the detailed compression information is not available.

For comparison purposes, the videolog images on the selected test sites (see Chapter 2) were also collected by using GTSV (the term of "GTSV data" is used hereafter in this chapter). The GTSV data used in this chapter compiles two images with a resolution of 2,448×2,048, a depth of 24 bits and a compression rate of 80%. The images are stitched together to create a panoramic view comparable to the GDOT data. The GTSV data in this study is captured at an interval of 5 m (16.4 ft.). The GTSV data is spatially associated with the GDOT data by matching the manually retrieved GPS locations from the proprietary data viewer. To facilitate the data quality

assessment and comparison between the GDOT data and the GTSV data, an assessment tool was developed, including a side-by-side view of the corresponding GDOT and GTSV data, the basic image information of the images, and the assessment indicators. The GUI of the developed assessment tool is shown in Figure 3.2.



Figure 3.1: Data Viewer for GDOT Videolog



Figure 3.2: The GUI of the Developed Image Quality Assessment Tool

2. Data Quality Evaluation Scheme

This section presents the data quality assessment scheme for the GDOT data. The GTSV data was evaluated using the same scheme for comparison purposes. Three assessment approaches

are proposed for the scheme to comprehensively evaluate the data quality, including image-based assessment, composition-based assessment, and object-based assessment. The scheme is designed so that both general characteristics of an image and the details of targets are considered. The details of each component are presented in the following sections:

- The image-based assessment aims at evaluating the quality of the entire image region. Five key no-reference indicators, i.e. blurriness, haziness, blockiness, noise, and entropy are selected and computed to represent the relative quality of the data. These quality indicators are further compared with the GTSV data.
- The composition-based assessment aims at the applicability of the images with respect to the task, i.e. extracting roadway features in the regions of interest (ROI). Factors that determine the applicability include camera intrinsic parameters, camera configurations, and environmental deficiencies, which are evaluated manually.
- The object-based assessment aims at evaluating the quality of certain image regions that correspond to objects of interest, e.g., traffic sign, guardrails, etc. The legibility of objects of interest are manually evaluated, including the legibility of the pictogram and characters on traffic signs, the legibility of the signs of the object marker (OM) at the beginning of guardrails, and the visibility of rumble strips.

3. Image-Based Quality Assessment

Due to the nature of image files, a general assessment approach for image files can be applied to GDOT data. The quality is evaluated with five no-reference indicators, i.e. blurriness, haziness, blockiness, noise, and entropy.

3.1 Blurriness

Blurry images are typically captured due to the loss of camera focus and/or camera vibration in conjunction with insufficient shutter speed during the data collection. The details of the images that construct distinctive image features of roadway objects may be potentially undermined due to the blurry images. Figure 3.3 shows progressive examples of the blurry effect due to different levels of camera vibration during the data collection. It can be observed that the blurriness has a

strong impact on the quality of the image, particularly on the legibility of the key features of the objects of interest, e.g., edges, characters, icons, etc. Therefore, the blurriness is one of the key indicators for videolog image quality that require evaluation.



Figure 3.3: Examples of Blurry Images

However, since baseline data (i.e., perfect data without any blurring) is impossible to acquire during data collection, a no-reference technique for computing blurriness (Crete et al., 2007) is introduced in this study. The derived indicator from this technique will provide a quantitative measure for the blurriness, which is subjectively calibrated based on human's perception. Figure 3.4 shows the flowchart of the method. Given an original image, if the variation between the original image and its derived image using artificial blurring operation is high, the blurriness of the original image is relatively low because the blurring operation destructs the sharpness of the original image more significantly. On the contrary, if the variation between the original image and its derived image using artificial blurring operation is low, the blurriness of the original image is relatively high because the blurring operation destructs the sharpness of the original image is relatively high because the blurring operation destructs the sharpness of the original image is relatively high because the blurring operation destructs the sharpness of the original image is relatively high because the blurring operation destructs the sharpness of the original image is relatively high because the blurring operation destructs the sharpness of the original image is relatively high because the blurring operation destructs the sharpness of the original less significantly.

To blur the original image, a low-pass filtering on both horizontal and vertical directions is applied to mimic the combined physical phenomena of the loss of camera focus and the camera vibration. Mean kernels h_v and h_h are used for the low-pass filtering in both directions, and the artificially blurred images, B_V and B_H , are the directional convolution of the original image I(x,y)with the corresponding kernels.



Figure 3.4: Flowchart of No-reference Blurriness Estimation (Crete et al., 2007)

$$h_{v} = \frac{1}{9} [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1], h_{h} = \frac{1}{9} [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$$

$$B_{V} = h_{v} * I, B_{H} = h_{h} * I$$

The variations within the original image, D_{IV} and D_{IH} , and the derived image from the artificially blurred images, D_{BV} and D_{BH} , are then computed using the local differentiation in both horizontal and vertical directions.

$$D_{I_V}(i,j) = \frac{dI}{dx} = |I(i,j) - I(i-1,j)|$$
$$D_{I_H}(i,j) = \frac{dI}{dy} = |I(i,j) - I(i,j-1)|$$
$$D_{B_V}(i,j) = \frac{dB}{dx} = |I(i,j) - I(i-1,j)|$$
$$D_{B_V}(i,j) = \frac{dB}{dy} = |I(i,j) - I(i,j-1)|$$

At each pixel (i, j), the variation between the original image and the derived image from the artificial artificially blurred images, D_{VV} and D_{VH} , are then computed as follows:

$$D_{V_V}(i,j) = \max(0, D_{I_V}(i,j) - D_{B_V}(i,j))$$
$$D_{V_H}(i,j) = \max(0, D_{I_H}(i,j) - D_{B_H}(i,j))$$

The overall variation of the whole image is computed by accumulating the pixel-wise variations within the original image and the pixel-wise variation between the original image and the derived image from the artificially blurred images as follows:

$$D_{I_VALL} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} D_{I_V}(i,j)$$
$$D_{I_HALL} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} D_{I_H}(i,j)$$
$$D_{V_VALL} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} D_{V_V}(i,j)$$
$$D_{V_HALL} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} D_{V_H}(i,j)$$

Finally, the estimated blurriness is defined as the maximum between the normalized variations along the vertical and horizontal directions as follows:

$$M_{blur} = \max(\frac{D_{I_VALL} - D_{V_VALL}}{D_{I_HALL}}, \frac{D_{I_HALL} - D_{V_HALL}}{D_{I_HALL}})$$

The estimated blurriness indicator, M_{blur} , represents the relative blurriness of the image ranging between 0 and 1. The value close to 0 represents an image with minimum blurring effect, whereas the value close to 1 represents an image with heavy blurring effect. A translation metric (Crete et al. 2007) is further defined to correlate between the estimated blurriness value and the human perception. Figure 3.5 shows the translation curve and the definition of the human perception category.

Using the proposed method, the GDOT data and the GTSV data collected on I-20 within Carroll County, east bound, was used to evaluate the performance. Figure 3.6 (a) shows the estimated blurriness of all the frames in the GDOT data and the GTSV data, represented by the red and blue curves respectively. The red background indicates a "Perceptible/Not Annoying" level of blurriness, and the green background indicates an "imperceptible" level of blurriness. It can be

observed that majority of the GDOT data shows a "Perceptible/Not Annoying" level of blurriness as compared with an "imperceptible" level of blurriness from the GTSV data. Figure 3.6 (b) shows a visual comparison of blurriness between the GDOT data and the GTSV data.



Figure 3.5: Translation Curve of Estimated Blurriness with respect to Human Perception

3.2 Haziness

Hazy images are typically captured due to aerosols suspended in the natural medium, e.g. dust, mist, etc., or the combination of incorrect camera color parameters, reversed lighting, and/or camera malfunctioning during the image acquisition. The details of the images that construct distinctive image features of roadway objects may be potentially undermined due to the image haze. Figure 3.7 shows an example of the hazy effect due to incorrect camera color parameters during the data acquisition and the corresponding normal image. It can be observed that haze has a strong impact on the quality of the image, particularly on the overall contrast and visibility of the objects of interest containing achromatic colors, e.g., guardrails (gray color), speed limit signs (black and white color), etc. Therefore, the haziness is another key video log image quality that requires evaluation.

Similar to the blurriness, baseline data (i.e., perfect data without any haze) is difficult to acquire during data collection; a no-reference technique for computing haziness (Guo et al., 2014; Sun et al., 2015) is introduced in this study. The derived indicator from this technique will provide a quantitative measure for the haziness. The atmospheric light image is first estimated to represent

the mask of the haze for the original image using the Retinex theory (Land & McCann, 1971). Then, the magnitude of the haziness is quantified by the accumulated intensity of the atmospheric light image.



(a) Estimated Blurriness for GDOT and GTSV Data



(b) Visual Comparison of Blurriness between GDOT and GTSV Data

Figure 3.6: Estimated Blurriness of GDOT and GTSV data on I-20 EB (Carroll County)



Figure 3.7: A Hazed Image due to Incorrect Camera Color Parameter

In computer vision, a general mode in describing the intensities of a hazy image is established by Narasimhan and Nayar (2003):

$$L(x, y) = L_0(x, y)e^{-kd(x,y)} + A(1 - e^{-kd(x,y)})$$

where L is the apparent luminance of the captured image, d(x,y) is the distance of the corresponding object at a scene point (x,y) from the camera, and A is the global atmospheric light constant, and $L_0(x,y)$ is the haze-free image. According to the atmospheric scattering model, haze reduces the visibility and contrast of the object in the scene by taking two effects, including the exponential decay by $e^{-kd(x,y)}$ and a white atmospheric veil by $A(1 - e^{-kd(x,y)})$. Without prior information, such as distance of the corresponding object and a scene point, the global atmospheric light constant, the magnitude of the hazing effect, can be hard to quantify following the Narasimhan and Nayar (2003) model. However, several assumptions can be made to simplify the problem, while only relative magnitude is needed, including:

- The exponential decay e^{-kd(x,y)} on the haze-free image can be ignored, assuming the scene depth d(x,y) is the same although unknown;
- White atmospheric veil A(1 e^{-kd(x,y)}) can be estimated purely based on the apparent luminance by assuming lower intensity will be perceived by the camera from the object from farther away, i.e., A(1 e^{-kd(x,y)}) = A(1 αL(x,y));
- The global atmospheric light constant *A* can be estimated based on the retinex theory by assuming the global atmospheric lighting constant *A* is equivalent to the illumination image in the retinex theory (Land & McCann, 1971) that can be estimated using Gaussian smoothing.

Therefore, the relative haziness can be represented by the equation below. A lower value represents an image with less hazing effect, while a higher value represents an image with more hazing effect.

$$M_{Haze} = A - A \cdot \alpha \cdot L(x, y)$$

Where

$$A = h * L(x, y), h = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

h is confined within a $w \times w$ window in the convolution. Because the haziness indicator is a relative value in this study whose magnitude does not represent any physical meaning, the constant α and σ can be arbitrarily selected. In this study, $\alpha = 1$ and $\sigma = 5$.

Using the proposed method, the GDOT data and the GTSV data collected on I-20 within Carroll County was used to evaluate the performance. Figure 3.8 (a) shows the estimated haziness for all the frames in the GDOT data and GTSV data (represented by the red and blue curves, respectively). The results clearly show that the GDOT data bears a more significant level of hazing effect compared with the GTSV data. Moreover, the results also show that the GTSV data bears a more consistent hazing effect throughout the complete data collection session compared with the GDOT data. Figure 3.8 (b) shows a visual comparison of haziness between the GDOT data and the GTSV data.



(a) Estimated Haziness for GDOT and GTSV Data



(b) Visual Comparison of Haziness between GDOT and GTSV on I-20

Figure 3.8: Estimated Haziness using GDOT and GTSV data on I-20 EB (Carroll County)

3.3 Blockiness

Blocky images are typically produced by heavy JPEG compression during data acquisition or post-processing. Due to the heavy compression, the block artifacts become visible and may overwhelm the key features of the objects of interest captured in the image. Figure 3.9 shows examples of the progressive blocking effect due to different levels of compression. It can be observed that the blocks have a strong impact on the quality of the image, particularly at the edges of the objects of interest, e.g., sign boundaries, guardrail stands, etc. Therefore, the blockiness is another key video log image quality that requires evaluation.



Figure 3.9: Blocking Effect due to Different Levels of Compression

The blocking effect appears in JPEG compressed images because JPEG compression is a block DCT (Discrete Cosine Transform)-based lossy image coding technique. The quantization operation is independently applied to the DCT coefficient in each 8×8 block, which produces discontinuous images along the block boundaries. To estimate such a blocking effect, an empirical method that correlates the actual block measurements from non-referenced images and the mean opinion score determined by manual review, is established.

Three measurements are computed from the non-referenced image, including the average difference across block boundaries B_h and B_v , the average absolute difference between in-block image samples A_h and A_v , and the horizontal zero-cross rate Z_h and Z_v .

$$B_h = \frac{1}{M(\lfloor N/8 - 1 \rfloor)} \sum_{i=1}^{M} \sum_{j=1}^{\lfloor N/8 - 1 \rfloor} |d_h(i, 8j)|$$

$$A_{h} = \frac{1}{7} \left[\frac{1}{M(\lfloor N/8 - 1 \rfloor)} \sum_{i=1}^{M} \sum_{j=1}^{N-1} |d_{h}(i,j)| - B_{h} \right]$$
$$Z_{h} = \frac{1}{M(N-2)} \sum_{i=1}^{M} \sum_{j=1}^{N-2} z_{h}(i,j)$$

Where

$$z_{h}(m,n) = \begin{cases} 1 & \text{horizontal zero cross at } d_{h}(m,n) \\ 0 & \text{otherwise} \end{cases}$$

Similarly, the vertical measurements, including B_v , A_v , and Z_v , can be computed in the same format. The average values of B_h and B_v , A_h and A_v , and Z_h and Z_v are then further computed from the horizontal and vertical directions. Finally, the blockiness indicator can be represented as a composite value of the averaged measurement values:

$$M_{Block} = \alpha + \beta \cdot B^{\gamma_1} \cdot A^{\gamma_2} \cdot Z^{\gamma_3}$$

Where α , β , γ_1 , γ_2 , and γ_3 are the model parameters that require calibration. In this study, the calibrated parameters by Zhou et al. (2002) are directly used for evaluating the blockiness of the image, including $\alpha = -245.9$, $\beta = 261.9$, $\gamma_1 = -0.0240$, $\gamma_2 = 0.0160$, and $\gamma_3 = 0.0064$.

The estimated blurriness indicator M_{block} represents the relative blurriness of the image ranging between 1 and 10. The value close to 10 represents an image with minimum blockiness, whereas the value close to 1 represents an image with heavy blockiness. A translation metric (Zhou et al., 2002) is further defined to correlate the estimated blockiness value and the human perception. Figure 3.10 shows the translation curve and the definition of the human perception category.

Using the proposed method, the GDOT data and the GTSV data collected on I-20 within Carroll County was used to evaluate the performance. Figure 3.11 (a) shows the estimated blockiness of all the frames in the GDOT data and the GTSV data and is represented by the red and blue curves, respectively. The red background indicates a "Slightly Annoying" level of blockiness; the yellow background indicates a "Perceptible/Not Annoying" level of blockiness; and the green background indicates an "imperceptible" level of blockiness. It can be observed that majority of
the GDOT data shows a "Perceptible/Not Annoying" level of blockiness, though some of the frames fall into a "Slightly Annoying" level, compared with an "imperceptible" level of blockiness from the GTSV data. Figure 3.11 (b) shows a visual comparison of blockiness between the GDOT data and the GTSV data.





3.4 Blurriness and Noise

Most digital cameras have an Image Signal Processor (ISP) to enhance the output of image sensor. One of the important functions of an ISP is to remove noise. The strong noise reduction removes noise sufficiently but makes detail and texture blurred. Since a trade-off exists between noise reduction and detail loss, criteria considers both are required. The method proposed by Choi et al. (2009) was adopted in this study, with its framework illustrated in Figure 3.12.

In this method, the blurriness is estimated by the difference between the intensity of current pixels and the average of neighbor pixels, which is then normalized by the average. If the intensity of the center pixel is closer to the average intensity of both side pixels, the center pixel is supposed to be on a blurred edge. As for noise, the method only measures the noises out of the edge, as they look more apparent than those along edges.



(a) Estimated Blockiness for GDOT and GTSV Data



(b) Visual Comparison of Blockiness between GDOT and GTSV Data

Figure 3.11: Estimated Blockiness using GDOT and GTSV Data on I-20 EB (Carroll

County)



Figure 3.12: Framework of Proposed Model (Choi et al., 2009)

Using the proposed method, the GDOT data and the GTSV data collected on I-20 within Carroll County was used to evaluate the performance. On average, GTSV scores higher than GDOT data on this metric, as shown in Figure 3.13.



Figure 3.13: Blurriness & Noise Comparison between GDOT and GTSV Data on I-20 EB (Carroll County)

According to the metric, a higher score is accompanied by better image quality. However, the conclusion is not always true for our dataset because asphalt pavements themselves are relatively noisy and inconsistent. A higher score, which implies less noise, may actually come from the uniform intensity caused by a large shadow, while small pieces of shadows may lead to lower scores. Figure 3.14 shows the results of representative image samples. Therefore, we will need to combine this metric with other criteria for comprehensive analysis.

3.5 Entropy

In information theory, entropy (more specifically, Shannon entropy) is the expected value (average) of the information contained in each message, where "messages" can be modeled by any flow of information. For JPEG images, the intensity on the red, green, and blue channel is expressed using 8 bits (a byte). For each dimension, there are $2^8 = 256$ possible values that a pixel can take. Ideally, the maximum entropy is achieved when the intensity follows the uniform distribution, $H = -\sum_{i=0}^{255} \frac{1}{256} \log_2 \frac{1}{256} = 8$.





score = -6.094, low score, low quality

score = -2.460, high score, low quality



score = -5.077, low score, good quality score = -2.946, high score, high quality

Figure 3.14: Examples of Blurriness and Noise Metric

However, in a real case, the bandwidth of intensity can be relatively narrower. As a result, the image contains less information than the ideal case. When it comes to display, the image has lower contrast and seems monotonous.

We calculate the source entropy of each image and connect the results to the image quality obtained from our observation. Typical examples can be found in Figure 3.15.



H = 4.061, low quality

H = 4.908, low quality



H = 5.885, low quality



H = 6.862, medium quality



H = 6.785, medium quality



H = 7.164, medium quality



H = 7.560, high quality



H = 7.770, high quality

Figure 3.15: Images at Different Entropy Level

Using the proposed method, the GDOT data and the GTSV data collected on I-20 within Carroll County was used to evaluate the performance. On average, GTSV data has a higher entropy than GDOT data, as shown in Figure 3.16. The sudden drop for GTSV data in Figure 3.16 results from sun glare and dark, which will be explained in next section.



Figure 3.16: Entropy Comparison between GDOT Data and GTSV on I-20 EB (Carroll County)

4. Composition-based Quality Assessment

Videolog images are often used by transportation agencies to extract roadway assets using automatic or manual methods. In this section, the quality of images are evaluated in terms of the capability of data extraction. Although the proposed image-based assessment method can objectively quantify the overall image quality, it does not represent the applicability of the data to support roadside asset data collection. Therefore, certain regions of the images require further investigation. In this context, the ROI is the only thing that is related to data extraction. Two aspects of an image are to be evaluated: camera configuration and environmental deficiencies.

4.1 Camera Configuration

There are five major concerns with respect to camera configuration: resolution, aspect ratio, shooting angle, white balance, and capture interval. Higher resolution would result in more pixels for the ROI, which means that the 2D approximation is more vivid and close to real-world entities. Figure 3.17 shows the comparison between GDOT and GTSV data; both images are captured at the same location. The proportion of the ROI in an image varies with the aspect ratio. As shown in Figure 3.17, the GTSV data focuses more on the roadway, whereas the GDOT Data contains more information of the surroundings.



GDOT Data





Figure 3.17: Comparison between GDOT and GTSV Data

The proportion of the ROI in an image also varies with different shooting angle. In GDOT data, the viewing angle is almost the same as the driver's, while the GTSV focuses more on roadway by looking down at a certain angle. According to the perspective rule, an ideal position for the ROI would be at the lower part of the canvas, which makes it larger. Therefore, when signs are the targets, GDOT data has a better shooting angle than GTSV data; however, when targets become rumble strips, GTSV data outperforms GDOT data.

As shown in Figure 3.17, there is an apparent distinction in white balance between the two data sets in that GDOT data is relatively dim and contains an excessive yellow tone compared to the GTSV data. Therefore, the contrast is reduced in the GDOT data because the color bandwidth is narrower. Moreover, the increase in color distortion may potentially bring challenges to color-based feature extraction.

Different mechanisms are adopted for interval control in capturing the video log images. For GDOT data, the intervals between two adjacent images are approximately 10 frames (30 frames = 1 second) in time. The intervals between two adjacent images are 5 m (16.4ft.) for GTSV data. Thus, the distance between two consecutive images of GDOT data depends on the speed of the vehicle, which is not fixed.

4.2 Environmental Deficiencies

The Alberta DOT identified five types of poor videolog images: sun glare, dark, lens debris, color corruption, and fog, among which the first three deficiencies can be found in GDOT and GTSV data (Tsai & Huang, 2010).

An image with sun glare, as the first image shows in Figure 3.15, is an image captured in strong sunshine or upon entering sunshine from a darker zone (e.g. under a viaduct bridge). A dark image, as the upper right image shows in Figure 3.15, is an image captured with insufficient light or at the entrance to a darker zone from a brighter one. The above two deficiencies can take place when data collection is conducted on a sunny day, especially when the exposure level is not adjusted properly. Sun glare and dark are the major deficiencies in GTSV data, whereas GDOT data survives because the data is collected on cloudy days.

An image with lens debris, as shown in Figure 3.18, is an image partially occluded by dirt or other lens imperfection/obstructions, making parts of the image unavailable/unclear to reviewers. Although no such dominant lens debris can be found in GDOT data, there is a light vertical line at the middle of the images for data on I-20 in Carroll County.



Figure 3.18: Examples of Lens Debris

5. Object-Based Quality Assessment

In this section, the ROIs are further investigated in detail by identifying key assets that are to be collected, such as traffic signs, guardrails, rumble strips, etc. The evaluation of the quality in these image regions will not only provide insight on the applicability of the data to support roadside asset management in the current practice, i.e., manual inventory and condition evaluation, but will also provide valuable insight for recommending the applicability of the data in future practices, i.e., automatic or semi-automatic inventory and condition. Subsequent chapters will present how the GDOT data can be immediately utilized given the current quality and the recommendations on how the data quality can be improved in the future data collection. In this subsection, the image legibility of the objects of interest, including traffic signs, guardrails, and rumble strips, is manually reviewed to reveal the qualitative indication of the GDOT data. The GTSV data is, also, studied for comparison purposes.

The examples of traffic signs, guardrails and rumble strips are shown in Figure 3.19, Figure 3.20, and Figure 3.21, respectively. In general, it can be observed that the legibility of traffic signs, guardrails, and rumble strips captured by the GDOT data are not as legible as the ones in

the GTSV data. Such an observation is consistent with the results derived from the image-based assessment. The legibility issues observed in the GDOT data can be attributed to three major reasons:

- Heavy JPEG compression creates excessive artifacts and heavily distorts the original image;
- Some of the image acquisition parameters, i.e., shutter speed, lens focus, exposure, and color processing, etc., might not be optimized;
- Some of the camera configuration parameters, i.e., camera orientation, acquisition interval, etc., may not be optimized.

The following are some of the detailed evaluations for specific objects of interest.

5.1 Traffic Sign

From the traffic sign inventory and condition evaluation perspectives, the key properties of a traffic sign, including location, type, dimension, and condition are required. Therefore, the features of a traffic sign corresponding to the key properties are required to be legibly captured by videolog images, including boundary (image edge), color, character, and pictogram, etc. However, as shown in Figure 3.19, many of these features are not clearly captured in the GDOT data (left side); in comparison, these features are better captured in the GTSV data (right side).

5.2 Guardrails

From the guardrail inventory and condition evaluation perspectives, the key properties of a guardrail include location and condition. Therefore, the features of a guardrail corresponding to the key properties are required to be legibly captured by videolog images, especially the guardrail edge. However, as shown in Figure 3.20, the guardrails captured in the GDOT data do not contain a clear edge in the image (left side). In addition, the OM sign that delineates guardrails is a unique feature that may be introduced in the future automatic/semi-automatic operations; however, it presents a very vague pattern with blurred edges and distorted yellow color. As a comparison, the right side images come from GTSV data.



Figure 3.19: Object-based Assessment – Legibility of Traffic Signs



Figure 3.20: Object-based Assessment – Legibility of Guardrail

5.3 Rumble Strips

From the rumble strip inventory and condition evaluation perspectives, the key properties of rumble strips include location and the condition. Therefore, the features of rumble strips corresponding to the key properties are required to be legibly captured by videolog images, especially the alternated pattern of the rumble bumps. As shown in Figure 3.21, it is hard to see rumble strips in any of the frames, and it is challenging to capture the alternated pattern due to the inadequate image quality and the camera orientation. For comparison purposes, the images on the right side come from GTSV data.



Figure 3.21: Object-based Assessment – Legibility of Rumble Strips

6. Summary

This chapter evaluated image quality from three aspects: image-based, composition-based, and object-based. Five quantitative indicators were recommended for image-based quality assessment. A prototype assessment tool was also developed. For full-scope image quality checking (QC), the computation speed can be improved by using parallel computing, such as GPU and computer clusters. Unlike the commonly used visual inspection and sampling method, these five indicators provide an objective, automatic, and full-scope alternative for quality checking of videolog images. Other than the image-based quality assessment, two other assessments are also recommended, which are based on how videolog images are used for roadway feature extraction. It should be noted that the image-based quality assessment. However, the later assessments are complements to the image-based assessment. In the current study, only visual assessments were investigated by using composition-based and object-based image QC. Further study is needed to develop the corresponding automatic approach.

References

Crete, F., Dolmiere, T., Ladret, P., and Nicolas, M. "The blur effect: perception and estimation with a new no-reference perceptual blur metric." 64920I-64920I-64911.

- Guo, F., Tang, J., and Cai, Z.-X. (2014). "Image Dehazing Based on Haziness Analysis." *Int. J. Autom. Comput.*, 11(1), 78-86.
- Land, E. H., and McCann, J. J. (1971). "Lightness and Retinex Theory." *J. Opt. Soc. Am.*, 61(1), 1-11.
- Narasimhan, S. G., and Nayar, S. K. (2003). "Contrast restoration of weather degraded images." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(6), 713-724.
- Sun, W., Wang, H., Sun, C., Guo, B., Jia, W., and Sun, M. (2015). "Fast single image haze removal via local atmospheric light veil estimation." *Computers & Electrical Engineering*(0).
- Zhou, W., Sheikh, H. R., and Bovik, A. C. "No-reference perceptual quality assessment of JPEG compressed images." *Proc., Image Processing. 2002. Proceedings. 2002 International Conference on*, I-477-I-480 vol.471.
- Choi, M. G., Jung, J. H., & Jeon, J. W. (2009). No-reference image quality assessment using blur and noise. International Journal of Computer Science and Engineering, 3(2), 76-80.
- Shannon, C. E. (2001). A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), 3-55.
- Tsai, Y., & Huang, Y. (2010). Automatic detection of deficient video log images using a histogram equity index and an adaptive Gaussian mixture model. Computer - Aided Civil and Infrastructure Engineering, 25(7), 479-493.

Chapter 4 Exploration of Guardrail Detection Using Videolog Images

This chapter explores the use of videolog images for guardrail detection. GDOT has a detailed guideline for guardrail inspection. By using videolog images, an inspector needs to view images frame by frame. After a guardrail is found, a closer inspection will be performed to collect the detailed information. This process could be very tedious and time-consuming because the portion of guardrails is small for the entire highway system. To avoid missing any guardrails, the inspector has to view all image frames from the starting point to the ending point of the entire highways. Thus, if the locations of all the guardrails were automatically identified, an inspector could just view a much smaller portion of total videolog images. For this purpose, this chapter explores the application of image processing and computer vision technology for automatic guardrail detection.

1. Introduction

Guardrails are railings or strong fences set on the edge of the road or in the middle of an expressway in order to prevent vehicles from straying to off-limit areas and causing serious accidents. Guardrails can, therefore, be found on both the left and right sides of roadways. They usually consist of horizontal rails and periodic, vertical posts. The distance between posts can vary.



Figure 4.1 Guardrail Starting Point and Appearance

For automatic detection purposes, the following characteristics of guardrails could be useful:

- Theoretically periodic "T"-like structure;
- Very little texture (mostly smooth);
- Guardrails are usually of a grayish color, very much like the road itself. The aspect can vary along the guardrail due to illumination conditions;
- Guardrails follow the curve of the road: usually they have a straight section, but they can also be curved;
- Guardrails can be either short or very long.
- Guardrails usually start with an OM sign, as shown in Figure 4.1.

In this study, we will focus on detecting the guardrail that is situated on the direct edge of the current road (right side). In order to do this, we will study the images taken from a camera directed slightly to the side of the driving direction. Notice that the same algorithm can be applied on the left-side guardrails just by mirroring the input image, using the hypothesis that position and orientation of the camera are (secularly) the same.

2. Literature Review

Not many articles can be found on guardrail detection. However, all of them seem to use some type of depth information in addition to the normal image-based information. For instance, Shcarwarchter et al. (2014) used stereo cameras, while Alessandretti et al. (2007) used laser data. The algorithms they describe always rely at one point on the straight and linear geometrical aspect of the guardrail.

In the study conducted by Shcarwarchter et al. (2014), a Sobel filter was first applied to the image to get its contours. Then, the Hough transform was used in order to retrieve the lines contained in the image. These lines were validated with respect to the depth data (the depth along the line should be linear). From these lines, ROIs were defined using the 3D information. Finally, each region was classified as guardrail or not, by using a bag-of-features approach. The

bag-of-features were created from the SIFT descriptor and pooled together according to their height in the 3D image (in order to have some geometrical ordering).

The study conducted by Alessandretti et al. (2007) used guardrail detection as a way to simplify vehicle detection. Guardrails were looked for in areas where the computed velocity (from a radar) was low, i.e. less than 5 m/s. A preprocessing step was applied using a Sobel filter: only edges with an orientation approximately matching a guardrail were considered. Then, a line-searching method was performed on either the left or right side of the image: the extremal edge pixel was searched and, then, if an almost continuous line starting from it was detected, the region was labeled as guardrail. Otherwise, the vehicle detection was performed in the region that was actually the aim of the method.

The approach applied by Seibert et al. (2013) is completely different, since the author's aim was to detect the road and its limits by describing its texture using Local Binary Patters (LBP), which is a texture-based area classification. They then classified the side of the road as soft shoulder, curb, or guardrail. In order to detect guardrails, some 3D information was extracted from the monocular camera using ego-motion and by tracking Harris corners via optical flow (with the assumption that the camera was only moving forward). For this, the camera needed to be calibrated.

Finally, Danescu et al. (2006) presented a method for side lane and guardrail detection using 3D information obtained from the stereovision and the detected geometry of the lane. Once the lane was identified and modeled, the lane side could be inferred from the model, and then the guardrail was searched according to 3D points' properties (mainly the height). A tracking step was added in the end to increase stability.

In our proposed method, some ideas from the above-mentioned references were utilized, in particular about the geometrical shape characterizing a guardrail in a 2D image. The big improvement was that no other methods/inputs needed to be added since the detection is only based on a stream of 2D images. The guardrail is tracked in the image by detecting its starting point and then continuing the detection according to the results of the previous image detection. In this way, the equipment cost for detection and the computational complexity of the method are drastically reduced.

3. Proposed Algorithm on Guardrail Detection

Based on the facts that 1) most of the time a guardrail starts with an OM sign (as shown in Figure 4.1, and 2), the rest of the guardrail is mostly linear and continuous, characterized by strong, vertical gradients on its top edge. The algorithm consists of two sides correspondingly.

- 1) Image-based detection:
 - a. Detection of potential guardrail signs;
 - b. Detection of linear segments of high vertical gradients in the image (referred to as "guardrail bodies").
- Detection by continuity along images: the above features are assembled in a continuous manner in order to detect a whole guardrail.

The flow chart of this approach is shown in Figure 4.2. An assumption on which the method is based is that the 2D-image input to the method is sampled from a video stream collected from equipment used in a moving vehicle. The only requirement concerns the camera orientation, which is supposed to be slightly toward the right side of the road. It is also preferable that the collecting vehicle is positioned in the rightmost lane so that other vehicles won't interfere with the detection.

3.1 Preprocessing

As a first step, some preprocessing is applied to optimize the condition of the image used for detection. The top third of the image is discarded taking into consideration the camera orientation. Thus, the background sky and part of the vegetation or other elements near the road are removed. Therefore, they will not interfere with the detection. Then, an equalization of the image is performed separately to enhance three color channels. The basic idea of the equalization step is to re-distribute the image intensity histogram over the entire available range, according to the initial distribution. This step is of primary importance because it often happens that the searched guardrail is shadowed by vegetation and not properly visible. By equalizing the image, this problematic case is solved, making the color better defined. In the case of an image with good lighting, instead, the equalization will have almost no effect, and this allows the next steps of the algorithm, including the color segmentation, to be applied within optimal conditions.



Figure 4.2 Flow Chart of Guardrails Detection Algorithm

3.2 Guardrail OM Sign Detection

• Step 1: Color Segmentation

In the first step, we want to narrow down the search area and be able to identify connected regions that are, potentially, guardrail OM signs. In order to do so, we consider one of the two most important constant features of guardrail OM signs: their color. Indeed, guardrail OM signs are always composed of yellow and black stripes. Although black is very difficult to characterize as a "color," yellow can be segmented out of an image. In order to do so, we define a range of interest in the first two channels of the HSV ("Hue-Saturation-Value") space. HSV is a popular space for color segmentation, since it separates intensity from color. In order to segment the yellow parts of the image, we apply two sets of thresholds on both the hue and the saturation channel. The following thresholds were chosen: HminY = 20; HmaxY = 40; SminY =

100; SmaxY = 240. With these thresholds, we can compute a binary image with points of interest (note that we only consider the lower two-thirds of the image) to minimize the computation cost without risking missing any guardrail sign). A few small morphological operations allow us to isolate regions of interest from the binary image, as shown in Figure 4.3 (b).



(a) Input Image (b) Segmentation (c) Region Identification

Figure 4.3 Color Segmentation Result

These areas of interest give us bounding boxes that might contain a guardrail road sign. The given bounding box examples are drawn in blue in Figure 4.3 (c). Note that to minimize missed detection, the previous thresholds were chosen as being very unrestrictive. However, this also means that a wider range of color may be detected (like orange or red).

• Step 2: Classification

Now that we have defined a few candidates for guardrail OM signs, we can train a classifier to separate and identify the real signs. To do this, we will make use of the specific texture of these signs. To account for these, we choose to take a descriptor inspired by the HOG (Histogram of Oriented Gradients). This seems like a reasonable choice, since a guardrail OM sign will have a very structured gradient distribution, while a region that does not contain a guardrail sign will tend to be less structured. Indeed, ideally, all high gradients on a sign will be in the same direction, which should show in the HOG. Since a guardrail OM sign is only one small coherent structured region, we feel that a simpler version of the HOG can be applied here. Instead of subdividing the considered region into cells and blocks, we choose to compute six different histograms of oriented gradients that we concatenate into one feature vector:

• One for each quarter of the bounding box

• One for half of the bounding box (left/right)

The histograms are computed by considering nine equally-distributed bins over 0-180 degrees, and each histogram is normalized. Each pixel in the considered area votes for the bin corresponding to its gradient orientation, weighted by its gradient magnitude. In total, we therefore have a 54-long feature detector.

From this feature vector, we train a random forest classifier with 100 trees to separate guardrail OM signs from others. A 10-fold validation on 2,558 different images (1,279 guardrail OM signs, 1,279 others) gives us a total confusion matrix of $\begin{bmatrix} 1,142 & 137 \\ 15 & 1,266 \end{bmatrix}$, which gives us 10% of FP and 1.1% of FN. The training was done by assigning a cost for misclassification of real guardrail signs 20 times higher than the one for misclassifying a non-guardrail sign. Because of this, the results still contain a lot of FP. The number is even increased by the fact that we run on an entire data set; the thresholds we use in segmentation tend to catch many regions that do not have guardrails. This choice was made because we know that the next step will allow us to validate the real guardrails and discard others.

3.3 Guardrail Body Detection

To find the whole body of the guardrail, we rely on its linear edges. We assume that the side of the guardrail (at least the top part) will show strong vertical gradients. In order to use this characteristic, we will segment the image in areas of interest that contain such gradients.

We start by computing the gradient magnitude and direction of the image. Blobs are extracted from the gradient study by keeping continuous areas of strong edges with correct orientation (around 80 degrees for the data set being test (as shown in the Figure 4.4). In order to do this, we compute an image combining the magnitude (m(i, j)) and orientation ($\theta(i, j)$) of the gradients into one image *G* with the following formula:

$$G(i,j) = m(i,j) \cdot e^{-\frac{\left(\theta(i,j) - \theta_{ref}\right)^2}{\sigma^2}}$$

where θ_{ref} is the targeted gradient angle (here 80 degrees), and σ is a parameter representing the accepted variance of the described distribution, empirically set to 20. Threshold is applied to this combined image, and a couple of morphological operations are done in order to extract the ROI.

A last step is then added to filter the extracted ROIs. The main direction of each extracted area is computed (with PCA), and only the linear ones with correct orientation are kept (we thus get rid of regions that have a main direction that is either too horizontal or too vertical). In addition, a filtering on the average color is performed in order to remove possible regions defined, for instance, by the white lanes marking the roadside.

In details, this filtering step is performed by computing the eigenvectors and eigenvalues of the covariance matrix of an N×2 matrix of the coordinates (row, column) of the N points constituting each of the extracted ROIs. The main orientation of the region, thus its direction and intersection point with the right side of the image, is defined by the eigenvectors; the inertia, that is a parameter describing the shape/linearity of a region, depends on the eigenvalues, as shown in the following equations:

$$m = -\frac{Var(r)}{Var(c)}; \ q = \frac{Var(r)Mean(r) + Var(c)Mean(c)}{Var(c)} - \frac{Var(c)}{\#cols \cdot Var(r)}; \ i = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

where

- *m* is the slope of the main direction of the region, and *q* is its intersection with the right side of the image;
- *i* is the inertia of the region;
- r is the column vector with the list of the rows coordinates of the points in the region;
- c is a column vector with the list of the columns coordinates of the points in the region;
- Var() and Mean() are variance and mean operators;
- *#cols* is the number of columns of the image under study;
- λ_1 and λ_2 are respectively the smaller and bigger eigenvalues of the covariance matrix.

The second filtering is performed on the average color of the region: $color = \frac{\sum_{i=1}^{N} c_i}{N}$, where *c* is the intensity value of the pixel i and N is the total number of pixels in the region considered. The

computed average color needs to be smaller than an empirically determined threshold, according to the fact that the guardrail body is gray, and it can get darker because of lighting problems, such as shadows produced by surrounding objects. The threshold is determined as the maximum of the average intensity of a sample set of several guardrails tested. This second filtering is important because it allows us to discard several regions that otherwise would look like a guardrail for the rest of the method.

After the above-described steps, as shown in Figure 4.4 (c), each image has a series of possible guardrail directions (green line) with their corresponding detected areas, and possible guardrail sign detection (small red rectangle).



Figure 4.4 (a) Gradient Magnitude of the reference image (in Figure 4.3 (a)) and (b) extracted ROI.

The ROIs are defined according to certain criteria:

- [slope, intercept]: defining the position and angle of the main orientation of the area detected. The intercept is defined as the intersection point between the line of main orientation and the right side of the image;
- size: defines the length of the region along the previously-mentioned orientation;
- inertia: the percentage of inertia is explained by the main orientation (the higher this value, the more linear the area);
- points: the list of the points in the area.

Notice that all the described steps are performed on an image equalized on the all three color channels. This results in an accurate detection disregarding the lighting conditions of the image, for example caused by vegetation's shadows.

3.4 Continuous Integration of Detection

With these detections, we are able to create a detection model for each guardrail detected. This model will be updated in each image and will hold the detected start and end images of the guardrail. The model of a guardrail is defined by:

- [slope, intercept]: defining the position and angle of the main orientation of the guardrail.
 The intercept is defined as the intersection point between the line of the main orientation and the right side of the image;
- points: the list of the points of the detected gradient regions associated with this guardrail in the last image;
- score: the confidence score associated with the guardrail (a number between 0 and 1);
- [beginning, end]: frame numbers for the beginning and end of the guardrail.

Initialization of guardrail

A guardrail model is initialized in two cases. First, when both a sign and an associated gradient ROI (as defined above) are detected. An ROI is considered as being associated with a sign if it is situated in the top/left of the sign, close to the sign, and with a main direction that intersects the sign. Second, if a region is detected satisfying some conditions of color, dimension, and inertia, the threshold values used for this purpose are defined as strictly characterizing only a guardrail in order to avoid false positive cases.

The initial slope, intercept, and points are set to those of the ROI. The initial score is set to 1. The beginning and end frame numbers are set to the current frame.

Update of a guardrail

At each frame, the update algorithm starts by looking for associations between the ROIs detected in the new image and the guardrail models inherited from the previous image. A correspondence between a guardrail model and an ROI is noted if

- Their slope and intercept are close enough and part of their areas (defined by the points parameter) overlap,
- Or, their slope and intercept are almost identical (this allows to recover the guardrail if there are gaps in the detection due to illumination or part-time occlusion).

If new ROIs are associated with a current guardrail model, it is updated in the following way:

- Its **slope** and **intercept** are set as a weighted average of its previous slope and intercept and the ones from the new ROIs. A fixed weight (0.3) is set for the previous model parameters, while the rest (0.7) is distributed to the new ROIs with respect to their size.
- The **points** are replaced by the union of the ones of the new ROIs.
- The score at instant t is computed as a weighted sum (with a ceiling at 1):

$$\circ \qquad score(t) = \alpha \cdot score(t-1) + \beta \cdot sz$$

- where *sz* is the cumulative size of the ROIs divided by the width of the image
- The end frame is set to the current frame.

If no ROIs are associated with the guardrail, the slope, intercept, points, and end-frame parameters stay unchanged. The score is computed as shown previously with the total size of ROIs equal to zero.

End of a guardrail

If the score goes under 0.3, the current end frame is considered as the end frame for the guardrail, and the model is deleted after saving the result.

4. Case Study

To evaluate the performance of the proposed algorithms on guardrail detection, two sets of videolog images data were selected on I-20 from milepost 12 to 45. One set of data was provided by GDOT; as a comparison, another set of data was collected using the GTSV. Because the image capture intervals are different in the two sets of data, the number of images are also different. The algorithms were tested using two data sets. The image-based results are listed in Table 4.1.

#Data set	Size	# positive samples	FP	FN	Precision	Recall	Accuracy
GTSV	17,545	4,863	0.4%	2.4%	99.5%	97.6%	99.1%
GDOT	11,879	3,592	13.7%	12.3%	84.9%	86.2%	88.8%

Table 4.1 Results of Guardrail Detection Algorithm

It can be seen that the proposed algorithms performed very well on the GTSV data with an accuracy of 99.1%. However, they performed worse on the GDOT data with an accuracy of 88.8%. In the meantime, the percentages of false positive (FP) and false negative (FN) cases on the GDOT data are above 10%, which are 0.4% and 2.4% for the GTSV data, respectively. Based on the test, the proposed algorithms are very promising for automatic guardrail detection as long as the quality of videolog image is satisfied.

The worsening in the performances of the GDOT data set is due to several issues. First, the image quality is lower than the GTSV data used. Moreover, the camera positioning and orientation are not optimal because the data is often recorded from the middle lane (instead of the outer lane), and passing vehicles interfere with the detection (as seen in Figure 4.5). This causes many FP (for instance, detection of a truck's sides) and FN cases (when the guardrail is covered by something else).



Figure 4.5 Interference of Passing Vehicle

5. Summary

In this chapter, the utilization of the videolog images in automatic guardrail detection was explored. The proposed algorithms performed very well on the GTSV data with an accuracy of 99.1% and acceptably on GDOT data with an accuracy of 88.8%. Thus, applying automatic guardrail detection to guardrail inventory is very promising. Though it is still a challenge to automatically collect all the detailed information for a guardrail, it can potentially save an inspector's time by automatically removing all irrelative videolog images that contain no guardrail. In comparison with GTSV data, there is still some room to improve the image quality of GDOT data. In addition, for purposes of collecting roadside assets, the data collection vehicle should drive in the outer lane instead of the middle lane so that the interference of passing vehicles can be minimized.

References

- Alessandretti, G., Broggi, A., and Cerri, P. (2007). Vehicle and Guard Rail Detection Using Radar and Vision Data Fusion, IEEE Transactions on Intelligent Transportation Systems, Vol. 8, No. 1.
- Danescu, R., Sobol, S., Nedevschi, S., and Graf, T. (2006). Stereovision-based Side Lane and Guardrail Detection, IEEE Intelligent Transportation System Conference.
- Seibert, A., Hahnel, M., Tewes, A., and Rojas, R. (2013). Camera based detection and classification of soft shoulders, curbs and guardrails, IEEE Intelligent Vehicles Symposium.
- Shcarwarchter, T., Schuler, M., and Franke, U. (2014). Visual Guard Rail Detection for Advanced Highway Assistance Systems, IEEE Intelligent Vehicles Symposium (IV).

Chapter 5 Exploration of Rumble Strip Detection Using Videolog Images

This chapter explores the use of videolog images for rumble strip detection. Like guardrails, rumble strips are another critical safety feature on highways. GDOT has no comprehensive inventory for rumble strips until now. However, the manual, in-field inventory and condition assessment is very labor intensive and time-consuming. Thus, it is worthwhile to explore an automatic method.

1. Introduction

Rumble strips are very important road safety features to alert drivers when they drift from their lanes. They produce a tactile vibration and audible rumbling transmitted through the wheels into the vehicle interior. In this study, only rumble strips on the right shoulder applied along the travel direction and parallel to the road's edge line are considered (see Figure 5.1).



Figure 5.1 Continuous rumble strips example

A rumble strip can vary in color, dimensions (both depth and width), and thickness of the groove. And it can contrast with the surrounding area. Moreover, it can be a continuous strip (Figure 5.1) or small sets of strips at a fixed interval (Figure 5.2).



Figure 5.2 Discontinuous rumble strips example

The following main characteristics of rumble strips are useful for developing automatic detection algorithms:

- Width: the groove can have a length up to several inches.
- Depth: depths vary, as do colors, too.
- Length: rumble strips can be imagined as a sequence of rectangles whose base is the width and height is the length; such heights can vary.
- Continuity: the strips can be continuous for, even, several miles; strips can be made of short sets of lines spaced several feet apart.
- Color: rumble strips are generally the same color as the roadway; they can be darker or lighter than the roadway, depending on the road materials and the depth of the grooves.
- Spacing in length: the distances between rectangles composing the rumble strips can vary.
- Spacing from the marker line: the rumble strips can be right next to the marker line or farther from it by some number of inches.

In this study, the detection of shoulder rumble strips will be focused, which are located on the right of the road. For this purpose, videolog images taken from a camera slightly directed to the right side of the driving direction are used.

2. Previous Work

To the best of our knowledge, there is no existing literature about using 2D videolog images to detect rumble strips. In a U.S. patent (Gordon et al., 2010), the inventors claimed that the presence of rumble strips can be detected based on the outcome of the frequency-based analysis of the wheel speed data, which was obtained from a wheel speed sensor. The wheel speed data can be modified before conversion to the frequency domain to reduce wheel-induced cyclic variation in wheel speed. The frequency-based analysis can use a fast Fourier transform (FFT) and a peak detection method that analyzes one or more peaks in the FFT data to determine if any are indicative of the presence of a rumble strip. The method can be carried out automatically in real time and used to alert the driver of the detection of the rumble strips. In our proposed method, only 2D videolog images are used to explore the applicability of the existing data in GDOT.

3. Proposed Algorithms

In this study, a set of semi-automatic algorithms, which are based only on 2D images, are proposed for rumble strip detection.

The following assumptions are made in this study. First, the images used for detection are all taken from the outer lane of the road, slightly facing the camera on the right; the camera orientation and positioning are fixed for the entire data stream. It was found that it is very difficult to detect rumble strips if images are taken from other lanes that are farther from shoulder. Second, the only two possible types of rumble strips considered in this work are the continuous and discontinuous shoulder rumble strips shown in Figure 5.1 and Figure 5.2. Third, the only two cases considered concerning were shaped as ribs or grooves, as shown in Figure 5.3.



Figure 5.3 Rumble Strip Shapes

The flow chart of the proposed method is illustrated in Figure 5.4.



Figure 5.4 Operation Flow of Rumble Strip Detection

3.1 Preprocessing

A rumble strip set consists of grooves in the pavement. Both the rumble strip, that is the object of interest to be detected, and the surrounding background are tones of gray. This allows a simplification of the problem of data collection by converting the image from RGB to grayscale (from three to one-dimensional space) without losing any information for the detection.

Then, an equalization of the entire image is performed in order to be better handle lighting problems, such as images that are too dark or too light and shadows. Moreover, smoothing with

a Gaussian filter is applied to reduce noise effects. Figure 5.5 shows two example input images, and Figure 5.6 shows the corresponding images after preprocessing.



Figure 5.5 Initial images for Rumble Strip Detection



Figure 5.6 Preprocessed Images for Rumble Strip Detection

3.2 Warping

Image warping is the process of digitally manipulating an image to apply some kind of distortion, such as, in the case for rumble strip detection, a homographic mapping of coordinates. The idea is to identify the ROI in the image by manually selecting four points on the input image and to map the described quadrilateral into a rectangle. This process might introduce some distortion, but it also adds some constant feature in the newly defined image for study, such as fixed shape and number of pixels. The limitation of this step is that the points to warp on are manually defined, and so the process is not fully automatic. Moreover, the results will depend on the region selected on the input image and its orientation and shape. With the above-mentioned assumption that camera position and orientation are fixed, the warping area selection could be a

process done only once on all sets of analyzed images, allowing a partially automated process. Figure 5.7 shows two cases for image warping with manually selected areas.



Figure 5.7 Warping Region Definition and Extraction

After the warping region is defined and extracted, three methods will be used to justify the possibility of rumble strips. Finally, a composite metric is defined for final decision.

3.3 Three Methods for Rumble Strips Detection

Based on the characteristics of rumble strips, three methods are incorporated in the proposed algorithms for rumble strip detection.

Method #1: Lines

Method #1 looks for stripe-shaped rumble strips. It is based on edge detection and line detection in the warped rectangular image.

The first step is to enhance the contrasts by applying a sharpening filter. This way, the following edges detection, performed using the Canny method with an optimized threshold, will give better results. On the binary image of the edges, some morphological operators are then applied to clean the result from noise and better shape the identified regions. The resulting image is used as an input to the line detection step based on a modified Hough method. By definition, the Hough method for line detection is based on the mapping from each point in the x-y coordinate space to a set of lines in the r- θ space (where r and θ are polar coordinates). An accumulation matrix is used to store votes produced by each x-y edge point in the r- θ space, and the maxima of such matrix will define the radius and slope of the lines in the image. For the purpose of rumble strip detection and according to the assumption already described, the method can be modified by reducing the angle range in which -30° and 30° are used in the proposed algorithms. Thus, only the almost horizontal lines will be detected, and the method will work better in terms of computation speed (the standard Hough method could be very slow for big images). For instance, if white lane marking appears in the image, the standard Hough method would detect it, altering the results of the detection, while the proposed algorithms still works properly. To avoid computational time problems, the number of detected lines in the output of the Hough method is fixed to 20.



Figure 5.8 Edge Detection and Hough Transform Output

If a stripe-shape rumble strip is present, the expected lines are almost parallel, as shown in Figure 5.9.



Figure 5.9 Output of Method #1

Method 2: Rectangles

Method #2 looks for rectangle-shaped rumble strips. The preprocessing step is different from Method #1, so there is no need to enhance the contrasts, but there is a need to smooth the image first to get more uniform regions, no matter the irregularity in the pavement intensity distribution. A color segmentation in then applied using the Otsu method to determine the applied threshold; the binary image to work on is the result of the segmentation with the smallest number of pixels in it. Since the rectangles in the rumble strips are sometimes lighter and other times darker than the background pavement, the segmentation will be above or below the threshold according to the smallest set. Again, a morphological operator is applied to clean the image from noise.

The connected components in the image are now extracted, and some of their properties are checked and compared to some empirically determined limits for the problem. Such properties are the orientation of the region that should be almost horizontal, the eccentricity, and the area and its extent that should be reasonable for a big-enough rectangle. The thresholds on the parameters are carefully selected in order to avoid detection of noise. The detection result will

be positive if some rectangle is detected. A bounding box is plotted around the detected rectangles, as illustrated in Figure 5.10.



Figure 5.10 Output of Method #2

Method #3: Grayscale Oscillation

Different from the previous two methods, which look at the latitude of rumble strips, Method #3 looks at the oscillation of gray level along the cross-section line using an orthogonal angle. Ideally, it is expected that the grayscale curve should be similar to square wave as rumble strips and gaps alternate along the line. In real cases, however, the curve is more similar to sinusoids, with an example shown in Figure 5.11.

The idea of Method #3 is to measure the similarity between the curve and sinusoids with the following criteria or dimensions: 1) number of wide waves per pixel, 2) proportion of wide waves; 3) number of vibrations with large amplitude, and (4) proportion of large-amplitude vibrations. The more similar to sinusoids, the more likely that rumble strips exist along the cross-section line. The duration and amplitude of each ascending trend and descending trend are counted, and then the number of trends with higher durations and the number of trends with larger amplitudes are counted.
For the warping area defined by the user, 10 cross-section lines separated at equal distance are picked. The similarity to sinusoids are calculated; the average of 10 scores is considered as the final score for Method #3.





Figure 5.11 Grayscale Variation along Cross-section Line of Rumble Strips

Final Decision

Each of the above three methods will result in a score. The final detection result is determined by the highest score, which should also be greater than an empirical threshold. The score of each method is defined as the ratio between the number of detected objects and a predefined number of objects expected to be detected when rumble strips present. For the stripe-shaped case in Method #1, the predefined number of expected objects is 20, representing the number of lines the Hough method is looking for; for Method #2, it is 10, i.e., 10 rectangles presented for rectangleshaped rumble strips. The empirical threshold for a score is set as 0.5. It means that rumble strips are detected by a method if the corresponding score is greater than 0.5. In Method #3, a wide wave is defined if its duration is longer than eight pixels; a high wave is defined if its amplitude is greater than 15. Then, if rumble strips are present, it is expected that there are 8 wide waves per 100 pixels and 50% of all waves are wide; in the meantime, it is expected that there are 10 high waves per 100 pixels and 50% of all waves are high. Figure 5.12 shows two detected rumble strips.



Figure 5.12 Detected Rumble Strips

4. Case Study

The proposed algorithms were tested using both GTSV data and GDOT data. As mentioned above, if data collection vehicle does not run in the outer lane, rumble strips are difficult to detect using videolog images. Figure 5.13 shows an example. The left image was collected by GTSV in the outer lane; the right one was collected by GDOT in the middle lane. The continuous rumble strips can be detected using the GTSV image. However, they are hard to see in the GDOT image because they were collected from the middle lane. Thus, if videolog images are not collected in the outer lane, such as the GDOT data collected on I-20 EB and WB in Douglas County, they are excluded for testing in the case study. For each dataset, a common warping region is defined at the beginning, which will be used for the entire dataset. Table 5.1 lists the test datasets for rumble strip detection.

Table 5.2 lists testing results on these six test datasets. Roughly speaking, the detection accuracy on continuous rumble strips (datasets #5 and #6) is much better than the one on discontinuous rumble strips (datasets $\#1\sim\#4$). Considering the same data source from the GTSV, datasets #1 and #2 have accuracies of 50.8% and 51.4%, respectively; datasets #5 and #6 have accuracies of 81.9% and 95%, respectively. The proposed algorithms works well on continuous rumble strips. However, they do not work well on discontinuous rumble strips.



Figure 5.13 Impact of Image Capturing Location on Rumble Strip Detection

Dataset Number	1	2	3	4	5	6
GTSV Data	I-20 WB Carroll County	I-20 EB Carroll County	-	-	I-20 WB Douglas County	I-20 EB Douglas County
GDOT Data	-	-	I-20 WB Carroll County	I-20 EB Carroll County	-	-
Type of Rumble Strips	Discontinuous	Discontinuous	Discontinuous	Discontinuous	Continuous	Continuous

 Table 5.1 Test Datasets for Rumble Strips Detection

 Table 5.2 Test Results for Rumble Strip Detection

Dataset Number	Number of Images	ТР	FP	TN	FN	Precision	Recall	Accuracy
1	2,907	38.1%	23.3%	12.8%	25.9%	62.1%	59.5%	50.8%
2	3,793	37.5%	20.7%	13.8%	28.0%	64.4%	57.3%	51.4%
3	2,371	2.9%	8.3%	19.4%	69.4%	25.9%	4.0%	22.3%
4	2,377	16.7%	15.7%	6.3%	61.3%	51.5%	21.4%	18.8%
5	2,408	78.8%	0.7%	3.1%	17.4%	99.1%	81.9%	81.9%
6	2,113	78.0%	1.0%	17.0%	4.0%	98.7%	95.1%	95.0%

The accuracy results from the datasets #1 and #2 are consistent (50.8% and 51.4%); they are also consistent on the datasets #3 and #4 (22.3% and 18.8%). However, the results from #1 and #2 outperform the ones from #3 and #4. The difference comes from the different image quality for rumble strip detection. Figure 5.13 shows an example. The left image was collected by the GTSV. The discontinuous rumble strips were detected. However, the same rumble strips are hard to see in the right image.





Figure 5.14 Impact of Image Quality on Rumble Strip Detection

Further investigation was performed to identify the issues of rumble strip detection. The following are some problematic situations that could lead to false detection.

Unlike continuous rumble strips, discontinuous rumble strips show as a small cluster of horizontal grooves. Its location in an image determines how well it can be detected. Figure 5.15 shows that if the cluster is near the top or at corner of the image, the proposed algorithms do not work well. The ideal location is shown in the left image in Figure 5.14, which is hard to achieve because videolog images are captured at a certain interval. This possibly explains the inferior performance of the proposed algorithms on discontinuous rumble strips. A possible solution is to increase the density of videolog images.

In homogeneous lighting conditions, whether too bright as sun glare or too dark as shadows, can produce unfavorable factors can lead to false detection. Figure 5.16 shows two cases of unfavorable lighting conditions. It can be seen that rumble strips are hard to see in each of these two cases. One solution is to choose a better time for data collection. A sunny day is not a good time, since sun glare and shadows are inevitable.



Figure 5.15 Unfavorable Locations of Discontinuous Rumble Strips in Image



Figure 5.16 Unfavorable Lighting Conditions for Rumble Strips Detection

Since rumble strips are shallow grooves in pavements, their detection results are very sensitive to the surrounding noise. The source of noise includes oil marks, foreign objects, and stripe-shaped shadows. The faded rumble strips will also increase the noise by making the target less obvious. Figure 5.17 shows some example noises that could lead to false detection.





Figure 5.17 Noise Issue in Rumble Strip Detection

5. Summary

In this chapter, the utilization of the video log images in semi-automatic rumble strip detection was explored. The proposed algorithms performed well on continuous rumble strips using the GTSV data. However, it remains a challenge for detecting discontinuous rumble strips using either GDOT or GTSV data. The following recommendations suggest future improvements.

- Instead of manual identification of rumble strip region, an automatic method needs to be developed to improve the accuracy of determining the rumble strip location.
- A stronger equalization in the warped image could be evaluated to reduce the impact of lighting conditions.
- A confidence factor could be developed based on the appearance of rumble strips in continuous image streams.

References

Gordon, T. J., Bareket, Z., Gilbert, M., & Hagan, M. R. (2010). Method for Detecting Rumble Strips on Roadways. U.S. Patent Application No. 12/783,233.

Chapter 6 Exploration of Sign Detection Using Videolog Images

Traffic signs provide vital guidance to road users regarding traffic regulation, adequate road hazard warnings, destination and other geographic information, and temporary road conditions. As one of the most important elements of a sign management system, a traffic sign inventory provides fundamental information. It is a collection of data containing essential traffic sign information, including locations and attributes (e.g. types, dimension, post, etc.).

Based on our previous work, this chapter explores the utilization of GDOT's existing videolog images for automatic or semi-automatic traffic sign detection and recognition.

1. Previous Work

The PIs have developed a generalized traffic sign imaging processing model for developing a sign inventory (Tsai et al., 2009). The generalized traffic sign model, using an object-relational model (ORM), is implemented as shown in Figure 6.1. The figure illustrates the computational flow for detecting potential traffic sign candidates from the input image files. In Figure 6.1, the elementary entity in the model is a polygon that is associated with various other entities. The traffic sign color value is limited to the 10 colors (black, blue, brown, green, orange, red, white, yellow, fluorescent yellow-green (FYG) and fluorescent pink (FP)) defined in the Manual on Uniform Traffic Control Devices (MUTCD); the 10 colors are specified in the form of an enumeration or range in a brace set. A traffic sign may have multiple sub-regions. Each sub-region has a monochromatic plain texture in one color, and such a unique constraint is represented by a bar across the role.

The purpose of the traffic sign detection algorithm is to reliably differentiate images with signs from images without signs. It is composed in three main steps: (1) traffic sign color detection using a statistical MUTCD-color model; (2) a closed convex polygon approximation method for detection of a traffic sign boundary; and (3) traffic sign decision rules. The MUTCD statistical color model (SCM) is built statistically using manually labeled traffic sign color samples as ground truth. The training process calculates the probabilities that a given input pixel belongs to each MUTCD-color. By using trained MUTD SCM, one input image can be converted into multiple MUTCD-color gray images, each corresponding to a color to be tested.



Figure 6.1 Generalized Traffic Sign Model for Traffic Sign Detection (Tsai et al., 2009)

From a given input, polygons are first detected by their area dimensions, convexity, and boundary shapes. Because a traffic sign can have multiple sub-regions, their nested spatial relationships are considered at this step to determine the outer boundary of a traffic sign. If a polygon meets these rules, it is classified by its internal color types and portions. Specifically, a traffic sign must have at least two of the 10 MUTCD colors. For example, a stop sign has a red background and white letters. When a polygon area complies with specific traffic sign rules, it finally becomes a traffic sign candidate. Figure 6.2 shows a flowchart of the polygon approximation algorithm. The candidates are later judged by their various properties, which are (1) equilateral deviation; (2) legend and background color distribution; and (3) video frame continuity.

After comprehensive testing conducted by the PIs (Tsai & Wang, 2013), it was found that automatic sign detection and recognition cannot be directly applied to establish a comprehensive sign inventory due to nontrivial FP cases. To address this issue, an enhanced traffic sign inventory procedure was proposed by the PIs in a national demonstration project sponsored by the US Department of Transportation (Tsai & Wang, 2013). The objective was to improve the efficiency of the manual videolog-based sign inventory process by incorporating the existing automatic methods. The proposed method is capable of fully utilizing the strength of the automatic methods and still employ the manual process to overcome false-detection issues. With

the advancement of automatic methods, the efficiency of the proposed procedure should, also, be improved accordingly. Figure 6.3 shows the complete processes of the traffic sign inventory, including the proposed enhanced procedures. There are three primary paths to implement the traffic sign inventory. The manual method (path #1) is the most inefficient approach because an inspector has to review images frame by frame. However, the fully automatic method (path #2) is not so reliable due the difficulty of eliminating all the false-detection cases. Thus, path #3 is a more practical and efficient means to semi-automatically collect sign data by incorporating both automatic and manual processes.

In the semi-automatic method, the raw data is first processed using the enhanced image-based and LiDAR-based traffic sign detection algorithms. Since GDOT currently has no LiDAR data, only imaged-based sign detection algorithms are evaluated in this research project. Although not all the traffic signs can be detected using the existing algorithms, manual process can only focus on the correction of the falsely detected signs. This means that the majority of the manual digitization effort for traffic sign extraction can be saved. In addition, as the removal of the incorrectly detected cases requires less effort compared to the effort for the extraction of undetected traffic signs, the efficiency of the whole detection process can be improved.

The enhanced procedure presented in path #3 for traffic sign inventory is not limited by the performance of any individual algorithm. As the performance of the algorithms improves in the future, the enhanced methodology demonstrates a better reliability and efficiency toward full automation. The application of the automatic algorithms can be flexibly determined by state DOTs based on their needs, data availability, financial situation, etc. Thus, the proposed enhanced methodology can be flexibly applied to the existing practices in state DOTs.



Figure 6.2 Flowchart of Traffic Sign Detection Algorithm (Tsai et al., 2009)



Figure 6.3 Framework of the Enhanced Traffic Sign Inventory (Tsai & Wang, 2013)

2. Case Study

The proposed algorithms were tested on both GTSV and GDOT data. Table 6.1 lists six test sites; in each test site, there are two sets of data from the GTSV and GDOT, respectively.

Test Site	1	2	3	4	5	6
GTSV Data	I-20 WB Carroll County	I-20 EB Carroll County	I-20 WB Cobb County	I-20 EB Cobb County	I-20 WB Douglas County	I-20 EB Douglas County
GDOT Data	I-20 WB Carroll County	I-20 EB Carroll County	I-20 WB Cobb County	I-20 EB Cobb County	I-20 WB Douglas County	I-20 EB Douglas County

Table 6.1 Test Datasets for Sign Detection

Table 6.2 summarizes the test results. It should be noted that the total number of images and the total number of signs in two datasets of each test site are not similar. The first reason is that GTSV data and GDOT data use different data acquisition intervals. Thus, the total number of images are different. The second reason is that these two sets of data are not exactly matched. From the point of view of algorithm assessment, it is not needed to exactly match these two

datasets. The comparison is still useful because they are acquired on the same road with a similar roadside environment. In terms of detection accuracy, GTSV and GDOT data perform equally (88%). On test site #2 and #3, GDOT data slightly outperforms GTSV data. However, on other test sites (#4, #5, and #6), GDOT data apparently outperforms GDOT data. The number of false positive in each test dataset counts all falsely detected objects. One image might contain several false positive objects. Thus, the total number could be greater than the total number of images, e.g. GDOT data on test site #4. For the same roadway, the GT dataset contains more images, as the interval is less than that of GDOT dataset's.

Test Site	Dataset	No. of Images	# of Signs	Detected Signs (%)	False positive
1	GTSV	4,372	92	81 (88%)	504
	GDOT	3,192	106	94 (88%)	513
2	GTSV	4,377	85	50 (59%)	814
2	GDOT	3,193	108	68 (63%)	169
3	GTSV	1,170	37	23 (62%)	343
	GDOT	726	37	24 (65%)	320
4	GTSV	1,168	49	35 (71%)	339
4	GDOT	726	42	17 (40%)	744
5	GTSV	3,248	86	70 (81%)	625
	GDOT	2,022	76	47 (62%)	1,017
6	GTSV	3,210	84	61 (73%)	495
	GDOT	2,023	72	37 (51%)	1,079

Table 6.2 Test Results for of Traffic Sign Detection

From the test results, it can be seen that not all the signs can be detected. That means some signs are missed by the automatic sign detection algorithms. Thus, path #3 in the framework of enhanced sign inventory (Figure 6.3) should be applied. Based on the study conducted by the PIs (Tsai & Wang, 2013), the semi-automatic method can save about 40% of manual review and data input time. When the performance of sign detection algorithms increases, the amount of saved time will, also, increase.

Through the investigation of the false detections, the following contributing factors were identified: 1) images captured when data collection vehicle runs in the middle lane; 2) improper camera view angles; 3) sign-like objects; and 4) inconsistent lighting conditions. These factors

could act independently or dependently. For GDOT data, when the images were captured from the middle lane, passing vehicles not only blocked the roadside signs, but also brought many false positives because of sign-like objects. Figure 6.4 shows an example that illustrates the false positives caused by factors 1) and 3).



Figure 6.4 False Positive Caused by Passing Vehicle

The testing results from the GTSV data show that some overhead signs are missed, which are caused by the improper vertical view angle. As shown in the left image in Figure 6.5, because the vertical angle to the pavement is set large, an overhead can only be seen from far away. However, its size would be too small to be detected. When the data collection vehicle approaches the overhead sign, only part of it can be seen (right image in Figure 6.5). As discussed in Chapter 6, a large view angle (for downward images) is good for detecting pavement surface features but not good for high objects. In contrast, a small view angle (for perspective images) is good for detecting overhead signs, but not good for features on pavement surfaces. If these two roadway assets are needed, extra cameras are needed.

The inconsistency in lighting conditions is the biggest challenge for the MUTCD SCM. As color distributions are widely dispersed because of different viewing conditions that deviate from the original MUTCD-color regions, the training samples may not cover the entire range of a particular color under all circumstances. As a result, the MUTCD-SCM can be inaccurate when the lighting condition differs from the training stage, which can lead to either false positives or false negatives. In this case study, only a small set of images were used for training the color

model, which might not cover all the lighting conditions. For a large-scale implementation, a bigger training set is needed. Thus, the detection accuracy could be further improved.





Figure 6.5 Improper Vertical View Angle

3. Summary

In this chapter, a potential utilization of the video log images in traffic sign detection is assessed. The automatic detection method presented in this work can be fed into the enhanced framework of sign inventory that was proposed by the PIs, and, thus, improve the efficiency and productivity of sign inventory. According to the test result from the GTSV and GDOT data, the algorithm is able to detect up to 88% of the traffic signs. However, the detection accuracy could be improved by introducing a large-scale color model training and enhancing the image quality. Then, the data collection productivity could further be improved. Since GPS location (longitude and latitude) associated with each image (it cannot be read using the proprietary data viewer software) and the intrinsic parameters of camera calibration are not revealed in GDOT data, it is difficult to accurately determine the sign location and its dimension. It is suggested that the relative information from data service providers be requested in the future.

References

Tsai, Y., Kim, P., & Wang, Z. (2009). Generalized Traffic Sign Detection Model for Developing a Sign Inventory. Journal of Computing in Civil Engineering, 23(5), 266-276.

Tsai, Y. and Wang, Z. (2013). A Remote Sensing and GIS-enabled Asset Management System (RS-GAMS). DTOS59-10-H-0003, US Department of Transportation, 2013

Chapter 7 Consideration of Data Specification and Quality Checking

Previous chapters present the assessment of videolog image quality and the exploration of using videolog images for collecting guardrails, rumble strips, and traffic signs. The objective is to assess how GDOT's existing videolog and pavement imaging data can be reused to support statewide maintenance practices. In the meantime, the results from the assessment and exploration can be used to support GDOT's future outsourcing of data collection by providing extra requirements for data quality and specifications. Thus, the data collected by vendors could be better reused. In this chapter, the data quality and specifications for both videolog and pavement imaging data are discussed.

1. Videolog Images

Videolog images capture the natural scene of roadway environments; they are used to extract pavement and roadside assets. For this purpose, a targeted asset should be positioned correctly and clearly in an image. Otherwise, the extraction process would be difficult and/or inaccurate. Natural scene images are ideal for visual inspection; this process is very time consuming and labor intensive. Thus, many efforts have been spent in research and industry to automate data extraction processes. However, the accuracy of the outcomes of automatic methods largely depend on the quality of input videolog images. To achieve satisfactory results, a rigorous data specification and quality checking process is needed. The following subsections discuss the factors that should be considered in a specification and quality checking process.

1.1 Image Composition

Photo composition in photography is important for achieving a certain level of art performance. Similarly, image composition in videolog on roadways is, also, critical to better extract pavement and roadside features. Image composition includes view angles (horizontal and vertical), resolution, and acquisition interval.

74

• View angles

Normally, three front camera are used in videolog to capture a roughly 180° panoramic view, including left and right shoulders and a perspective front view, as shown in Figure 7.1 (a). To avoid distortion, wide-angle lenses should not be used. The use of the perspective front view makes it easy to capture high-rise objects, such as bridge clearance and overhead signs. However, due to the small vertical angle to the pavement (Figure 7.1 (b)) of the perspective view, pavement features such as distresses (raveling, cracking, etc.) and rumble strips, are hard to extract. Figure 7.2 illustrates how vertical view angles affect the appearance of pavement surface. The left image has a larger vertical angle to the pavement surface; the right one shows a perspective view. The rumble strips on the right shoulder are much clearer in the left image, as are the pavement details. In the right image, due to the small vertical angle to the pavement surface, it is difficult to discern pavement details. Thus, the left image is more suitable for extracting pavement surface features. In GDOT's existing videolog images, all the cameras target the perspective view. Thus, the performance on detecting rumble strips is not good. If the side view images (e.g. front right) are used to inspect guardrail conditions, the guardrails are often out of range (see Figure 7.3). To better capture pavement surface features, it is suggested adding a downward view, as shown in Figure 7.1 (b). This camera can be positioned either forward or backward. On the GTSV, the downward camera is installed backwards. In addition, for a better inspection of roadside assets, the current view angles to pavement on side views should be increased a little in comparison to the current settings.

• Resolution and acquisition interval

Image resolution is controlled by the minimum pixel size of the targeted roadway features. For example, for traffic signs, mile markers are smaller than most other signs. Thus, for a well-positioned mile marker image, the pixel size should not be too small. Otherwise, it cannot be detected by an automatic sign detection algorithm. Tsai et al. have done an experimental test on sign detection rate vs. sign size as shown in Figure 7.4. It can be seen that the detection rate drops under 90% when the pixel size is less than 40. Since videolog images are used to extract various roadway assets, a more comprehensive study is needed to determine the minimum resolution of each image. Another consideration is the image file size. Increased resolution will also result in a higher demand for hard disk storage space.



Figure 7.1 Camera View Angles



Figure 7.2 Impact of Different Vertical View Angles



Figure 7.3 Roadside Feature Out of Range



Detection Rate vs. Traffic Sign Size

Figure 7.4 Impact of Sign Pixel Size on Detection Rate (Tsai, et al., Unpublished)

Image acquisition interval is another factor that needs to be considered in terms of roadway feature extraction. Figure 7.5 shows a stop sign that is captured by four consecutive images. Due to the nature of natural scene images, an automatic sign detection algorithm cannot guarantee a 100% of detection rate. Thus, if a sign appears in several images, the possibility that it could be detected increases. On the other hand, signs with larger pixel sizes have better chances of being detected. Thus, if image resolution is big, the acquisition interval could be reduced. Currently, there is no a theoretical study on choosing optimal image acquisition and resolution. However, a trial-and-error method could be applied to determine a good combination of image acquisition interval and image resolution. The average acquisition interval of GDOT's existing videolog images is about 26.4 ft. (8 m), which is

larger than the one for the GTSV images (5 m). Based on the experimental test on sign detection, it is suggested the current image acquisition interval be reduced.



Figure 7.5 A Sign Captured in Consecutive Images

Lane Location

Based on the exploration of the automatic detection of guardrails, rumble strips, and traffic signs, it was found that the lane location in which data collection vehicle drives is important for collecting roadside assets. Figure 7.6 shows three cases of false detection caused by the images captured from the middle lane instead of the outer lane. For guardrails and traffic signs, the interference from the passing vehicles causes false detections, as shown in Figure 7.6 (a) and (c). For rumble strips, if images are captured from the middle lane, the roadside rumble strips are hard to discern, as seen in Figure 7.6 (b). Of course, passing vehicles will also cause problems for rumble strip detection because they will block the view to the right shoulder.

Since roadway assets are usually installed on shoulders, it is suggested collecting videolog images from the outer lane. For interstate highways, some roadside assets, such as signs and rumble strips, are also installed on the left shoulder. To extract these assets, another round of videolog images in the inner lane is suggested.





(a) Guardrail Detection

(b) Rumble Strip Detection



(c) Sign Detection



1.2 Image Quality

Image quality directly affects how well the image can be used for extracting roadway assets, especially when automatic methods are used. In Chapter 3, five types of image-based quality indicators, blurriness, haziness, blockiness, noise, and entropy, were introduced. In literature, blurriness and blockiness are correlated to human perceptions, and thresholds can be directly used for image quality checking. Other indicators can also be used to compare with other known good-quality images. Based on the experimental tests presented in Chapter 3, the blurriness of GDOT's images on the test roadways shows between "perceptible" and "slight annoying." The blockiness also shows the same quality, i.e., in the scope between "perceptible" and "slight annoying." In contrast, the tested GTSV images show between "imperceptible" and "perceptible." Thus, there is still some room to improve the quality of videolog images in future data collection.

In addition to the quality issues caused by a camera's intrinsic problems, Chapter 3 also discussed the issues caused by environments, such as sun glare, dark, lens debris, and fog. Tsai & Huang (2010) have also developed algorithms to automatically detect these defects.

Some of the above image quality issues could be alleviated by following a good practice of data collection that defines when, where, and how to collect videolog images. GDOT also defines the detailed requirements on how videolog images should be collected. The following lists some examples:

- To avoid lighting issues, it is strongly suggested not collecting videolog images near sunrise or sunset. Strong sunshine would create shadow that brings trouble for automatic data extraction. Thus, cloudy or partially could days are ideal for collecting videolog images.
- The data collection must not take place if there is precipitation falling, if the roadways are wet, or if there is visible precipitation, such as snow, ice, hail, standing water etc., on the roadway, sidewalks, and medians, or in other areas of the roadway rights-of-way, that would hamper data collection or data extraction. Data collection should not take place if there is any severely inclement weather that would hamper data collection: fog, flying debris due to storms or severe winds, extreme darkness due to storms, lightening, or other personnel hazards, etc.
- A gray card should be prepared as a middle gray reference. Before formal collection is started, a number of pretest images should be collected with the gray card included on the canvas. These pretest images will provide reference for color correction and rebalancing.
- Data collection vehicles should be driven in outer lanes to avoid interference from passing vehicles. If targeted roadside assets exist at the left side of the road, another round of data collection is suggested by running data collection from the inner lane.
- Images must be collected as clear digital images and free of distortion, overexposure, underexposure, or obstructions that would prevent the extraction of the roadway of characteristics, asset and inventory data, or other data items normally gathered via videolog.

2. 3D Pavement Imaging Data

In comparison with videolog images, 3D pavement imaging data is relatively new and captured by the emerging 3D laser technology. Nevertheless, more and more state DOTs have interest to adopt this technology due to its advantages over 2D videolog images. Currently, the pooled fund study TPF-5(299) set the goal to improve the data quality and analysis. The PIs have conducted a technology overview on validating 3D pavement imaging data for TPF-5(299) (Tsai & Wang, 2015). In addition, the PIs have conducted comprehensive study to validate the use of 3D laser technology for pavement condition data collection through two national demonstration projects sponsored by USDOT (Tsai & Wang, 2013; Tsai & Wang, 2014). The following summarizes data quality checking for 3D pavement imaging data based on the PIs' past research.

2.1 Data Specification

The triangulation principle, which is also commonly known as the laser triangulation, is used for 3D pavement imaging data acquisition. Figure 7.7 illustrates the concept of the system diagram. In the 3D laser system, the laser light projected is a fine laser line. A high-intensity, areascanning camera with an image sensor, such as a complementary metal oxide semiconductor (CMOS) sensor or charge-coupled device (CCD) sensor, is placed at a known distance and at an oblique angle (θ) with respect to the laser line projector. The camera takes images of the laser light. Then, the deformations of the laser line on the object are analyzed to evaluate the elevation for each point with a known horizontal position on the object. Because the laser sensors are coupled with a distance measurement instrument (DMI), the system can obtain point locations along the longitudinal direction. Consequently, a complete three-dimensional set of points of the object's surface can be acquired. These 3D points can be used to identify pavement distresses and to measure the geometry of the distresses.



Figure 7.7 3D Laser System (Tsai & Wang, 2013)

The specification regarding 3D pavement imaging data includes range measurement accuracy, transverse data resolution, and longitudinal data resolution. For example, the 3D laser system installed on the GTSV has a 0.5 mm range measurement accuracy. The transverse resolution is 1 mm. The longitudinal resolution is 5 mm when the vehicle runs at 100 km/hr. The longitudinal resolution is dependent on vehicle speed because the shuttle speed of the digital camera in the 3D laser system is fixed. For the system installed on the GTSV, the shuttle speed is 5,600 Hz. That means there are 5,600 transverse profiles captured in a second. To increase the longitudinal resolution, e.g. 1 mm, but maintain the highway speed, the camera shuttle speed has to be increased. However, higher resolution also demands higher storage space and higher computing power. Thus, it is suggested that specifying the 3D data specification be based on the need for extracting pavement distress data. Based on the PIs' study (Tsai & Wang, 2013), the current specification of the 3D laser system installed on the GTSV is capable of detecting cracks with widths greater than 2 mm. If 1 mm hairline cracking were desirable, higher specification would be needed.

2.2 Data Quality Validation

The quality of 3D pavement imaging data is affected by various factors that are either intrinsic or external. For example, due to the nature of a laser beam, it is hard for a 3D laser system to measure the range of a shining surface, such as new pavements with moisture or shining aggregates. In addition, a 3D laser system can only measure a certain range of distance. That is why the measurements are often out of range on a bumpy road. Thus, a rigorous quality checking is needed for 3D pavement imaging data, even though the data specification is clearly known.

The quality checking of 3D pavement imaging data is often performed by validating the data accuracy. Normally, an object is measured and compared with known dimensions and ranges. A set of acceptance thresholds are defined in terms of the difference between the actual measurement and the known values. This object could be an artificial calibration/validation board or a test site with known rutting and/or cracking that can be measured by 3D pavement imaging data.

Through the research project sponsored by USDOT, Tsai et al. developed a calibration/validation board as shown in Figure 7.8. The board is used to quantitatively evaluate the accuracy of 3D pavement imaging data, especially in the depth direction, which consists of three knows steps of 0.5 mm, 1.0 mm, and 2.0 mm and 27 grooves with known dimensions. If the difference between a measurement and the known value is greater than a threshold value (that was not standardized yet), a calibration/validation process, including sensor configuration checking and system setup checking, such as checking the tire pressure and the alignment of all sensors.



Figure 7.8 Calibration/Validation Board Developed by Tsai, et al.

The PIs also performed validation on 3D pavement imaging data accuracy by measuring rut depth and detecting cracks using lab and field tests (Tsai & Wang, 2013). Figure 7.9 shows the artificial rutting and cracking (left images), and an actual rut depth test site and a real pavement

crack that were used for data accuracy validation. It is suggested for GDOT to adopt a validation board and establish a test site with known distresses, such as rutting, cracking, and IRI. During the contract period, the data service provider should perform periodic testing on the validation board and test site to validate the data accuracy.



Figure 7.9 Validation on Rut Depth and Cracks

3. Data Format

In GDOT's existing data, videolog images are in JPEG format that can be read by different software and programming libraries. Thus, the detection of guardrails, rumble strips, and traffic signs can be explored in this research project. However, the GPS location (latitude and longitude) along with each image was not revealed; it can only be accessed by using the vendor-provided proprietary data viewer software. In addition, the parameters for camera calibration were not revealed either. The intrinsic parameters, such as focal length, optical center, and skew coefficient, for camera calibration are used to map the 2D location and dimension information to real-world 3D measurements. For example, for a detected sign, if the GPS location of the image and the parameters for camera calibration are known, the sign's real-world location and

dimensions can be determined. Therefore, to fully reuse the videolog images for extracting other roadway assets, requesting an open format of image GPS locations and camera calibration parameters is strongly suggested for the future outsourcing contracts.

On the other hand, GDOT's pavement imaging data is stored in a proprietary format. Only the proprietary data viewer software can be used to display the data. Thus, it is hard for anyone other than the data service provider to use the data for extracting pavement distresses (e.g. cracking and raveling). It will limit the use of the raw data. As discussed in Chapter 2, with the growing interest from state highway agencies, more and more vendors are manufacturing hardware and providing data services, but they normally develop and use proprietary software and technology to process, display, and report the collected data. Using proprietary formats for storing 3D pavement imaging data faces the challenge of meeting transportation agencies' different data requirements and makes it difficult to unify data analysis, reporting, sharing, and evaluating. FHWA has initiated a project to develop a common and interchangeable data format for pavement imaging data. It is reasonable for GDOT to request an open format of pavement imaging data in future contracts. Thus, GDOT has the flexibility to choose better vendors for roadway feature and pavement distress extraction.

References

- Tsai, Y., & Huang, Y. (2010). Automatic detection of deficient video log images using a histogram equity index and an adaptive Gaussian mixture model. Computer - Aided Civil and Infrastructure Engineering, 25(7), 479-493.
- Tsai, Y., and Wang, Z. (2009). Technology Overview on Validating 3D Transverse Profile Data and Measurement of Pavement Surface Distresses. Pooled Fund Study: TPF-5(299).
- Tsai, Y. and Wang, Z. (2013). A Remote Sensing and GIS-enabled Asset Management System (RS-GAMS). DTOS59-10-H-0003, US Department of Transportation.
- Tsai, Y. and Wang, Z. (2014). A Remote Sensing and GIS-Enabled Asset Management System (RS-GAMS) Phase 2. RITARS-11-H-GAT, U.S. Department of Transportation.

Chapter 8 Conclusions and Recommendations

GDOT's videolog and pavement imaging data is a valuable asset that has enabled GDOT to fulfill the needs of an HPMS. On the other hand, roadway asset inventory and pavement condition data are essential to support statewide asset management and maintenance programs. Thus, the return on investment would be further increased if GDOT's existing and future videolog and pavement imaging data could be utilized for extracting roadway assets and pavement condition data. For this purpose, this research project explored the utilization of GDOT's existing videolog and pavement imaging data for extracting guardrails, rumble strips, and traffic signs. In the meantime, through the comprehensive investigation and analysis of GDOT's existing data and the project's exploration results, suggestions were made for data specification and data quality control that can be added to GDOT's existing requirements for outsourcing contracts. Therefore, future data could be better utilized to support the statewide asset management and maintenance. The following summarize the research outcomes and major findings:

- Videolog and pavement imaging data is a big investment for GDOT. Though it has supported GDOT in fulfilling the requirements of HPMS, the large volume of data is not currently fully utilized in GDOT to support the statewide asset management and maintenance programs. To improve the return on investment, there is a need to apply the data for supporting the extraction of other roadway assets. Due to the use of proprietary format on pavement imaging data, its current application is limited to display data using the proprietary data viewer software. The videolog images are in JPEG format and can be used for extracting guardrails, rumble strips, and traffic signs.
- 2) Image quality determines how accurately and reliably roadway assets can be extracted. Five image-based quantitative indicators, blurriness, haziness, blockiness, noise, and entropy, were recommended for image-based quality assessment. A prototype assessment tool was also developed. These five indicators provide an objective, automatic, and full-scope alternative for the current subjective, manual, and sampling-based quality checking of videolog images.

- 3) In addition to the image-based quality assessment, composition-based quality assessment, such as camera configuration, environmental deficiencies, and object-based assessments, such as how images could support the extraction of guardrails, rumble strips, and traffic signs, are also important to support accurate and reliable extraction of roadway assets.
- 4) The proposed algorithm for extracting guardrails is very promising. Test results showed that an accuracy of 99.1% can be achieved by using the GTSV data, and an accuracy of 88.8% can be achieved by using GDOT data. Though it is still challenge to automatically collect all the detailed information for a guardrail, doing so can potentially save an inspector's time by automatically removing all irrelative videolog images that contain no guardrails.
- 5) The detection of discontinuous rumble strips remains a challenge using the existing videolog images because the image acquisition interval and the camera vertical view angle are not suitable. For continuous rumble strips, the test results showed promising results with a detection accuracy of 81.9% and 95% on two test datasets collected by the GTSV.
- 6) The test results on sign detection showed various detection accuracies using different testing datasets. However, the automatic detection results can be incorporated into an enhanced framework of sign inventory, proposed by the PIs in a previous research project, to improve the efficiency and productivity in sign data collection.
- 7) Based on the exploration of using videolog images for extracting various roadway assets and the PIs' previous research, requirements were discussed for videolog images in terms of their composition and quality, and 3D pavement imaging data in terms of data specification and data quality validation. Fulfilling these requirements will help improve data quality, and thus, better utilize the data to support statewide asset management and maintenance.

The following recommendations are for future research and implementation:

 Due to the change of roadway assets over time, the collection of videolog and pavement imaging data is a recurring need for GDOT. To enhance data quality, it is suggested that the quantitative image quality indicators that are proposed in this research project be incorporated into GDOT's future quality checking procedures. An automatic, full-scope method can be developed to replace the current manual, sampling-based approach. To better utilize it for statewide roadway asset management and maintenance, the suggested dataspecification-related factors need further study and should be incorporated into GDOT's future outsourcing contracts.

- 2) The testing on guardrail detection showed very promising results. Testing on a large scale is recommended to further improve the performance of the proposed algorithms. Though the manual process is still needed to collect the detailed guardrail condition information, the automatic detection results can be combined with the manual data collection process to improve the overall efficiency and productivity, and thus, reduce costs.
- 3) The algorithms for rumble strip detection need further refinement. For example, instead of manual identification of a rumble strip region, an automatic method needs to be developed to improve the accuracy of determining rumble strip locations. A stronger equalization in the warped image could be evaluated to reduce the impact of lighting conditions. A confidence factor could be developed based on the appearance of rumble strips in continuous image streams. Regarding discontinuous rumble strips, other than the improvement of algorithms, there is a need to study the optimal image acquisition interval in order to avoid the misplacement of rumble strips in videolog images.
- 4) To implement sign data collection by incorporating the automatic sign detection procedures, a large-scale, color model training is needed to improve the sign detection accuracy. In addition, to accurately and reliably determine sign locations and dimensions, the GPS location of each videolog image and the intrinsic parameters of camera calibration need to be acquired from the data service provider.
- 5) 3D pavement imaging data is ideal for collecting various pavement distresses, such as cracking, rutting, and raveling, to support pavement management, maintenance, and rehabilitation. If the data can be accessed publically, instead of only by proprietary data viewer software, more flexibility would be gained for data collection. Thus, the cost would be reduced.

88