

Innovative Techniques for Analyzing Cyclist Behaviour and Predicting Cyclist Safety

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree

Doctor of Philosophy
In
Civil Engineering

by

Ali Kassim
B.Sc. University of Mosul
M.A.Sc. University of Jordan

Department of Civil and Environmental Engineering
Carleton University
Ottawa-Carleton Institute of Civil and Environmental Engineering

September

© 2014, Ali Kassim

Abstract

Cycling is an important mode of travel which offers powerful solutions to chronic traffic problems of congestion and emissions. Therefore, it is imperative for the road designers to have reliable information on the cyclist behaviour and characteristics to design safe and efficient cyclist facilities. The behaviour of the cyclist can be measured in different ways, including, cyclist crossing speed, cyclist violations, and cyclists- vehicle interactions.

The essential focus in this thesis was on cyclists, especially their speed measurement and analyzing cyclist-vehicle interactions. In this thesis, a reliable video analysis technique to measure cyclist speed at signalized intersection was developed. This technique enables the automated observation of large volume of naturalistic cyclist movements in an accurate and resource-efficient manner.

As for cyclist safety, relying solely on collision records to analyze cyclist safety is challenged by inherent limitations in collision data. These limitations can be quantitative and/or qualitative. Traffic conflict techniques have been used as a proactive and integrated approach to collision-based road safety analysis. However, traditional traffic conflict studies are mostly field-based studies and due to the subjectivity of field observers, errors were common when manually counting and deciding/judging whether a given traffic event is a conflict. In this thesis, interactions between motor vehicle and cyclists at signalized intersections were characterized using an objective conflict indicator; Post-Encroachment Time (PET).

The thesis also produced a sizeable database of 806 hours of video data for cyclist movements. A total of 19,058 cyclists and 48,632 vehicles were observed within a total

period of 57 days. The following contributions were achieved in this thesis: [i] video tracking system was improved from an open-source feature-based vehicle tracking system in order to track cyclist and vehicles and produce a trajectory database, [ii] different analysis methods were developed to measure cyclist crossing speed, [iii] automated measurement of cyclist crossing speed, [iv] investigation of the effect of different potential factors affecting cyclist crossing speed, [v] development and validation of an automated method to measure PET between cyclists and motor vehicles using video analysis techniques, [vi] different statistical techniques were investigated to utilize PET observations to measure a cyclist safety.

Acknowledgements

First of all, profusely and all thanks are due to my God who enabled me to achieve this research work.

This thesis could not have been done without the support and advice of my two supervisors, Dr. Karim Ismail and Dr. Yasser Hassan. I would like to express my deeply heartfelt gratitude to Dr. Karim Ismail, for his supervision, prudent guidance, friendly and untiring efforts to support me over the past five years. I would like to convey my gratefulness and sincere appreciation to Dr. Yasser Hassan for his guidance, supervision and encouragement.

I would also like to express my great thanks to the members of the examination committee, Dr. Adel Sadek, Steven Prus, Ata Khan and David Bell. I would like to give my sincere appreciation to the Carleton University and especially the Civil and Environmental Engineering department.

I would like to thank all the people that have helped me during the data collection, especially the people that allowed me to install the video cameras in their apartments and buildings. I appreciate their help and kindness towards me.

I am grateful to have such great friends that were there for me when I needed help, they are Munther Hussain, Othman Nasir and Lai Zheng. I am indebted to all those people and other members whom I have not mentioned them here for helping me.

Finally, this work cannot be done without the unconditional support of my wife “*Sahar*” and lovely kids “*Sara and Amina*”.

Table of Contents

Abstract.....	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	x
List of Figures.....	xv
List of Appendices.....	xxii
1. INTRODUCTION.....	1
1.1 Background.....	1
1.2 Problem Definition	4
1.3 Research Objectives	6
1.4 Research Contributions	7
1.5 Organization of the Thesis.....	8
2. LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Cyclist Speed Measurements.....	10
2.2.1 Cyclist speed measurements	13
2.2.2 Semi-automated cyclist speed measurements	15
2.2.3 Automated cyclist speed measurements.....	24
2.2.4 Identified limitations	25
2.2.4a Limited movement directions	25
2.2.4b Cyclists selection	26
2.2.4c Seasonal variation	26
2.2.4d Completeness of result reporting	27
2.2.4e Device limitations	27

2.2.4f	Measurement validation	29
2.2.5	Case for computer vision techniques	30
2.3	Approaches to Study Cyclist Safety	30
2.3.1	Background	32
2.3.2	Cyclist safety treatments	36
2.3.2a	Cyclist lane treatment.....	36
2.3.2b	Advanced stop and bike-box treatments	40
2.3.2c	Other treatments	44
2.3.3	Conflict indicators	46
2.3.4	Post-Encroachment Time	48
2.4	Summary.....	49
3.	AUTOMATED MEASURING OF CYCLIST SPEED.....	54
3.1	Data Collection.....	55
3.1.1	Study site and equipment	55
3.1.2	Video data collection.....	56
3.1.3	Surveying works and camera calibration	57
3.1.3a	Surveying works	57
3.1.3b	Camera calibration.....	58
3.1.4	Selection of tracking parameters	60
3.2	Speed Measurements	65
3.2.1	Cyclist water trail (CWT).....	66
3.2.2	Active crossing area (ACA).....	67
3.2.3	Cyclist video-measured speed.....	70
3.3	Validation of Cyclist Speed Measurements.....	71
3.3.1	CWT field-measured speed.....	72
3.3.2	ACA field-measured speed	73

3.4	Data Analysis and Results	75
3.4.1	Descriptive statistics.....	75
3.4.2	Factors affecting crossing speed	77
3.4.3	Comparison of cyclist speed results	79
3.5	Summary.....	84

4. AUTOMATED MEASURING OF CYCLIST-MOTOR VEHICLE POST ENCROCHMENT TIME AT SIGNALIZED INTERSECTIONS..... 86

4.1	Data Collection.....	87
4.1.1	Study site and equipment	87
4.1.2	Video data collection.....	89
4.1.3	Surveying works and camera calibration	89
4.1.3a	Surveying works	90
4.1.3b	Camera calibration.....	91
4.2	PET Measurement Methods	92
4.2.1	Manual video timer measurement (MVTM) method.....	95
4.2.2	Manual frame count measurement (MFCM) method.....	97
4.2.3	Automated measurement (AM) method.....	98
4.3	Automated PET Measurement.....	101
4.3.1	PET measurement algorithm.....	101
4.3.2	Detailed procedures in the PET measurement algorithm.....	103
4.3.3	Adjusting the initial PET.....	106
4.4	Results and Discussions	109
4.4.1	Descriptive statistics of PET	109
4.4.2	Validation of AM method	111
4.4.3	Accuracy of MVTM method.....	112
4.5	Summary.....	115

5. DEVELOPMENT AND APPLICATION OF STATISTICAL MODELING TO MEASURE A CYCLIST SAFETY	118
5.1 Data Collection.....	119
5.1.1 Site Description.....	119
5.1.2 Video Data Collection Equipment	123
5.1.3 Video Data Collection.....	125
5.2 Performance Parameters.....	127
5.2.1 Background	127
5.2.2 Movements.....	127
5.2.3 Surrogate Measures of Safety	127
5.2.4 Observation Methodology.....	129
5.3.5a Cycling against traffic.....	133
5.3.5b Stopping after stop-line.....	134
5.3.5c Ride in a crosswalk	135
5.3.5d Out-of-path stop.....	135
5.3.5e Riding out-of-path.....	136
5.3.5f Cross stop-line on red signal	136
5.3.5g Vehicle turns left or right and stops on cyclist path.....	137
5.3.5h Vehicle fails to yield to the cyclist.....	139
5.3 Data Analysis.....	141
5.3.1 Different measure of exposure	141
5.3.2 Comparative statistical analysis	144
5.4 Regression Analysis of Count Data.....	155
5.4.1 Counts models.....	155
5.4.2 Effect of cyclist volume, vehicle volume, treatment and season on PET counts...	156
5.4.3 Effect of pedestrian volume on PET Counts.....	168

5.5	Statistical Modeling Based on Extreme Value Theory.....	173
5.5.1	Background	173
5.5.2	Return level value	174
5.5.3	Threshold selection	175
5.5.4	Modeling results.....	176
5.5.5	Correlation analysis.....	179
5.6	Summary.....	181
6.	CONCLUSIONS AND RECOMMENDATIONS.....	186
6.1	Summary and Conclusions	186
6.1.1	Summary	186
6.1.2	Conclusions	198
6.2	Recommendations and Future Work	202
	References.....	205

List of Tables

Table 2-1: Comparison of different techniques used for cyclist speed measurements.....	50
Table 3-1: Absolute value of the difference between actual field distances using measuring wheel and estimated distances using calibrated camera parameters.	61
Table 3-2: Average overall errors between linear field measurement (shown in Figure 3-3) and estimated distances using camera calibration output in different days and seasons.	62
Table 3-3: Description of the configuration file tracking parameters*	64
Table 3-4: Summary of the tracking parameters.....	65
Table 3-5: Validation test between cyclist field-measured and video-measured speed for the first approach (16 crossing events).	72
Table 3-6: Summary results of cyclist field-measured and video-measured speeds.	74
Table 3-7: Fit Distribution of the field-measured and video-measured speeds.	77
Table 3-8: ANOVA results for effect of different factors on cyclist crossing speed.	78
Table 3-9: ANOVA results of the cyclist speed unsigned and signed error between field-measured and video-measured speed.....	79
Table 3-10: Summary of the cyclist speed measurements using different devices.	80
Table 3-11: Meta analysis results of the average cyclists' speed at the road sections and the signalized intersections.	83
Table 4-1: Coordinates of the bounding box nodes.....	100
Table 4-2: Directions of vehicle movement and corresponding adjustments to initial PET.	108
Table 4-3: Analysis results of PET using MVTM, MFCM and AM methods.	110

Table 4-4: Fit distribution of the PET measurement methods.....	111
Table 4-5: Frequency of the PET levels by MVTM, MFCM and AM methods.	113
Table 4-6: Agreement evaluation using kappa values under the PET levels of (0, 1], (1, 2] and (2, 3] seconds.	114
Table 5-1: Video data collection during Fall -2011 and Summer -2012 seasons.....	126
Table 5-2: Average hourly volume during the Fall and Summer 2011 seasons.....	145
Table 5-3: Statistical results of the average hourly volume during Fall -2011 and Summer -2011 seasons.....	145
Table 5-4: Data summary of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.....	147
Table 5-5: Statistical results of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.....	147
Table 5-6: Data summary of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.....	148
Table 5-7: Statistical results of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.....	148
Table 5-8: Data summary of the average of cyclist left-turn interaction for Fall -2011 and Summer -2012 seasons.	150
Table 5-9: Statistical results of the average of cyclist left-turn interactions for Fall -2011 and Summer -2012 seasons.....	150
Table 5-10: Summary of right-turn and left-turn interactions rate for Fall -2011 and Summer -2012 seasons.	151

Table 5-11: Statistical results of right-turn interaction rate for different seasons at Metcalfe intersection.....	152
Table 5-12: Summary of findings seasonal change in proportion of cyclists involved in right-turn or left-turn interactions.	152
Table 5-13: Summary of the cyclist violation rate for Fall -2011 and Summer -2012 seasons.	154
Table 5-14: Statistical results of the cyclist violation rate for Fall -2011 and Summer -2012 seasons.	154
Table 5-15: Summary of the motor vehicle violation rate for Fall -2011 and Summer -2012 seasons.	154
Table 5-16: Statistical results of the motor vehicle violation rate for Fall -2011 and Summer -2012 seasons.	155
Table 5-17: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 3] seconds.....	159
Table 5-18: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 2] seconds.....	160
Table 5-19: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 1] second.....	161
Table 5-20: Suggested the best three Negative Binomial models to accommodate the PET counts.	162
Table 5-21: Summary of regression analysis of PET (0, 3] seconds when vehicle yields to the cyclist (positive).....	163

Table 5-22: Summary of regression analysis of PET (0, 2] seconds when vehicle yields to the cyclist (positive).....	163
Table 5-23: Summary of regression analysis of PET (0, 1] second when vehicle yields to the cyclist (positive).....	163
Table 5-24: Summary of regression analysis of PET (0, 3] seconds when vehicle fails to yield to the cyclist (negative).....	165
Table 5-25: Summary of regression analysis of PET (0, 2] seconds when vehicle fails to yield to the cyclist (negative).....	165
Table 5-26: Summary of regression analysis of PET (0, 1] second when vehicle fails to yield to the cyclist (negative).....	165
Table 5-27: Summary of negative binomial regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds.	167
Table 5-28: Pedestrian volume at six signalized intersections (8:30 AM to 4:00 PM)..	169
Table 5-29: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds.....	169
Table 5-30: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds when vehicle yields to the cyclist (positive).	170
Table 5-31: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds when vehicle fails to yield to the cyclist (negative).....	171
Table 5-32: Summary of the pedestrian effects on PET counts.....	172
Table 5-33: Summary of estimation results of the return level value at the study area..	178
Table 5-34: Summary of the right-turn interaction rates, intersections return level values and violation rates.	180

Table 5-35: Spearman's correlation rank coefficient results.	180
Table 5-36: Spearman's correlation rank coefficient results: Right-turn interaction versus the average cyclist violation per type.....	181
Table B-1: Sample calculation of the interpolation algorithm.....	236

List of Figures

Figure 2-1: Schematic diagrams for the pneumatic tube counter device at the field after cyclist lanes were marked (modified from (Hunter, <i>et al.</i> , 2009)).	16
Figure 2-2: Schematic diagrams: (a) of off-street and (b) on street transportation facility (reproduced from (El-Geneidy, <i>et al.</i> , 2007)).	19
Figure 2-3: Predicted cyclist travel speeds under three different types of facility (reproduced from (El-Geneidy, <i>et al.</i> , 2007)).	19
Figure 2-4: Conflict scenarios between cyclist and motor vehicle (reproduced from (Summala, <i>et al.</i> , 1996)).	33
Figure 2-5: Conflict scenarios between cyclist and motor vehicle (reproduced from (Phillips, <i>et al.</i> , 2011)).	34
Figure 2-6: Cyclist lane treatment. (a) marked cyclist lane. (b) segregated cyclist lane (photos taken by the thesis author).	37
Figure 2-7: Schematic diagram of advanced stop line (ASL) (reproduced from (Koorey & Mangundu, 2010)).	40
Figure 2-8: Schematic diagram of a bike box at signalized intersection (reproduced from (Weigand, 2008)).	41
Figure 2-9: Schematic definition of TTC and PET.	47
Figure 2-10: Scenarios of conflict movements between cyclists and motor vehicles at signalized intersection. (a) Right-turning motor vehicles versus cyclists. (b) Through motor vehicles versus left-turning cyclists. (c) Left-turning motor vehicle versus through cyclists (reproduced from (Huang & Jianping, 2004)).	49

Figure 3-1: Site locations: signalized intersection of Stewart and Waller streets (Google maps, 2011).....	55
Figure 3-2: Camera position at the site location. (a) building and level at which the camera was housed. (b) video camera set-up.....	56
Figure 3-3: Field measurement distances at the signalized intersection.....	58
Figure 3-4: Visualization of camera calibration. (a) orthographic grid. (b) same grid projected on a sample video frame.	60
Figure 3-5: Flow chart diagram of cyclist video tracking system.	63
Figure 3-6: Video processing loop.....	64
Figure 3-7: Sample frames showing features identified by the tracker. (a) before the tracker system applied. (b) after the tracker system applied.....	65
Figure 3-8: Cyclist track measurement using the CWT method during the validation experiment at study area. (a) video sample frame. (b) test bicycle.....	67
Figure 3-9: The active cyclist crossing area at the site location.	68
Figure 3-10: The field common cyclist paths measurements at the site location. (a) shows the reference paths. (b) and (c) show the color marks used in the field to delineate the paths.	69
Figure 3-11: Schematic diagram illustrates object features: (a) features ($t1, t2, \dots, tn$) of the moving object in frame i . (b) new positions of moving features in frame $i+1$	71
Figure 3-12: Relationship between cyclist field-measured speed and the error value between cyclist field-measured and video-measured speeds (483 crossing events).....	74
Figure 3-13: Relationship between cyclist field-measured and video-measured Speeds.	75
Figure 3-14: The frequency distribution of 483 cyclists' field-measured speed.	76

Figure 3-15: The frequency distribution of 483 cyclists' video-measured speed.....	76
Figure 3-16: Average cyclists' speed distribution used different measurement techniques.	82
Figure 3-17: Average cyclist's speed distribution at the road sections and signalized intersections.	82
Figure 4-1: Location of the study intersections (Google, 2013).....	88
Figure 4-2: Camera views. (a) camera view of the Lyon-Laurier intersection. (b) camera view of the Bay-Laurier intersection.	88
Figure 4-3: Camera level positions at the site location. (a) at Laurier-Lyon intersection. (b) at Laurier-Bay intersection. (c) camera setup.	90
Figure 4-4: AutoCAD intersections drawing. (a) Laurier-Lyon intersection. (b) Laurier- Bay intersection.	92
Figure 4-5: Illustration of camera calibration. (a) field distance measurements. (b) projected distance using camera calibration. (c) orthographic grid. (d) world grid at the area of study (signalized intersection at Laurier and Lyon streets).	93
Figure 4-6: Illustration of camera calibration. (a) field distance measurements. (b) projected distance using camera calibration. (c) orthographic grid. (d) world grid at the area of study (signalized intersection at Laurier and Bay streets).	94
Figure 4-7: Schematic diagram: right-turn and left-turn interactions.....	95
Figure 4-8: Four different cyclist-motor vehicle interaction scenarios.	96
Figure 4-9: Object features, bounding box, nodes and mask in the tracker system. (a) the masked area, and the bounding box. (b) the cyclist features and bounding box nodes. (c)	

the motor vehicle features and the bounding box. The coordinates in world and the x-axis increases to the right, y-axis increase to the bottom.	99
Figure 4-10: Schematic of the distance between different positions of vehicle’s nodes and different positions of cyclist’s nodes.	106
Figure 4-11: Example critical conflict node and correct conflicting part of an interacting vehicle. (a) motor vehicle direction. (b) equivalent point of “y”, as denoted point “y”. (c) right conflicting part of the interacting motor vehicle.	109
Figure 4-12: Scatter plot of PET measurements using the AM and MFCM methods....	112
Figure 5-1: Location of the study intersections (Google, 2013).	120
Figure 5-2: Laurier- Metcalfe signalized intersections. (a) photo taken at the intersection. (b) AutoCAD intersections drawing.	120
Figure 5-3: Location of the control (un-segregated) signalized intersections. (a) Lyon- Somerset intersection. (b) Murray-Sussex intersection (Google, 2013).	121
Figure 5-4: Lyon- Somerset signalized intersection. (a) photo taken at the intersection. (b) AutoCAD intersection drawing.	122
Figure 5-5: Murray- Sussex signalized intersection. (a) photo taken at the intersection. (b) AutoCAD intersection drawing.	122
Figure 5-6: A high definition camera (image resolution is 1920 x 1080).	123
Figure 5-7: Mobile camera (image resolution is 320 x 240).	124
Figure 5-8: Stationary camera (image resolution is 640 x 480).	124
Figure 5-9: Illustration of the safety hierarchy (reproduced from (Hydén, 1987)).	128
Figure 5-10: Observation methodology attributes.	129
Figure 5-11: Cycling against traffic violation.	134

Figure 5-12: Cyclists stopping beyond stop-line.	134
Figure 5-13: Cyclists using the crosswalk.	135
Figure 5-14: Cyclists stopping out of path.....	135
Figure 5-15: Two examples of two cyclists riding out of path.	136
Figure 5-16: Cyclist crossing on red signal.	137
Figure 5-17: Two examples of vehicle turn violation. (a) vehicle turns to the left and stopped on cyclist's path. (b) vehicle turns to the right and stopped on cyclist's path...	138
Figure 5-18: Two examples of vehicle fail to yield to the cyclist. (a) vehicle turns to the left and fails to yield to the cyclist. (b) vehicle turns to the right and fails to yield to the cyclist.	139
Figure 5-19: Schematic diagrams: Evasive actions by cyclists in a interaction situation. (a) and (c) swerving without violation. (b) swerving with riding out-of-path violation. (d) swerving with ride in cross-walk violation.	140
Figure 5-20: Cyclist volume as a candidate measure of exposure.....	143
Figure 5-21: Turning vehicles as a candidate measure of exposure.	143
Figure 5-22: The product of turning vehicle volume and through cyclist volume as a candidate measure of exposure.	144
Figure 5-23: Mean residual life plot (Metcalf intersection).....	176
Figure 5-24: Threshold stability plot for scale and shape parameters (Metcalf intersection).....	177
Figure 5-25: Probability and quantile plot (Metcalf intersection).	178
Figure A-1: Sample of video recording information methodology.	178
Figure A-2: Sample of the tracker system configuration file.	178

Figure A-3: Video frame samples: before and after tracker system application.	220
Figure B-1: Object bounding box. (a) before adding the bounding box with center tracks only. (b) after adding the bounding box.....	222
Figure B-2: Features of an object and the connections, edges, between them.	223
Figure B-3: The rear of a vehicle and a pedestrian being over-grouped as one object...	224
Figure B-4: No over-grouping between objects.....	226
Figure B-5: Video mask. (a) before adding the mask. (b) after adding the mask.....	178
Figure B-6: Video restrictive mask. (a) before adding the restrictive mask. (b) after adding the restrictive mask.	229
Figure B-7: Configuration file before the improvement.....	231
Figure B-8: Configuration file after the improvement.....	232
Figure B-9: Explanation of tracked and untracked object.	233
Figure B-10: Sample output txt of an object with partial tracks, the data loss are represented with a? symbol.....	234
Figure B-11: Visualization video sample with two different ids (120, 194) for one object.	236
Figure C-1: Layout of Laurier- O'Connor signalized intersection.	238
Figure C-2: Layout of Laurier- Bank signalized intersection.	238
Figure C-3: Layout of Laurier- Kent signalized intersection.....	239
Figure C-4: Layout of Laurier- Lyon signalized intersection.	239
Figure C-5: Layout of Laurier- Bay signalized intersection.	240
Figure C-6: Layout of Laurier- Percy intersection.	240

Figure D-1: Probability and quantile plots of Laurier-O’Connor signalized intersection.	241
Figure D-2: Probability and quantile plots of Laurier-Bank (eastbound) signalized intersection.....	241
Figure D-3: Probability and quantile plots of Laurier-Bank (westbound) signalized intersection.....	242
Figure D-4: Probability and quantile plots of Laurier-Kent signalized intersection.	242
Figure D-5: Probability and quantile plots of Laurier-Lyon signalized intersection.....	243
Figure D-6: Probability and quantile plots of Laurier-Bay signalized intersection.....	243
Figure D-7: Probability and quantile plots of Laurier-Somerset signalized intersection.	243
Figure D-8: Probability and quantile plots of Laurier-Sussex signalized intersection. ..	244

List of Appendices

Appendix A : Sample of Tracker Inputs and Outputs for Speed Measurement	218
Appendix B : Tracker System Improvement	221
1- Object position.....	221
2- Speed parameter	223
3- Mask operation	227
4- Minimum features.....	230
5- Configuration file	230
6- Partial tracks	233
7- Stitching Object	235
Appendix C : Layout of Study Site Intersections	238
Appendix D : Graphical Diagnostic of Probability and Quantile Plots	241
Appendix E : Published Papers and Conference Presentation.....	245

1. INTRODUCTION

1.1 Background

Cycling is a popular and energy-efficient mode of travel used throughout the world (Giancarlo, *et al.*, 2010). The popularity of cycling is growing as an attractive mode of urban travel (Litman, *et al.*, 2006). Cycling offers many benefits to the public for example; cycling can relieve pressure on existing urban roads due to usage of road surface (Parkin, *et al.*, 2007; Koorey & Mangundu, 2010). Also, cycling activities have a negligible environmental impact (Rezendes, 2006; Kendrick, *et al.*, 2011). In addition, cycling offers benefits to the users, such as imposing a positive health effect and fitness improvements (Koorey & Mangundu, 2010).

The safety of cyclists is becoming a national and a worldwide concern. For example, Transport Canada reported that across Canada, 60 cyclists were killed in traffic accidents with motor vehicles each year during the period from 2004 to 2006 (Transport Canada, 2010). Therefore, it is imperative for the road designers to have reliable information on the cyclist behaviour and characteristics to design safe and efficient cyclist facilities. The behaviour of the cyclist can be measured in different ways, including, cyclist crossing speed, cyclist violations, and cyclist's-motor vehicle interactions.

Different techniques can be used to measure the speed of cyclists, and can be classified into manual, semi-automated and automated techniques. The manual technique requires the presence and participation of an observer during data collection and data extraction, which makes it an expensive and labour-intensive technique. The semi-automated and automated techniques do not require the presence of the observer during

the field data collection. The boundary between semi-automated and automated techniques can be a challenge to define in precise terms, but what sets them apart is the amount of human input during both stages of data collection and processing. Several devices can be used for the purpose of measuring the speed of cyclists, *e.g.*, radar gun, stopwatch, pneumatic tube counter, Global Positioning System (GPS), bike speedometer, and video camera.

Several previous studies focused on the measurement and analysis of the speed of cyclists in different route configurations, *e.g.*, urban trails used by hikers, cyclists and joggers (Virkler & Balasubramanian, 1998), public roads (Akira, *et al.*, 2012), roundabouts (Zaki, *et al.*, 2013), and signalized intersections (Lee, *et al.*, 2011; Lucas & Kanade, 1981). It is important to note that the majority of the studies which focused on measuring cyclist speed are based on manual and semi-automated speed measurements; either in the field or from video recordings. This is a significant challenge when large samples are required. In this thesis, automated image analysis from a video camera using computer vision was used to measure cyclist speed. This automated technique enables the automated observation of large volume of naturalistic road user's movements in an accurate and resource-efficient manner. Computer vision is defined as "*enterprise of automating and integrating a wide range of processes and representations used for vision perception*" (Ballard & Brown, 1982).

Besides behavioural data, the assessment of road safety can be based on collision data or on surrogate measures of safety. Traditionally, road collision statistics are considered as the main data source for road traffic safety analysis (Chin & Quek, 1997; Lareshyn, *et al.*, 2010). Thus, cyclist-involved collisions were investigated in several

cyclists safety studies at signalized intersections (Huang, *et al.*, 2012; Jensen, 2008; Wang & Nihan, 2004; Wang, *et al.*, 2008), un-signalized intersection (Summala, *et al.*, 1996; Räsänen, *et al.*, 1999), roundabouts (Sakshaug, *et al.*, 2010; Mette & Tove, 2008) and road-cyclist-path intersections (Hunter, 2000; Wood, *et al.*, 2009). An alternative approach to road safety assessment is based on surrogate measures of safety. The motivation for the use of such surrogate measures of safety in analyzing road safety in general, and more specifically in examining cyclist safety, is the inherent limitations in collision data. These limitations can be quantitative (scarcity of collision data or under-reporting) and/or qualitative (inaccurate geo-coding, incomplete records or limited level of detail about the collision event).

Surrogate safety measures are a group of techniques for road safety analysis which depend on data other than road collisions (Ismail, 2010). These measures have been used in practice to assess the safety of current and proposed roadways (Sayed, *et al.*, 1994; Persaud & Mucsi, 1995; Huang & Pant, 1994; Rao & Rengaraju, 1998). There are several surrogate measures to evaluate the safety of road users, *e.g.*, traffic conflicts (Parker & Zegeer, 1989; Thompson & Perkins, 1983) speed, delay and accepted gaps (Tarko, *et al.*, 2009).

Traffic conflicts can be marked either by evasive actions or temporal proximity between road users. A representative definition of evasive action based traffic conflict is “*an event involving two or more road users, in which the action of one user causes the other user to make an evasive maneuver to avoid a collision*” (Parker & Zegeer, 1989). A representative definition of proximity based traffic conflict is “*an observable situation in which two or more road users approach each other in space and time to such an extent*

that there is a risk of collision if their movements remain unchanged' (Amundsen & Hydén, 1977). However, errors were common, when the observers decide/judge if a given traffic event is an interaction or not. In addition, traffic interaction studies depending on manual counting are time-consuming, labour-intensive and prone to observation inaccuracies. Therefore, to eliminate the subjectivity from traffic interaction analysis, different objective proximity indicators have been suggested as surrogate measures of safety (Songchitrusa & Tarko, 2006; Gettman & Head, 2003; Archer, 2004; Kiefer, *et al.*, 2005; Vogel, 2003). Traffic conflicts will then be registered if the values of conflict measure are less than a predetermined threshold (Sayed *et al.*, 2013).

Among the different proximity indicators available in the literature, Post-Encroachment Time (PET) offers a simple way to measure proximity without the need for the prediction/extrapolation of future positions or speeds of road users. This indicator can objectively measure the proximity of road users to collision in situations where a collision was not observed (Tarko, *et al.*, 2009). In the context of cyclist safety, PET between a cyclist and motorist can be defined as the time difference between the moment at which the first road user leaves an area of potential collision and the moment at which the second road user arrives at this area while possessing the right of way (Allen, *et al.*, 1978).

1.2 Problem Definition

The design and regulation of dedicated urban cycling facilities require careful analysis of safety considerations, construction cost, environmental impacts, and other operational requirements. However, the level of engineering knowledge available for the design of cyclist facilities is yet to compete with that of other modes of travel. One of the

key shortfalls in engineering knowledge lies in understanding cyclist behaviour, and safety implications as a result of the design decisions.

Techniques used in the literature by utilizing traditional devices in speed data collection for pre-selected (participating) cyclists may not be representative of the general population. Therefore, it becomes more appropriate to use a technique that offers an appealing solution to cover a large-scale of speed measurements. Also, if cyclists were aware of the process of measurement, such as the case in using Global Positioning System (GPS) receivers, the results might be biased and would represent non-naturalistic cyclist behaviour. In addition, traditional devices will not provide an overall view of the study area to include all road user's with no ability of revisiting the data in the future for more analysis. A more efficient technique, in terms of resource utilization and sample size, is presented in this thesis.

An available automated open- source feature-based vehicle tracking system that uses computer vision techniques (Google, 2011; Saunier & Sayed, 2006; Lucas & Kanade, 1981) needed to be improved in order to be used in this thesis to track vehicles and cyclists based on several shortcomings. Shortcomings such as; [i] providing only one position to the object (road users), thus lacking the ability to provide a detailed profile of the object, [ii] over-segmentation, which means that more than one object can be seen representing real road users, [iii] over-grouping, when two or more different objects are grouped or merged, incorrectly, by the tracker system and considered as one object, [iv] partial tracks, which occur when the number of features goes below the threshold set by the user for the object to be tracked and detected. Many improvements have been made to the tracker system in this thesis by adding new functionalities and mechanisms.

For safety evaluation, depending solely on the reported collision data for traffic safety analysis would introduce several possible shortcomings. First, collisions are relatively rare events. Therefore, the safety evaluation requires extended observation time to monitor stable trends. For example, a before-and-after safety evaluation is typically based on a few years of collision records in order to detect changes in safety. Second, it is relatively difficult to observe, record (by means of a video camera or auditory sensors), and analyze the failure mechanism leading to collisions in comparison to traffic conflicts (Hydén, 1987). Third, underreporting of collisions, the level of which depends on collision severity and the type of road users involved, is another shortcoming of reported collisions data (Laureshyn, *et al.*, 2010). Fourth, collision data does not typically provide complete information about the collision process. For example, pre-collision road users' behaviour is usually not reported (Laureshyn, *et al.*, 2010). Important developments in the data collection and statistical analysis of surrogate measure of safety for cyclists are achieved in this thesis.

1.3 Research Objectives

Based on the previous section, the main study objectives are described as follows:

- 1- To examine the strengths and weaknesses of the techniques used by the traditional devices utilized in previous research to measure cyclists speed.
- 2- To improve the automated open-source version of the tracker system to overcome the shortcomings which were mentioned in Section 1.2.
- 3- To develop different analysis techniques to measure cyclist crossing speed. All techniques need to be non-intrusive and should be capable of collecting a large dataset.

- 4- To identify and impact of potential factors affecting cyclist crossing speed.
- 5- To develop a method for automated analysis of cyclist-vehicle PET to assess cyclist safety using video analysis techniques and examine its accuracy. A PET of (0, 3] second is considered as a threshold of interactions that will be studied in this thesis.
- 6- To examine the reliability of the PET as a surrogate safety measure and the factors affecting using different statistical techniques.

1.4 Research Contributions

The following contributions are achieved in this thesis:

- 1- Video tracking system was improved from an open-source feature-based vehicle tracking system in order to track cyclists and vehicles and produce a trajectory database.
- 2- Automated and semi-automated analysis techniques were developed to measure cyclist crossing speed. The focus on cyclist speed was motivated because: [i] speed is one of the most important behavioural parameters which reflect mobility and perceived safety, [ii] cyclist speed is utilized in traffic signal design when dedicated phases are allowed to cyclists.
- 3- The effect of different potential factors on cyclist crossing speed was investigated.
- 4- Automated method to measure Post Encroachment Time (PET) between cyclists and motor vehicles using video analysis techniques was developed and validated.
- 5- Different statistical techniques were used to investigate PET observations to measure a cyclist safety. In addition, evaluations based on other surrogate measures of safety, *e.g.*, violations were investigated.

In total, this thesis makes a significant contribution to the analysis techniques used for behavioural and safety studies of cyclists. The thesis also produced a sizeable database of 806 hours of video data for cyclist movements. A total of 19,068 cyclists and 48,632 vehicles were observed in this study. The observations were made within a total period of 57 days (30 days in fall season and 27 days in summer season) and covered 10 intersections.

1.5 Organization of the Thesis

This thesis is broken down into six chapters to cover main component of this study. Chapter One presents the introduction. Studies related to cyclist's speed measurements and evaluating cyclist safety are presented in Chapter Two. The reliability of using computer vision techniques in measuring cyclist speed and different factors affecting cyclist speed are investigated in Chapter Three. In Chapter Four, an automated method to measure PET between cyclists and motor vehicles using automated video analysis techniques is presented. In addition, testing the accuracy of the automated video analysis techniques compared to manual methods is also presented in Chapter Four. While in Chapter Five, different statistical techniques were used in order to measure a cyclist safety. Finally, the main conclusions with the summary and recommendations of this study are highlighted in Chapter Six.

2. LITERATURE REVIEW

2.1 Introduction

The occurrence of cyclist-involved collisions has recently caught the attention of decision makers and transportation engineers. Cyclist risk exposure can be classified into three types; risk of common occupancy of space with motor vehicles, aerodynamic effect of truck/bus passing a cyclist and risk of collision while turning (Khan & Langlois, 2011). The growing importance of improving cyclist safety has been motivated by grassroots support for sustainable and energy-efficient modes of travel. Reducing the incidence as well as the consequences of this type of collisions requires an adequate knowledge about all factors affecting cyclist collisions. It is commonly recognized that most of those factors can be classified as attributable, in part or in total, to the characteristics of: the cyclist, the driver, the roadway or traffic facility, and the vehicle. While hypothetically, cyclist-involved collisions, as shorthand will be hereafter called cyclist collisions, could involve cyclist and different types of road users, serious collision typically involve cyclists and motor vehicle. This will be the focus of the research described in this thesis.

The technical literature contains several studies on the evaluation of various approaches to improve cyclist safety. These approaches can be classified as [i] public education, [ii] enforcement, and [iii] engineering measures related to the physical aspects of the roads and/or traffic control. Despite the importance of public education, and enforcement approaches, they are not considered within scope of this research. The focus of the research will be related to the transportation engineering approach to create safety treatments. This approach concerns modifications to the road infrastructure and traffic

control to enhance cyclist safety. Examples, of engineering safety treatments are segregated bike lanes, sharrows, shared paths, signage, and bike boxes. These treatments will be defined and reviewed in subsequent sections. In this chapter, peer-reviewed studies related to cyclist safety are reviewed. Studies that are concerned with the first and second approaches, education and enforcement, will not be considered in this review.

This chapter is organized in two main sections. The first section presents the review related to the topic of cyclist speed measurements. This review includes: techniques used to measure cyclist speed, advantages of the proposed video analysis technique relative to techniques used in the literature. The second section presents the review related to the topic of evaluating cyclist safety treatments. This review includes: evaluations of common safety treatments and a description of the advantages of the proposed conflict analysis technique relative to evaluation techniques in the literature. In addition, whenever available, critical aspects of reviewed studies are presented.

2.2 Cyclist Speed Measurements

The behaviour of the cyclist can be measured in different ways. One of those measures is cyclist crossing speed. Crossing speed may reflect the overall level of mobility offered to cyclists. This measure also influences the design of signal controls dedicated to cyclist safety. Furthermore, cyclist crossing speed is an important factor that affects the duration of exposure to the risk of collision with motor vehicles at urban traffic intersections.

Several previous studies focused on the measurement and analysis of speed of cyclists in different route configurations, *e.g.*, urban trails used by hikers, cyclists and joggers (Virkler & Balasubramanian, 1998), public roads (Akira, *et al.*, 2012),

roundabouts (Zaki, *et al.*, 2013), and signalized intersections (Lee, *et al.*, 2011; Lucas & Kanade, 1981). Knowledge of cyclist speed can be important in the following contexts:

- 1- Designing safe and efficient cyclist facilities. For example, cyclist crossing speed, which is the cyclist speed when crossing an intersection, is a key parameter in the timing of traffic signal intervals serving heavy cyclist volumes and/or dedicated exclusively for cyclists.
- 2- Improving cyclist safety by reducing interaction severity between cyclists and vehicles. The knowledge of cyclist speed is important in assessing the proximity to collision, *e.g.*, in calculation of time-to-collision.
- 3- Studying cyclist behaviour at different locations and its impact on the occurrence of any speed violation or abnormality in speed profiles. In addition, cyclist speed can be a good indicator of perceived safety at the intersection; faster crossing probably reflects a high safety perception.
- 4- Selecting and evaluating cyclist lanes. Cyclist speed is considered an important factor that affects the choice of appropriate cyclist lanes, *e.g.*, marked lane, segregated lane (generally separated from motorized traffic), and shared path with other road users. More facility segregation, either by built barriers or wider lanes, can allow faster cyclist speed. Therefore, the effectiveness of cyclist lanes can be gauged by measuring cyclist speed along them.
- 5- Cyclist crossing speed can be an important parameter in assessing intersection sight distance.

There are different devices used to measure the speed of cyclists, *e.g.*, radar gun, stopwatch, pneumatic tube counter, Global Positioning System (GPS), bike speedometer,

and video camera. It is important to note that the majority of the studies which focused on measuring cyclist speed are based on manual speed measurements; either in the field or from video recordings. This is a significant challenge when large samples are required and when continuous (or every short time period) speed measurement is required, especially along curved paths. Selecting an appropriate technique depends on several factors, *e.g.*, sample size, required accuracy, and cost.

The objective of this section is to investigate whether there is a consensus on design values of cyclist speed. This section critically reviews the techniques reported in the literature to measure cyclist speed. Furthermore, this review covers the accuracy of the measurement devices and highlights limitations of the reviewed studies. These identified limitations are classified as: limited range of movement directions, selection method of observed cyclists, seasonal variation in the measurements, completeness of results reporting, reporting of equipment limitations, and measurement validation. The study further compares the performance of these measurement techniques against automated computer vision techniques. This comparison is carefully made to take into account the particularities of cyclist movements.

The following sections categorize speed measurement techniques into manual, semi-automated, and automated technique. A number of techniques were reported in the literature to measure cyclist speed while utilizing some level of automation in recording the positions of cyclists and/or travel time. The boundary between semi-automated and automated techniques can be a challenge to define in precise terms. In this review, what sets them apart is the amount of human input during both stages of data collection and

processing as explained earlier. It will be further observed that automated techniques are less frequently applied, potentially much less than is hoped in practice.

2.2.1 Cyclist speed measurements

Any technique involving a human input during the field data collection and analysis to estimate cyclist speed is defined as a manual cyclist speed measurement technique. In this technique, the devices that were commonly used were a stopwatch, measuring tape and a radar gun. This type of the technique requires the presence and participation of the observer during data collection and data extraction. A few manual techniques to measure cyclist speed were reported in the literature. These techniques relied mainly on the manual measurement of time that elapses while a cyclist is traversing a segment of known length.

In a study by (Thompson, *et al.*, 1997), a radar gun and a stopwatch were used to measure speeds of cyclists belonging to different age group along a road closed to motorized vehicles during a recreational event in Seattle, Washington. Age and gender were documented for cyclists who agreed to participate in the study. Measurement of cyclist speed using a radar gun and a stopwatch were conducted along a 91.44 meter (300 feet) flat road section. One hundred and fifty two cyclists were observed in this event over weakened days. The stopwatch was used to measure the time required for a sample of cyclists to cross the road section. Two observers were placed at each end of the road segment and they were provided stopwatch and a means of communication. The radar gun was installed on a tripod and positioned 0.30 – 0.61 meter (1-2 feet) from the cyclist path. The results found that the majority of riders were travelling at a relatively moderate speed ranging from 8 to 24 km/h, with a mean speed for all ages of around 14.8 km/h.

The mean speed for females and males were found to be 14.16 km/h and 15.45 km/h respectively. Results found that there was $\pm 4\%$ error radar reading at 11.27 km/h or less and no accuracy mentioned for higher speeds. The reported error is vaguely reported and does not clarify whether the error bounds are measured or estimated based on quantile values. The study did not comment on the difference in cyclist speeds as measured by the radar gun and stopwatch. Paired t-test was used to evaluate the difference between average measured and estimated speed. The estimated speed was obtained by a survey from the cyclists themselves. Results found that there was no significant difference between mean measured and estimated speed by individual year of age ($p = 0.1$ to 0.9).

In a study by (Virkler & Balasubramanian, 1998), a stopwatch was used to measure the time required by cyclists to cover a short section 8.0 – 16.0 meter of the trail road in two sites in Columbia, Missouri and Brisbane, Australia. The trail road in Columbia site was level and has a crushed rock surface, while the trail road in Brisbane site was level and has a smooth asphalt surface. The trail road width in Columbia and Brisbane sites were 3 and 2.9 meter respectively. Data was collected on weekend days at the Columbia site and on Tuesdays at the Brisbane site. The results showed that the mean cyclist speed in Columbia site was 21.42 km/h with standard deviation 7.56 km/h. The minimum and maximum cyclist speeds in Columbia site were 4.43 km/h and 36.76 km/h respectively. It was also found that the mean cyclist speed in Brisbane site was 20.74 km/h with standard deviation 4.79 km/h. The minimum and maximum cyclist speeds in Brisbane site were 7.31 km/h and 35.66 km/h respectively. The cyclist's counts in both Columbia and Brisbane sites were 35 and 105 cyclists respectively.

In a study by (Akira, *et al.*, 2012), an experimental bicycles was used in order to determine the free travel cyclist speed on flat public roads in Tokyo. The free travel cyclist speed was defined as the speed that is not affected by traffic volume, traffic signals, or road width. Results found that the average cyclist free travel speed for 120 cyclists was 16.8 km/h and the maximum was 21.7 km/h with standard deviation 2.1 km/h. At the narrow street, the results showed that the cyclist travel speed between 14 km/h and 15 km/h for 96 cyclists. In addition, the result showed that the cyclist free travel speed was almost 12 km/h when the cyclist shared the space with the pedestrian. The study did not comment on the difference in cyclist free travel speeds due to gender and age.

2.2.2 Semi-automated cyclist speed measurements

Semi-automated cyclist speed measurement technique does not require the presence of the observer during the field data collection. In this technique, the devices that were commonly used such as; a pneumatic tube counter, Global Positioning System (GPS) and image analysis frame by frame manually from a video camera. The following subsections present key methodologies and findings using semi-automated technique.

In study by (Hunter, *et al.*, 2009), a pneumatic tube counter “*MetroCount* device” was used to count and calculate the average speed of bicycles at two different locations in Petersburg, Florida. A MetroCount is a vehicle classifier/counting system that consists of counter device, rubber tubes, and computer. In this study, cyclist speeds were measured at two traffic corridors composed of a total of 19 intersections. Data was also collected before and after cyclist lanes were marked along these corridors. One of these corridors located in residential and a few commercial firms. The second corridor is located in a

commercial area. Two counter tubes were used at each intersection along each corridor in both directions and were placed as such to encompass the entire cyclist lane as shown in Figure 2-1. Before lanes were marked, the tubes were placed to cover about one-fourth of the traffic lane. The spacing between each two tubes was 0.99 meter (39 inches). Tubes were placed 30.48 to 60.96 meter (100 to 200 feet) approximately in downstream from each signalized intersection. Speed data was extracted from the counter device before and after cyclist lanes were marked. In order to achieve free flow traffic condition before the cyclist lanes were marked, the study excluded any observation from the speed where the gap between cyclist and vehicle was 3 second or less. Results showed that the average cyclist speed before and after the cyclist lane were 17.7 to 19.3 km/h respectively. It was also found that there was a significant change in cyclist speed at the level of 0.05 after cyclist lanes were added to both corridors. It can be concluded that this technique is more suitable for measuring the speed of the cyclist along a straight path.

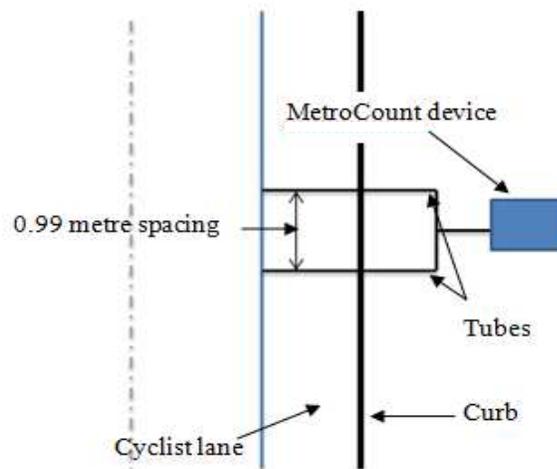


Figure 2-1: Schematic diagrams for the pneumatic tube counter device at the field after cyclist lanes were marked (modified from (Hunter, *et al.*, 2009)).

In a study by (John & Rotheram, 2010), a handle-bar Global Positioning System GPS devices were used and installed on the bikes of 16 cyclists who participated in this study. This was done in order to measure the speed and acceleration of cyclists during summer 2008 in Leeds, UK. A total of 518 cyclist speed observations were recorded. The effect of flat and different road grade on cyclist speed and on acceleration were studied. Two linear regression models were developed to represent the effect of different road grades on cyclist speed and on acceleration rates. The results showed that cyclist speeds ranged from 6 to 40 km/h, with a mean speed of 21.6 km/h and a standard deviation of 5.98 km/h. The results of linear regression models for speed and acceleration indicated that the effect of an upgrade on speed was more pronounced than that of a downgrade. It was also found that the mean speed is increased by 0.86 km/h for every additional 1% of downgrade. While the model suggests that the mean speed is reduced by 1.44 km/h for every additional 1% of the upgrade. The results showed that the implication of the linear model is that for a comparatively modest downgrade of 3%, the eighty-fifth percentile speed is 25 km/h. The study did not comment on the accuracy of GPS measurements. It is expected that the accuracy of the GPS data depends on the continuity of the satellite signal capture.

In a study by (El-Geneidy, *et al.*, 2007), an analytical model was developed to predict cyclist travel speed using real-time data obtained from GPS-equipped bikes in Minneapolis. Eight cyclists were involved in this study. Data was collected over three weeks in October 2005. On-street, off-street, and mixed traffic were three different scenarios of transportation facility considered to create two regression models for cyclist travel speed. An on-street; when the cyclists use a marked lane for cycling as a part of the

street. An off-street; when the cyclists use a dedicated path for cycling and walking. The off-street is not considered part of the street as shown in Figure 2-2. While the mixed traffic facility; when the cyclists use the regular streets and travel in mixed traffic without a dedicated path for cycling. The dependent variable in these models was cyclist speed, while the total trip length, segment length, number of signalized intersections, age, gender, and average daily traffic were used as independent variables. Independent variables such as total trip length segment length, and the gender were found positive and statistically significant on cycling speed in all models. Meanwhile, only the variable associated with off-street cycling, as one of the three scenario types tested, was found to be significant (at the level of 0.01) and positively related to cyclist speed. Both the number of signalized intersections, average daily traffic, and the age were not statistically significant in the models at the level of 0.01. The results indicated that the average cyclist travel speed ranged between 15.62 and 16.25 km/h. Predicted speeds under on-street and mixed traffic conditions were lower than those predicted for off-street conditions as shown in Figure 2-3.

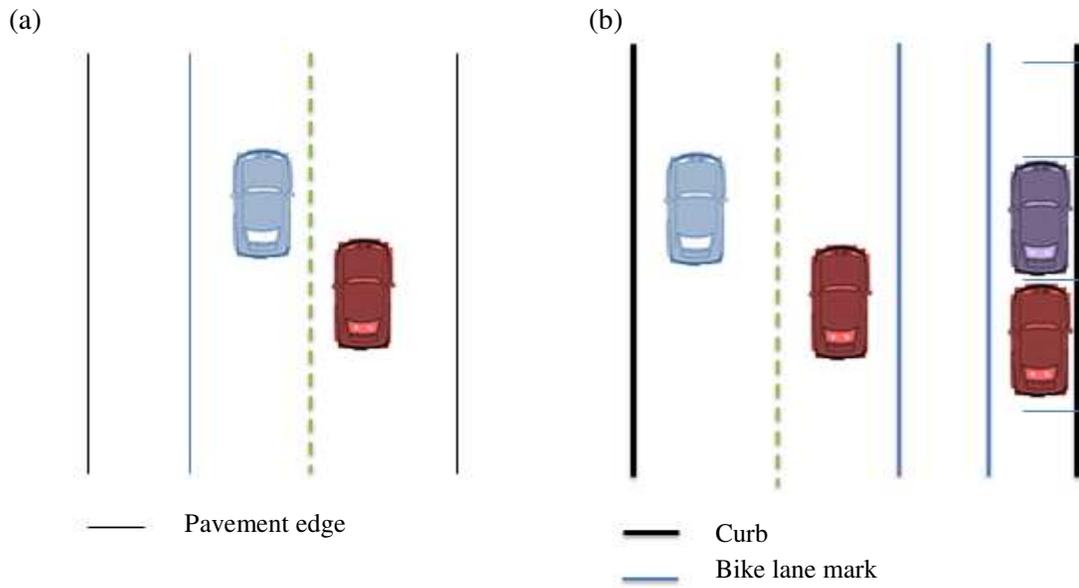


Figure 2-2: Schematic diagrams: (a) of off-street and (b) on street transportation facility (reproduced from (El-Geneidy, *et al.*, 2007)).

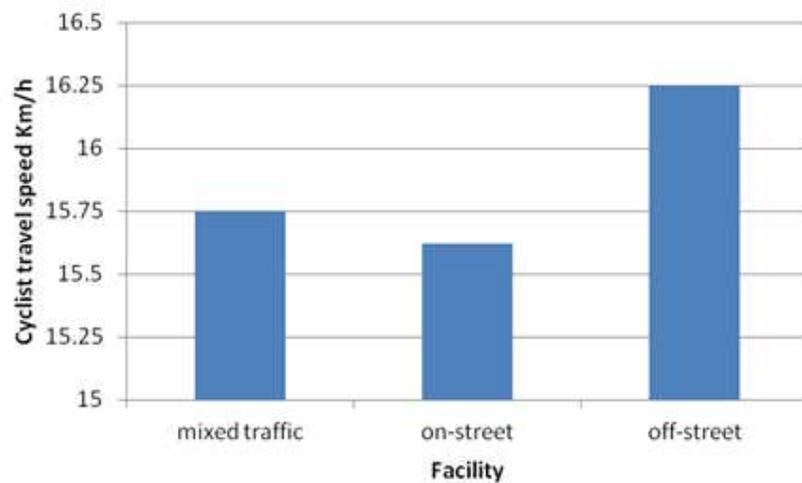


Figure 2-3: Predicted cyclist travel speeds under three different types of facility (reproduced from (El-Geneidy, *et al.*, 2007)).

(Dill & Gliebe, 2008), studied the behaviour of a regular sample of cyclists in the City of Portland, Oregon using GPS devices. The regular sample comprised frequent cyclists who bike at least one day per week. The purpose of the study was to help the

urban planner find the easiest ways to increase the human physical everyday activity by using a cyclist instead of motor vehicle. GPS devices were used and installed on the bikes of 164 cyclists who participated in this study. Data was collected over one week through March and November 2007. The results showed that the overall average speed was 17.38 km/h and the standard deviation was 3.2 km/h. While the averages running speed was 17.86 km/h. The average running speed only includes the time the cyclist is in motion. In addition, the results showed that men drove their bikes faster than women. The average running cyclist speeds for men and women were 18.67 km/h and 15.77 km/h, respectively.

(Witte & Wilson, 2004), studied the cyclist speed using GPS and bike speedometer. GPS accuracy was tested for cyclist travelling at constant speeds on running track and on a straight road. A non-differential GPS was attached to the cyclist helmet, while the speedometer was linked to the bike frame. The length of the running test track was 400 meter. The track consists of two curves with different radii and straight lines. A total of 5,060 GPS speed values were recorded. The results showed that the accuracy of GPS was within 0.2 m/s for 45% of the total observations, and within 0.4 m/s for a further 19% of the total observations. The results showed the cyclists speed was between 7.56 km/h and 38.88 km/h. It was also found that there was a significant change in cyclist speed error; where error is the difference between GPS and speedometer. The cyclist speed error decreased when the number of satellites used increased ($p < 0.000$). Results showed that the speed errors increased on circular paths, especially with the small radii, compared to a straight path. This increase in speed error is a result of a tendency to underestimate speed.

In a study by (Sarosh & Winai, 2001), video data collection and frame-by-frame annotation technique were used to determine cyclist speed along a selected street section on exclusive cyclist path in Denver, Colorado. The exclusive cyclist path length was 91.44 meter (300 feet) and 3.0 meter wide. Two cameras were used and placed on the sidewalk on the bridge crossing the exclusive cyclist path. Data was collected over two days and for six hours per day in April, 2000. Cyclist speed was calculated manually by identifying and transfer cyclist screen coordinate position to roadway coordinate position using a transformation technique every 0.5 second interval from captured video frames. The video frame capture rate (frame frequency) was 15 frames per second. The results of analyzing the speeds of 53 cyclists showed that the cyclist speed ranged from 17.17 to 35.9 km/h. It was also found that the average cyclist speed was 24.83 km/h with a standard deviation of 4.33 km/h.

(Huang & Jianping, 2004), studied the cyclist crossing speed at an urban signalized intersection in Beijing, China. Two video cameras were used to monitor left-turn and through cyclist movements. The video frame capture rate (frame frequency) was 25 frames per second. The traversing time was calculated by multiplying the number of frames that display the cyclist while traversing a 15 to 20 meter distance by the frame rate. The sample size of recorded through- and left-turn cyclist movements was made of 376 and 94 observations, respectively. The results showed that the cycling crossing speed for the through movement was in the range of 5.28 km/h to 20.41 km/h with an average value of 11.44 km/h. For the left-turn movement, the range of cyclist crossing speeds was between 6.94 km/h and 22.30 km/h with an average value of 13.21 km/h. The standard deviation of through and left-turn cyclist movement were 3.01 km/h and 3.33 km/h

respectively. Chi-square test was used to analyze speed distribution of cyclist speeds performing these two movements. The Normal distribution was found to fit the cyclist speed distribution using the Chi-square test. The results showed also that the average cyclist crossing speeds for male and female cyclists were 11.26 km/h and 11.58 km/h for through movement, and 12.91 km/h and 13.43 km/h for left-turn movement, respectively. The effects of gender did not appear to have a clear effect on average cyclist crossing speed at the level of 0.01. The reported overall average crossing speed was 11.44 km/h and 13.21 km/h for through and left respectively.

(Opiela, *et al.*, 1980), studied the cyclist behaviour at signalized intersections in order to measure cyclist crossing speed. Video cameras were used to monitor cyclists. The speed calculation technique utilized the number of frames and frame rate similar to the study by (Huang & Jianping, 2004). The results showed that the cyclist crossing speed ranged from 3.40 to 39.6 km/hour. Using the same technique to monitor the users of bicycles, (Wayne, 1997) measured cyclist crossing time over a known distance at 16 different intersections in Pinellas, Florida. Camera was placed on the roadway in order to cover a distance from 6.4 to 32 meter. The crossing time was measured using digitized timing with the accuracy of 0.1 second. The intersections were composed of two groups: signalized and un-signalized intersections. The results showed that the cyclist crossing speed for 442 cyclists was found to be 12.7 km/h. No statistical analysis was presented in the study for the difference of cyclist crossing speeds on signalized and un-signalized intersections.

(Singh, 2012), studied the behaviour of cyclists in term of speed and flow over a 19 to 24 meter road section in South Kensington and Chelsea. Video data was collected

before and after a new marked line was implemented. The marked line makes cyclists shared space with vehicles. Five cameras were used for data collection into two phases. First, two cameras were used before the marked line was implemented, covered 20 hours during weekday. Second, three cameras were used after the marked line was implemented are covered 30 hours during weekday. The time frame was 10 hours per day of data collection. A stopwatch was used to measure the time required for cyclists over the road section. The results showed that the mean cyclist speed of 187 cyclists before the construction of the new marked line was 22.3 km/h with a standard deviation 4.8 km/h. While the result shows that the mean cyclist speed of 194 cyclists in the vehicle zone after the new marked line implemented was 19.8 km/h with a standard deviation 4.3 km/h. It was also found that the mean cyclist speed of 20 cyclists in the pedestrian zone after the new marked line implemented was 16.2 km/h with a standard deviation 3.3 km/h. The results reflect that the shared space impact on cyclist speed. The cyclists that travelled in the vehicle zone cycled higher than the cyclists in the pedestrian zone and statistically significant at the level of 0.05.

(Davies, *et al.*, 2003), studied the behaviour of cyclists in vehicle restricted areas (VRAs). VRAs are defined as pedestrian areas or those areas of the road which are free from motor vehicles. Data was collected at four sites in three towns in England, Cambridge, Hull, and Salisbury using video camera. Data was collected for a weekday from 8:00 AM to 6:00 PM and 1,681 cyclists were recorded. Cyclist speeds were calculated using a stopwatch and a fixed distance in the video. The distance was between 8.0 to 13.0 meters. The speed analysis showed that the overall average speed was 16.0 km/h with standard deviation 0.2 km/h. The cyclist average speed in Cambridge was 16.7

km/h at Hull was 10.9 km/h, and at Salisbury was 12.1 km/h. The results showed that the cyclists change their behaviour in response to the change in the level of pedestrian flow. An increase in the level of pedestrian flow reduced the speed of cyclists travelling in the same areas.

2.2.3 Automated cyclist speed measurements

When the human input is relatively very limited during the field data collection and data extraction, the method is classified as an automated cyclist speed measurement. An example, automates image analysis from video camera using computer vision technique. Computer vision is defined as “*enterprise of automating and integrating a wide range of process and representations used for vision perception*” (Ballard & Brown, 1982). In this technique, data can be retrieved after recording using a video camera. In addition, it does not require the presence of the observer during data collection and data extraction. The use of computer vision techniques for measuring cyclist speed is relatively recent. This technique relied mainly on an automated tracking of cyclists and other road users.

Recently, a study by (Zaki, *et al.*, 2013), computer vision was used in order to count cyclists volume and to measure cyclists speed from automated video analysis. Data was collected for two days from 9:00 AM to 9:00 PM on March 2011 at the roundabout in the main campus of the University of British Columbia, Canada. A video camera was used and 734 cyclists were recorded. Two scenarios of cyclist group size were studied. The first scenario, when there is only one cyclist in the group. The second scenario when there is more than one cyclist in the group. Results showed that the cyclist mean speed according to the first scenario was 16.78 km/h and the standard deviation was 13.93 km/h. While according to the second scenario, the cyclist mean speed was 14.26 km/h

and the standard deviation was 12.6 km/h. In order to validate the results using the computer vision technique, 70 cyclists were selected. Results showed that the manual calculated average cyclist speed for 70 cyclists was 14.1 km/h. The time that's required for the cyclists to travel along a measured distance was extracted manually from the video footage. The average cyclist speed using the automated system was found to be 13.41 km/h. The evaluation result showed that the coefficient of determination, R^2 , between the automated and manual methods was found to be 0.86.

2.2.4 Identified limitations

The studies reviewed in the previous section utilized different techniques to measure cyclist speed. This section presents a summary of the identified limitations of the reviewed studies and cyclist speed measurement techniques. This critical presentation cites previous studies along with identified patterns of limitations.

2.2.4a Limited movement directions

A few of the reviewed studies focused on curvilinear movements, *e.g.*, while making a left-turn or right-turn. This may be due to limitations in the technique that is used. Traditionally, manual methods that were used to measure the cyclists speed and require to presence the observers in the field, they did not measure the cyclist speeds in different directions at the same time of the measurement. For example, studies by the authors (Thompson, *et al.*, 1997; Virkler & Balasubramanian, 1998; Akira, *et al.*, 2012) studied the cyclist speed in a straight road section. This is due to the [i] use of the devices where they can only measure the cyclist speed in a straight line (*e.g.*, stopwatch and pneumatic rubber tube), or [ii] use measurement devices suitable for spot or momentary speed observations (*e.g.*, radar gun).

2.2.4b Cyclists selection

The sampling strategy of selecting cyclists may not be representative of the general population. This shortcoming is pronounced when manual or semi-automated techniques are used. The observed cyclists may have to consent to participate in the study, *e.g.*, using GPS, which may cause selection bias. For example, studies by (Thompson, *et al.*, 1997; Virkler & Balasubramanian, 1998; Akira, *et al.*, 2012; Hunter, *et al.*, 2009), cyclists were asked to participate in the studies, in order to measure cyclist speed. The use of GPS equipped to the bikes can lead to speeds which may not be from naturalistic behaviour. Furthermore, sampled cyclists may have been engaged in the data collection effort due to shared attributes. This leads to potential sample bias. Unfortunately, there is no simple remedy to this limitation for techniques which require conscious participation of observed cyclists.

2.2.4c Seasonal variation

All the studies that were reviewed may have in common a shortcoming of the limited observation period. That is, they do not capture seasonal variation in cyclist behaviour. Due to practical difficulties, researchers often use a short and concentrated period of observation to represent the actual field conditions. For example, study by (John & Rotheram, 2010), the data was collected during summer without mentioning the total hours or days, studies by (Thompson, *et al.*, 1997; Sarosh & Winai, 2001; Zaki, *et al.*, 2013), the data was collected within two days, studies by (El-Geneydy, *et al.*, 2007; Dill & Gliebe, 2008), the data was collected within a few weeks. While studies by (Akira, *et al.*, 2012; Witte & Wilson, 2004; Huang & Jianping, 2004; Opiela, *et al.*, 1980), didn't mention the time that the data was collected. Few studies in the literature

review mention time-of-day but fail to show the differences between cyclist speeds during peak and non-peak periods (Virkler & Balasubramanian, 1998; Singh, 2012; Davies, *et al.*, 2003).

2.2.4d Completeness of result reporting

Some of the reviewed studies suffered from incomplete or inadequate reporting of study methodology and/or results. For example, the studies by (John & Rotheram, 2010; El-Geneidy, *et al.*, 2007), did not report the characteristics and level of accuracy of the GPS receivers used or the accuracy of the speed measurements. In the study by (Wayne, 1997), no comments were presented regarding the statistical significance of the difference between average cyclist crossing speed in signalized and un-signalized intersection.

2.2.4e Device limitations

Different devices may be used for collecting data related to cyclist speed, such as GPS, radar gun, stopwatch, and video camera. Each one has its own advantages and disadvantages and can be selected depending on the project budget, sample size, and level of accuracy. The justification of selecting specific measurement equipment was not adequately described in most reviewed studies.

When the traffic is heavy, the radar gun is not a reliable instrument to measure the speed of road users with acceptable accuracy. Radar gun was used by (Thompson, *et al.*, 1997) and positioned between 0.30- 0.61 meter (1-2 feet) from the cyclist path to measure cyclist speed. It is noted that this distance was chosen for the purpose of preventing any interference or disconnected the signal from the radar gun and thus will affect the measurement results. In this case, the behaviour of the cyclists may be affected by the

visibility of the radar gun. It is impossible to re-measure or to check the data measurement of the events using the radar gun. It is difficult to measure a road user speed along the track because the radar gun can be used to measure the spot speed only of the road users. The radar gun cannot be used to study the relationship between different road users, *e.g.*, study the interaction between road users. Therefore, the radar gun is suitable for low volume condition and during peak period will be difficult. In addition, the radar gun signal is affected by the size of the road users, *e.g.*, large vehicle reflects a strong return signal (Duane, *et al.*, 2002).

Although, using a stopwatch is a quick and inexpensive method for collecting speed data, but it still remains not very accurate. The accuracy of the measurement is dependent on the accuracy of the observers. When one observer will be used for collecting data in order to reduce the percentage of the error between different observers, the observer should be positioned higher than the study area. Therefore, this method is not appropriate in all location and under different weather conditions.

Although, the pneumatic tube sensors is used for longer data collection time period than the radar gun and stopwatch but this method is still not suitable to measure the speed of road users along a track. The pneumatic tube sensors can provide only spot measurements. In study by (Hunter, *et al.*, 2009), the counting of cyclist is mentioned in more details more than the measurement of cyclist's speed using a pneumatic tube sensor. Therefore, this method is more suitable for counting road users than the measurement of road user's speed. In addition, it is difficult to apply pneumatic tube sensors in shared lane (Federal Highway Administration, 2011). It is difficult to classify road users using this method.

Traditionally, the cost of using a radar gun and pneumatic tube sensors are more than other methods, *e.g.*, stopwatch and video camera. The radar gun and pneumatic tube sensors are always in need of maintenance, furthermore the radar gun method requires the presence and participation of the observer during data collection. This is considered an additional labour cost compared by collecting data by installing video cameras.

A GPS was a good device to measure the travel cyclist speed, especially in open areas and for a long trip (John & Rotheram, 2010; El-Geneidy, *et al.*, 2007; Dill & Gliebe, 2008). An area with denser high buildings causes missing GPS observation data, more intersection and signal delay (Christopher & Min, 2009). The accuracy of the GPS data depends on the continuity of the satellite signal capture. It is known that the positioning accuracy can be impacted within urban canyons and in the vicinity of high buildings. Furthermore, GPS is a device for speed measurement which collects data specific to cyclists participating in the survey. The accuracy of the GPS can be affected by weather changes and the deflection of the obstructions signal as well as the speed measurement errors increase around curved paths (Townshend, *et al.*, 2008; Gray, *et al.*, 2010).

2.2.4f Measurement validation

Data validation is an important tool for quality control and validation of results. Validation of field measurement can be performed by comparing the results of using different measurement techniques and what may be regarded as near-error-free measurements. All of the reviewed studies did not include a specific section on measurement validation and accuracy assessment except (Thompson, *et al.*, 1997; Witte & Wilson, 2004; Zaki, *et al.*, 2013).

2.2.5 Case for computer vision techniques

The automated cyclist speed measurements using computer vision techniques might overcome these limitations. Monitoring movements of road users are non-intrusive (as opposed to GPS and speedometers) and very likely will result in analyzing naturalistic cyclist movements of random, not selected cyclists. Observed subjects are typically unaware of video monitoring, especially when cameras are properly installed. This advantage has made video recording a commonly used technique to observe the behaviour of road users. Any mistakes during the cyclist speed measurement from the video can be repeated many times without affecting the behaviour of the cyclist. The variability of cyclist speed with time can be measured along straight as well as curved paths. This is possible because computer vision techniques can report speed every frame. The typical frame rate is approximately 30 frames per second, with some cameras capable of much smaller time resolution. Video analysis permits continuous tracking of cyclists to calculate a sequence of instantaneous speed observations over a finite period of time. Lastly, the accuracy of computer vision techniques for speed measurement is promising and is actively improved.

2.3 Approaches to Study Cyclist Safety

The safety of vulnerable road users, such as cyclists and pedestrians, is becoming a national and a worldwide concern. For example, Transport Canada reported that across Canada, 60 cyclists were killed in traffic accidents with motor vehicles each year during the period from 2004 to 2006 (Transport Canada, 2010). The assessment of road safety can be based on collision data or on other surrogate measures of safety. Traditionally, road collision statistics are considered as the main data source for road traffic safety

analysis (Chin & Quek, 1997; Laureshyn, *et al.*, 2010). Thus, cyclist-involved collisions were investigated in several cyclist safety studies at signalized intersections (Huang, *et al.*, 2012; Jensen, 2008; Wang & Nihan, 2004; Wang, *et al.*, 2008), un-signalized intersection (Summala, *et al.*, 1996; Räsänen, *et al.*, 1999), roundabouts (Sakshaug, *et al.*, 2010; Mette & Tove, 2008) and road-cyclist-path intersections (Hunter, *et al.*, 1999; Wood, *et al.*, 2009). An alternative approach to road safety assessment is based on surrogate measures of safety. The motivation for the use of such surrogate measures of safety in analyzing road safety in general, and more specifically in examining cyclist safety, is the inherent limitations in collision data.

Depending solely on the reported collision data for traffic safety analysis would introduce several possible shortcomings. First, collisions are relatively rare events (Spek, *et al.*, 2006), which require extended observation time to monitor stable trends. For example, before and after safety evaluation is typically based on a few years of collision records in order to detect change in safety. Also, it is relatively difficult to observe, record (by means of a video camera or auditory sensors), and analyze the failure mechanism leading to collisions in comparison to traffic conflicts (Hydén 1987). Second, underreporting of collisions, the level of which depends on collision severity and the type of road users involved, is also a shortcoming of reported collisions data (Laureshyn, *et al.*, 2010). Third, collision data does not typically provide complete information about the collision process. For example, pre-collision road users' behaviour is usually not reported (Laureshyn, *et al.*, 2010).

Surrogate measures of safety are used in order to evaluate the safety of road users, especially when the availability of collision data is limited in quantity or quality. These

measures have been used in practice to assess the safety of current and proposed roadways (Sayed, *et al.*, 1994; Persaud & Mucsi, 1995; Huang & Pant, 1994; Rao & Rengaraju, 1998). Traffic conflict techniques represent one of the key surrogate measures of safety, among other measures (Thompson & PERKINS, 1983). However, due to the subjectivity of field observers, errors were common when manually counting and deciding/judging whether a given traffic event is an interaction. In addition, traffic interaction studies depending on manual counting are time-consuming, labour-intensive and prone to observe inaccuracies. Therefore, to eliminate the subjectivity from traffic interaction analysis, different objective proximity indicators have been suggested as surrogate measures of safety (Songchitruksa & Tarko, 2006; Gettman & Head, 2003; Archer, 2004). Recently, automated data collection using video techniques was introduced in order to improve the accuracy of traffic conflict data collection and to reduce the analysis time (Sayed, *et al.*, 2013; Huang & Jianping, 2004).

2.3.1 Background

A static video camera was used by (Summala, *et al.*, 1996) to observe and record the head movements of motor vehicle drivers while approaching two un-signalized T-intersections in Helsinki. The researchers reviewed the pattern of head movement to find out whether the motor vehicle driver visually searched for the cyclist while the latter is crossing the road. This review was performed in order to find out whether the driver intended to yield to the cyclist. Two main types of conflicts were possible at the monitored intersections. The first type was when a motor vehicle turning right enters or leaves the intersection. The second type was when the motor vehicle turning left enters or leaves the intersection. Each conflict type has four scenarios for a cyclist-vehicle

collision to occur at cyclist crossing as shown in Figure 2-4. Each scenario describes a situation of collision where a cyclist coming from either the left side or the right side of a motor vehicle. Data was collected during non-peak hours at both intersections in the fall season of 1990. The results showed that the motor vehicle driver gives less attention to the right leg of T-intersection while turning right either enters or leaves the intersection.

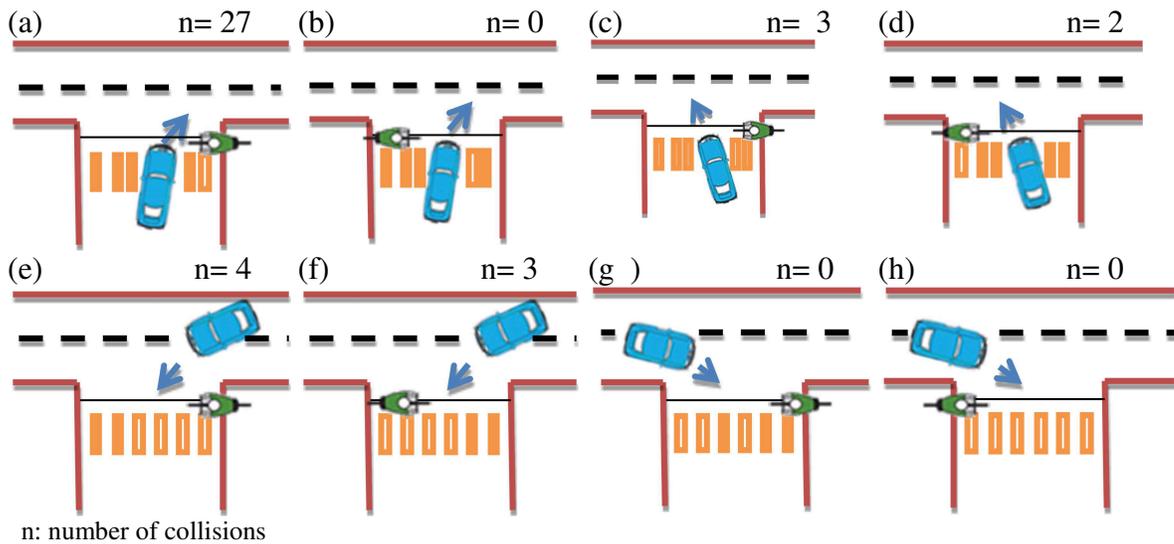


Figure 2-4: Conflict scenarios between cyclist and motor vehicle (reproduced from (Summala, *et al.*, 1996)).

The effect of added two-way cyclist lanes for cyclist safety was studied by (Phillips, *et al.*, 2011). Video camera recording was used to collect data at a T un-signalized intersection in Oslo, Norway. Eight types of cyclist-vehicle conflicts were defined according to the relative movement of the cyclist and the vehicle. The four possible movements of monitored motor vehicles were: left-turn and right-turn movements from side road to main road or vice versa as shown in Figure 2-5. The two possible movements of a cyclist were either up or down on main road.

Conflict observations were conducted during three different periods in order to identify changes in the number of conflicts occurring before, during, and after the new cyclist lane was marked in June 1997. The overall observation time was 57.0 hours, divided as follows: 3.0 hours in 1997, 25.5 hours in 2001, and 28.5 in 2007. Two strategies were adopted when collecting the traffic data. The first strategy was when the cyclist and/or motor vehicle driver yields in order to avoid a collision. While the second strategy was when cyclist and/or motor vehicle driver stops or brakes suddenly in order to avoid a collision. The results showed that a significant reduction in cyclist-motor vehicle conflict was observed especially after 4 years subsequent to the implementation of the new cyclist lane.

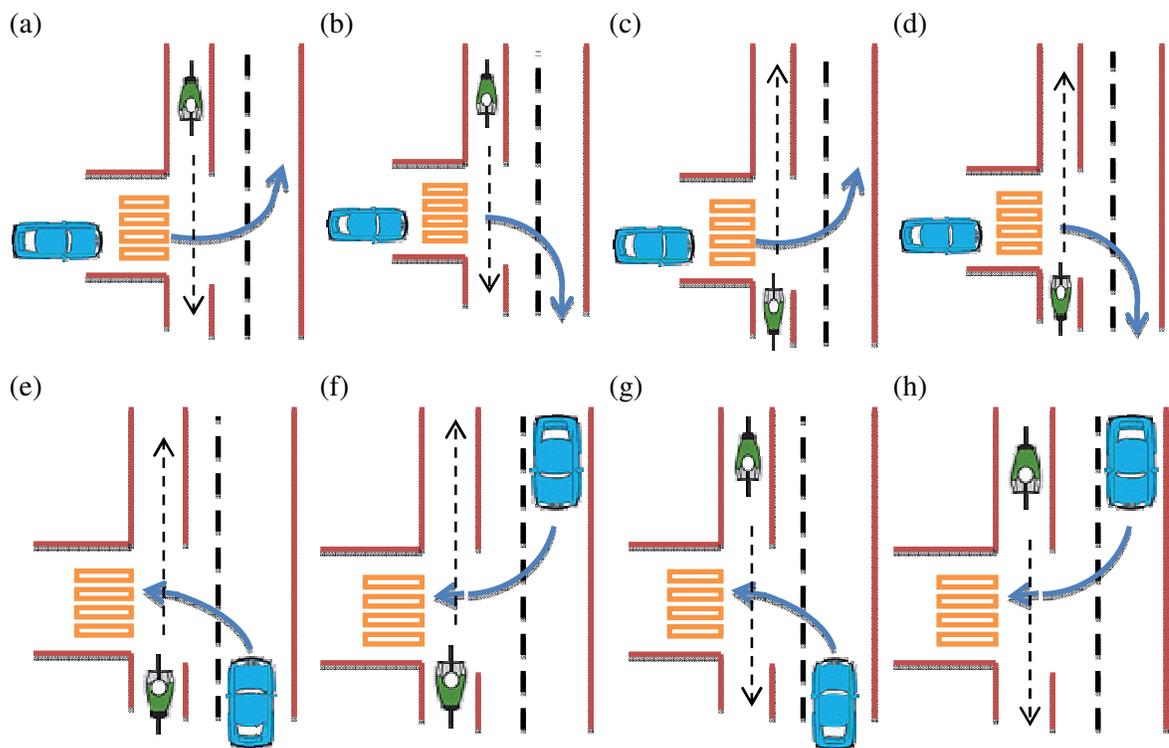


Figure 2-5: Conflict scenarios between cyclist and motor vehicle (reproduced from (Phillips, *et al.*, 2011)).

In another study on the safety impact of cyclist lane, (Räsänen & Summala, 1998), analysed cyclist-vehicle collisions at treated sites in Finland. Fifteen collision scenarios, or chain of events leading to a potential collision, were used to characterize different types of cyclist-vehicle collisions. Each scenario is defined based on classifications of different turning movement of motor vehicles as well as cyclist crossing. Collision data was obtained from the detailed police reports on 188 cyclist-vehicle collisions for the period from 1990 to 1994. In particular, the reports were reviewed to classify collisions and to find out the possible causes of the collision. The results showed that the majority of collisions happened when a motor vehicle driver turning right failed to notice the cyclist coming from the right side of motor vehicle driver along cyclist crossing. The percentage of this majority was 19.6% from overall cyclist-vehicle collisions by scenarios. According to whether the driver or the cyclists were mutually aware, the results were 66% and 51% for cyclist and vehicle driver respectively. It was reported that 17% of cyclist-vehicle collisions happened because both of them did not notice the other before the collisions. The results suggest two main issues that may produce cyclist-vehicle collisions: [i] in some cases vehicle driver may fail to notice a cyclist coming from an unexpected direction and [ii] cyclist mistakenly thought that the vehicle drivers should give a way to the cyclist.

(Wang & Stamatiadis, 2010), studied several contributing factors that may impact cyclist safety at un-signalized intersection in Kentucky. Factors such as characteristics of cyclists, drivers, roadways, and collisions were used as independent variables in a collision severity model. Collision severity is treated as a dependent variable. The model was used to investigate the significance of these variables in explaining cyclist safety.

Collision severity was categorized into four severity levels: slight injured, non-incapacitating, incapacitating, and fatal. A total of 1,242 of collision events were extracted from police reports for the period of 2000 to 2009. The results showed that there was a significant variance among contributing factors on collisions severity. Cyclists older than 25 were found to have increased probabilities of being involved in high severity collisions. Road characteristics such as: icy road, non-traffic control, and one lane approaches were significantly related to the increase in the probability of high severity collision. In addition, through-the-collision characteristics such as rain, fog, cloudy weather, and unlighted darkness were significantly correlated to the increased probability of high severity collision.

2.3.2 Cyclist safety treatments

Safety is considered a fundamental requirement in any transportation system design. It can be argued that the frequency of collisions between cyclists and motor vehicles is proportional to the demand for cycling a mode of travel. With the growing demand for cycling in Canadian cities, existing shortcomings in traffic facilities will become more pronounced. To mitigate the risk of collision, a relatively large number of cyclist safety treatments are available as recommendations in the literature. A number of those treatments have been implemented and evaluated. This section presents reviews of evaluation studies for common cyclist safety treatments.

2.3.2a Cyclist lane treatment

Cyclist lanes are sections of roadway dedicated for cyclist use that can be outlined by pavement marking, premium pavement materials, or physical boundaries. There are many types of cyclist lanes that can be used to provide the cyclist a safe carriageway

either crossing the road or driving beside other road users. Cyclists lanes can be marked with dashed or continuous white line or cyclist lane symbol with barrier are common types of dedicated cyclist lanes as shown in Figure 2-6. A shared lane is a lane shared by both cyclists and vehicles in areas when the dedicated cyclist lane is not possible, due to insufficient width of road. Most of these treatments have been implemented in the downtown area of Ottawa.



Figure 2-6: Cyclist lane treatment. (a) marked cyclist lane. (b) segregated cyclist lane (photos taken by the thesis author).

Various techniques were used to measure the impact of cyclist lanes on the proximity between cyclist and vehicle traffic. (Parkin & Meyers, 2010), studied the effects of cyclist lanes implemented at three sites in Lancashire, England with different overall road width and different posted speed limits. Video data was collected during peak-hour periods at all three sites by a properly instrumented cyclist. Video cameras were used to measure the lateral spacing (side gaps) between motor vehicles and cyclists.

The results showed that the lateral spacing between the cyclist and vehicles in the absence of a bike lane was greater than when a bike lane was outlined.

Empirical before-and-after evaluation of segregating cyclist lanes and marking cyclist lanes was studied by (Jensen, 2008) in Copenhagen, Denmark. The number of collisions and injuries were used as performance measure to evaluate segregated cyclist lanes and marked cyclist lanes. The period of reviewed data was between 1976 to 2004 over 110 Km of road links and 170 intersections. A total of 10,396 cyclist collisions were found, including 4,432 fatalities and severe injuries. The results showed that collisions and severe injuries decreased by 10% and 4% respectively after segregated cyclist lanes were outlined on road links between intersections. Meanwhile, adding marked cyclist lanes on road links and intersections caused an increase in collisions and injuries. At road links, collisions and injuries increased by 30% and 27% respectively. While at intersections, injuries increased by 14%, and no collisions recorded at intersections after marked cyclist lanes added.

(Kendrick, *et al.*, 2011), studied the impact of cyclist lane characteristics on cyclist exposure to traffic on a three-lane one-way street in Portland; mixed-user traffic volume was noticed in the study area. Cyclist lanes were changed from marked cyclist lane to segregated cyclist lane and the street geometry was changed from three vehicle lanes to two vehicle lanes. Data collected over non-consecutive four days between 2009 and 2010. Different equipment was used including: [i] Monitoring equipment (not explained by the authors, but most likely video monitoring), [ii] traffic counters (tubes), and [iii] two particle counter sensors. The ultrafine particle concentration levels, which can be obtained by sensors, were used as performance measures to evaluate segregating cyclist

lanes and marking cyclist lanes. Ultrafine particle number concentrations were used as an indicator for lateral distance between cyclist and vehicle. Ultrafine particle number concentrations were found to be lower in the segregating cyclist lane than the marking cyclist lane. Therefore, the results showed that the segregating cyclist lane provides more protection of exposure to traffic than the marking cyclist lane due to lowering exposure concentrations of ultrafine particles.

(Hunter, *et al.*, 1999), studied the impact of coloured cyclist lane treatment on cyclist safety in Portland, Oregon. Their Data was collected using video cameras and a field survey at 10 conflict areas with high potential of cyclist and motorist interaction. Traffic data was collected over the period from 1997 to 1999. Video recordings were used to investigate whether signage and coloured cyclist lane changed the behaviour of the vehicle drivers and cyclists in order to reduce the conflicts between the two modes. Depending on manoeuvres made by cyclists and vehicles, the sites categorized into three groups of conflict areas: [i] exiting roadway and ramp, [ii] auxiliary right turn lane, and [iii] intersecting road way or ramp. Traffic data was collected during two peak periods on each site before and after treatment constructed. The purpose of field survey was to collect the opinions on the treatment plan from the cyclists and vehicle drivers who are using the sites. The videotape analysis showed that the higher numbers of vehicle motorist yielded to cyclists and slowed or stopped before entering the coloured cyclist lane. In addition, the majority of cyclists followed the coloured cyclist lane and fewer cyclists turning their heads to scan for traffic or using hand signals. The survey results showed that the 76% from both cyclists and vehicle riders felt the locations with the coloured pavement were safer.

2.3.2b Advanced stop and bike-box treatments

Another specific context for the spatial separation between cyclists and motor vehicles is called Advanced Stop Lines (ASLs). An ASL is a storage space, coloured or uncoloured, located in front of motor vehicles at intersection to allow cyclists to stop and resume movement free from conflict with motor vehicles (Weigan, 2008). A schematic of ASL is shown in Figure 2-7, in several cities in the U.S. and Europe, bike boxes are popular configurations of ASL (Hunter, 2000; Loskorn, *et al.*, 2013) as shown in Figure 2-8. The purpose of the bike-box design is to reduce the possibility of cyclist collisions, especially right-hook collision. A key advantage of bike-boxes is to enhance the visibility of cyclists when stopped in front of motor vehicles at signalized intersections. This separation potentially allows safer cyclist movement at intersections and lower risk of conflict between cyclists and motor vehicles. Bike-boxes also provide cyclists with a head start when the light turns green.

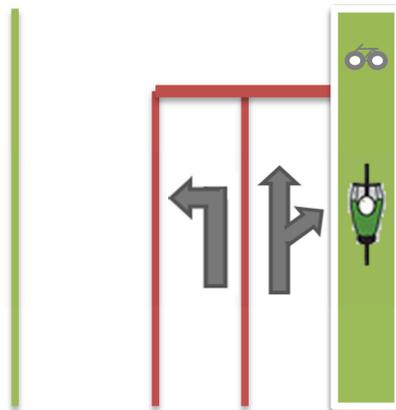


Figure 2-7: Schematic diagram of advanced stop line (ASL) (reproduced from (Koorey & Mangundu, 2010)).

Behaviour of cyclists at advanced stop lines was studied by (Allen, *et al.*, 2005), at twelve sites in London, England. The results revealed that there were good separation margins provided by the advanced stop line while the cyclists stopped in front of motor vehicles at intersections.

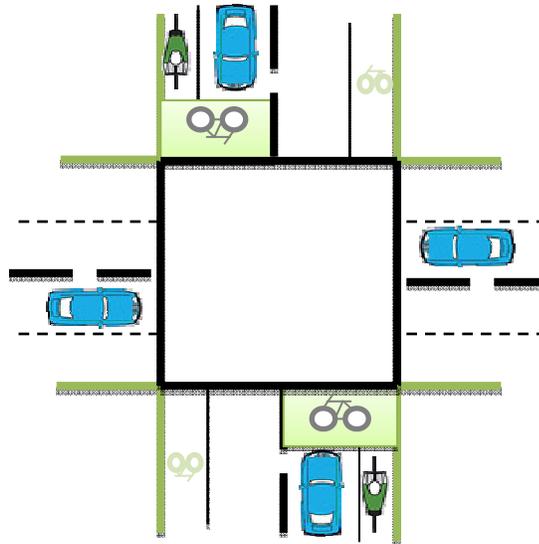


Figure 2-8: Schematic diagram of a bike box at signalized intersection (reproduced from (Weigand, 2008)).

(Koorey & Mangundu, 2010), studied the impact of colour cyclist lane on vehicle driver behaviour at 18 different signalized intersections in Christchurch, New Zealand. In this study, signalized intersections featured were Advanced Stop Lines (ASLs) and bike boxes. In addition, the study examined the effect of varying traffic lane width on the behaviour of drivers. Manual surveys and video recording were used to observe the positions of motor vehicles in relation to both coloured and uncoloured ASLs and bike boxes. Nine sites were selected as ASLs with varying traffic lane widths; six uncoloured and three coloured surfacing. Also nine sites were selected as bike boxes; six uncoloured

and three coloured surfacing. One person was selected to observe and record data manually over three months from 2008 to 2009. Data was collected during peak hours from 7-9 am or 4-6 pm/per site. Three categories were used in order to define the position of vehicle related to cyclist facilities. These categories include: [i] non-encroachment inside cyclist facility, [ii] partial encroachment inside cyclist facility, and [iii] complete encroachment inside cyclist facility. In order to test the three categories, statistical analysis was used and a comparison was made between the effects of coloured and uncoloured surfacing on driver behaviour. The results showed that drivers were less likely to encroach on coloured marked areas in comparison to uncoloured ones and the majority properly used the marked facilities. The results showed that the driver compliance improved at coloured sites and the rate of encroachment decreased significantly; especially by turning vehicles. In addition, there was a statistically significant difference between motor vehicle encroachments at narrow and wide lanes. This study reported that narrow traffic lanes (≤ 5.0 m), vehicle encroachment for both coloured and uncoloured lanes were found to be 10.8% and 19.4%, respectively. At wide traffic lanes (> 5.0 m), vehicle encroachments for both coloured and uncoloured were found 26.6% and 33.8% respectively.

(Dill, *et al.*, 2012), studied the effectiveness of bike boxes at signalized intersections at ten signalized intersections in Portland, Oregon. Video data was collected over two time periods: [i] January through to March in 2008 before cyclist facility installed, and [ii] through April to June in 2009 after cyclist facility installed. To check the visibility of cyclist at bike boxes, two lane colour markings were used. Seven bike boxes were painted with green colour and three were uncoloured. The objective of the

video observation was to investigate the compliance rate of motor vehicles and cyclists in response to the new bike boxes. The results showed that 73% of the stopping vehicles did not encroach on the bike- box. In addition, the number of conflicts at the bike-box places decreased and cyclists yielding behaviour increased.

The effect of bike box treatments on cyclist safety was studied by (Loskorn, *et al.*, 2013), at two four-leg signalized intersections in Austin, Texas. The cyclist facility was used to segregate cyclists from other road users. Video image processing was used to collect the traffic movement data over the period between 2009 and 2010. In order to measure cyclist- vehicles conflicts, three different stages were used during traffic data collection. The first stage included intersections clear of bike boxes for five months. The second stage spanned three months, during which, uncoloured bike boxes were implemented. The third stage spanned four months, during which, coloured bike boxes were implemented. At the beginning of the second phase, data was collected two weeks after the treatment had been implemented. In order to observe the natural behaviour of cyclist and vehicle drivers, measures were taken to avoid alerting road users of changes to the intersection control. The authors in this study defined safety (conformal) behaviour as follows; [i] the cyclist uses the cyclist lane in order to approach the intersection, [ii] cyclists stops inside the bike box in front of vehicles, [iii] no encroachment by vehicles into either the stop line or bike boxes, [iv] vehicles enter the intersection after cyclist departure, [v] cyclists make a general regularly or legal movement. The results showed that there was a significant effect of bike boxes on the behaviour of cyclists and motor vehicle drivers. Drivers were less likely to encroach on uncoloured bike boxes compared to approaching stop lines without boxes. The rate of encroachment further decreased

significantly after colouring. In the same study, another intersection the motor vehicle drivers encroachment increased by 34% after the bike boxes were marked, but decreased by 10% after they were coloured. In addition, the study reported that "*no conflicts or collisions between motorists and bicyclists were observed during any phase of the study*". In general, this conclusion invites criticism into the operational definition of a conflict adopted by the authors in this study. This definition was not explicitly described in this study.

2.3.2c Other treatments

Study by (Summala, *et al.*, 1996) implemented two types of a physical treatment plan in order to reduce the number of conflicts between cyclists and vehicles at six intersections in Helsinki, Finland. Two cameras were used and data was collected before and after treatments were implemented. First treatment includes two control warnings: [i] painted triangular warning signs on the ground before the intersection, and [ii] cyclist sign posted at the approaches to the intersection. The second treatment includes three control warnings applied on different intersections: [i] hump cyclist path, [ii] elevated cyclist crossing path, and [iii] stop sign posted before cyclist path. These control measures were used to attract attention of vehicle drivers by stimulating visual scanning routines at intersections. Head movements of motor vehicle driver were studied before and after the treatment plan. The results showed that there was an improvement after the treatment plans implemented. The number of drivers who stopped and made attempts to scan the cyclist crossing increased after the treatment was implemented. In addition, the improvement was noticeable when using hump and painting compared to other treatments.

(Jensen, 2008), studied the effect of marked 1 to 4 (a maximum total of four crossings) blue cycle crossings on cyclist safety at 65 signalized intersections in Copenhagen, Denmark. Data of cyclist collisions and injuries for the period from 1976 to 2004 were studied before and after treatments were implemented at intersections. Collision prediction models were developed. The study showed that cyclist safety is affected by the following: [i] number of blue cycle crossings, [ii] number of intersection legs, [iii] intersection size, and [iv] traffic volume. Small intersections were found to have the best safety effect using blue cycle crossing compared to the larger intersection area. It was argued that this is due to having too much focus by motorists on blue cyclist crossing rather than focusing on other traffic control. The results showed that one blue cycle crossing significantly reduced the number of intersection collisions and injuries by 10% and 19% respectively. There were no improvements when blue cycle crossings were used.

Another type of treatment that provided motivation for the study of cyclist crossing speed was the use of dedicated signal phase for cyclists. (Korve & Niemeier, 2002), studied the effect of adding a cyclist phase to the existing three phases. In this study, the stated purpose of adding signal phase was to reduce cyclist-vehicle conflicts at three-legs T signalized urban intersection in Davis, California. The total cycle length was increased by 21 second and the signal operation was changed from semi-actuated to fully-actuated. Detection of cyclist was performed using cyclist loop and cyclist push button. High traffic volume of both cyclists and vehicles were noticed at the monitored signalized intersection. Traffic volume data was collected and collision data was categorized in into two scenarios depending on: [i] the type of collision, and [ii] and the

parties involved. Ten collisions were recorded between cyclists and vehicles over a 35 month period from 1992 to 1994. Benefit-cost analysis was used to evaluate the modified signal phases. The purpose of the analysis was to compute how much improvement could be achieved in terms of: [i] reduction in collisions between cyclist and vehicle to improve the cyclist safety, [ii] changes in travel time delays, [iii] physical cost due to a new construction, and [iv] vehicle emissions changes. The annual cost of collisions was estimated with and without cyclist phase depending on the type of highway and geographic location. The results found there was a reduction in cost when the cyclist's phase added, taking into account the increase of cyclist volume in the future. In addition the number of conflicts between cyclist and vehicles were reduced.

2.3.3 Conflict indicators

Due to the frequent traffic conflict manoeuvres occurring at intersections, it is considered to be among the highest risk locations on the road network (Sobhani, *et al.*, 2012; Wang & Nihan, 2004). Traffic conflicts between cyclist flows and motor vehicle flows are more frequent and more observable than the collision events (Wang & Nihan, 2004). A conflict can be defined as “*an observable situation in which two or more road users approach each other in space and in time to such an extent that there is a risk of collision if their movements remained unchanged*” (Amundsen & Hydén, 1977).

There are several traffic conflict indicators that were used as alternative approaches to safety analysis (Archer, 2004; Gettman & Head, 2003; Kiefer, *et al.*, 2005; Vogel, 2003; Zheng, *et al.*, 2014). Several indicators were used to quantify the degree of proximity of road users and also in some situations to identify traffic conflicts such as Time-to-Collision (TTC) and Post-Encroachment Time (PET).

TTC is defined as “the time that remains until a collision between two vehicles would have occurred if the collisions course and speed difference are maintained” (Hayward, 1972). In the context of cyclist-vehicle conflicts, TTC represents the time that elapses from current moment till both the cyclist and the vehicle reach an extent of proximity that the former’s safety is compromise if their movements remain unchanged. Post-Encroachment Time (PET) can be defined as the time difference between the moment at which the first road user leaves an area of potential collision and the moment at which the second road user arrives at this area while possessing the right of way (Allen, *et al.*, 1978). An interaction zone can be defined as an area within which the positions of the cyclist and the interacting vehicle are deemed perilously close. Figure 2-9, shows a schematic diagram for both TTC and PET.

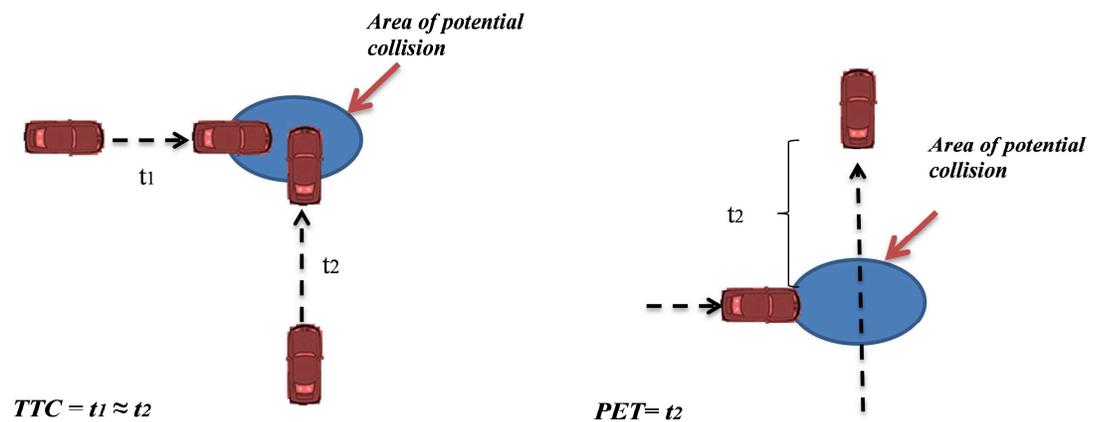


Figure 2-9: Schematic definition of TTC and PET.

2.3.4 Post-Encroachment Time

PET has been used in previous studies to identify traffic conflicts (Cooper, 1984; Hunter & Rodgers, 2012; Peesapati, *et al.*, 2013; Van der Horst, *et al.*, 2014). In order to count the number of conflicts, a threshold needs to be selected in order to count observations with PET values below this threshold. Furthermore, PET was successfully used to point for limitations in collision record when the magnitudes of this surrogate measure and collision frequency were not consistent (Peesapati, *et al.*, 2013).

The study of cyclist safety using PET is relatively limited. (Huang & Jianping, 2004), measured the proximity between motor vehicles and cyclists at signalized intersections in Beijing, China. Three scenarios of conflict movements were studied between cyclists and motor vehicles. The first scenario is when the motor vehicle is turning right and the conflicted cyclist is crossing the intersection, through/or left. The second scenario is when the motor vehicle is travelling in the through direction while the cyclist is turning left. The third scenario is when the motor vehicle is turning left while the opposite cyclist is travelling in the through direction as shown in Figure 2-10. In order to identify a conflict, two time durations were manually measured from the image sequences of the traffic events involving a cyclist and a vehicle. The term lag that the author used to measure the proximity in the study has the same description of the PET. The camera frame rate was 25 frames/second, and therefore the precision of the estimated lag was 0.04 second. The results showed that the average lag was 2.93 second and the minimum was 0.52 second. However, the study didn't use any automated PET measurement.

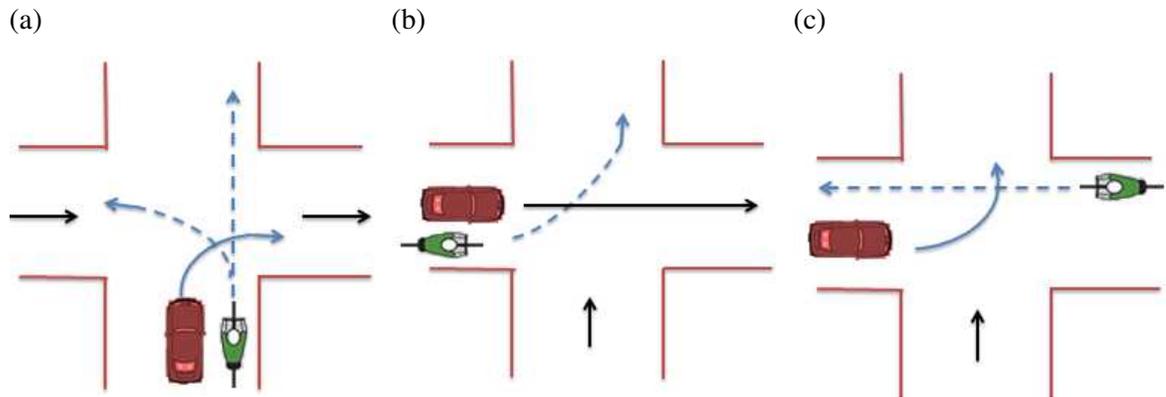


Figure 2-10: Scenarios of conflict movements between cyclists and motor vehicles at signalized intersection. (a) Right-turning motor vehicles versus cyclists. (b) Through motor vehicles versus left-turning cyclists. (c) Left-turning motor vehicle versus through cyclists (reproduced from (Huang & Jianping, 2004)).

2.4 Summary

This chapter presented a comprehensive review of the studies and techniques to measure cyclist speed. As well as, this chapter presents the review related to the topic of evaluating cyclist safety treatments.

The cyclists in the majority of the reviewed studies may pre-select to participate in advance. The sampling strategy of selecting cyclists may not be representative of the general population. The observed cyclist might have been aware of being monitored. The cyclists who were provided with any type of tracking equipment; for example, GPS and speedometer; for the purpose of collecting speed data, it can be argued that the data collection strategy may not yield naturalistic behaviour. To get the naturalistic cyclist behaviour it is preferred to be interfering with other road users and without any prior notice. In addition, it is not easy to measure the cyclist speed as a motor vehicle using a

specific technique; this is due to the technique is not designed for cyclist speed measurements. Judgments have been made regarding device limitations and usage in techniques. A summary of the comparison of different techniques for cyclist speed measurement is shown in Table 2-1.

Table 2-1: Comparison of different techniques used for cyclist speed measurements.

Technique Speed condition	Radar gun	Stop- watch	Speedo- meter	Pneumatic tube sensor	GPS	Video camera imaging	Computer vision
Non-intrusive	◐	●	○	○	○	●	●
All movement directions	○	◐	●	○	●	●	●
Environmental independency	◐	◐	◐	◐	◐	●	●
High accuracy	◐	○	○	◐	◐	●	●
Cost efficient	○	◐	◐	○	○	●	●
Coverage area	○	○	○	○	○	●	●
Global observation	○	○	○	○	○	●	●
Large sample size	○	○	○	●	○	●	●
Quality check	○	○	○	○	○	●	●
High traffic volume	○	○	●	○	●	●	●
Human input	○	○	○	●	○	◐	●

●technique meets the conditions, ◐technique partially meet the conditions and ○technique doesn't meet the conditions.

As shown in Table 2-1, the automated cyclist speed measurements using computer vision technique might overcome the devices' limitation that are listed in Section 2.2.1.

It is important to note that using the computer vision analysis technique with video recording will help in providing opportunities to closely study the data in later stages. In addition to that, video analysis permits continuous tracking of cyclists to calculate a speed profile over a finite period of time or over a specific path.

Different treatments have been implemented and evaluated in the reviewed studies in order to reduce the number of interactions between cyclists and motor-vehicles and to improve cyclist safety such as; marking lanes, segregated cyclist lanes, advanced Stop Lines, bike-boxes, colour cycle crossings, signal phase for cyclists and different control warnings of cyclist's signs.

One of the results, in the studies reviewed showed that the segregating cyclist lane provides more protection of exposure to traffic than the marking cyclist lane. In addition, the results showed that collisions and severe injuries decreased by 10% and 4% respectively after segregated cyclist lanes were outlined on road links between intersections. Results of another reviewed study revealed that there were good separation margins provided by the advanced stop line while the cyclists stopped in front of motor vehicles at intersections.

The treatment effects of adding bike-boxes in reviewed studies showed that the number of conflicts at the bike-box places decreased and cyclists yielding behaviour increased. In addition, the rate of encroachment further decreased significantly after bike-boxes were coloured. The result of the impact of colour cyclist lane at the intersection was clear through the reviewed studies. The results of the reviewed studies showed that drivers were less likely to encroach on coloured marked areas in comparison to

uncoloured ones and the rate of vehicles encroachment decreased significantly; especially by turning vehicles.

Another reviewed study found that factors such as; number of blue cycle crossings, number of intersection legs, intersection size, and traffic volume were an influence on cyclist safety. The reviewed study results showed that one blue cycle crossing significantly reduced the number of intersection collisions and injuries by 10% and 19% respectively. In addition, small intersections were found to have the best safety effect using blue cycle crossing compared to the larger intersection area. It was argued that this is due to having too much focus by motorists on blue cyclist crossing rather than focusing on other traffic control. Another study was found that the number of conflicts between cyclist and vehicles were reduced when adding a dedicated cyclist phase.

Different cyclist-motor collisions scenarios were investigated in the reviewed studies. The results of reviewed studies showed that the majority of collisions happened when a motor vehicle driver turning right failed to notice the cyclist coming from the right side of motor vehicle driver along cyclist crossing

Different surrogate measures of safety can be used to evaluate the safety of road users. The assessment of road safety based on collision data would introduce several possible shortcomings, such as; [i] collisions are relatively rare events, [ii] underreporting of collisions, [iii] collision data does not typically provide complete information about the collision process. Traffic conflict is considered as one of the key surrogate measures of safety, among other measures. Recently, automated data collection using video techniques were introduced in order to improve the accuracy of traffic conflict data collection and to reduce the analysis time. Traffic interaction studies depending on

manual counting are time-consuming, labour-intensive and prone to observe inaccuracies. In order to eliminate the subjectivity from traffic interaction analysis, different objective proximity indicators have been suggested as surrogate measures of safety.

3. AUTOMATED MEASURING OF CYCLIST SPEED

Measuring road user speed is an important task in traffic data collection. With growing interest in sustainable modes of travel, many traffic control measures are developed to serve cyclists. This study investigates the use of computer vision techniques for measuring cyclist speed. The developed technique enables the automated observation of large volume of naturalistic cyclist movements in an accurate and resource-efficient manner. The use of computer vision techniques for measuring cyclist speed is relatively recent. The key advantages of video monitoring and computer vision techniques are:

1. Video monitoring is less susceptible to the impact of tall physical objects, (*e.g.*, high-rise buildings that abound in urban environments as compared to traditional GPS tracking techniques). Note that sophisticated GPS technologies which provide high-accuracy in urban canyons, such as real-time kinematic (RTK GPS) (Lee, *et al.*, 2011), require much more complicated set-up preparations than the computer vision technique presented in this study.
2. Monitoring movements of road users are non-intrusive and will likely result in analyzing naturalistic cyclist movements. Observed subjects are typically unaware of video monitoring. This advantage has made video recording a commonly used technique to observe the behaviour of road users.
3. The variability of cyclist speed with time can be measured along straight as well as curved paths.

3.1 Data Collection

3.1.1 Study site and equipment

For the purpose of measuring cyclist speed, a static high-definition commercial-grade video camera (1920 × 1080 pixels at approximately 30 frames/second) was used. In order to have a full-view of the intersection, the camera was placed at the 12th floor of a building overlooking the four-leg signalized intersection of Stewart and Waller streets near the Mackenzie King Bridge, located in the Downtown of Ottawa. Figure 3-1 shows the location of this intersection. Data collection was non-intrusive because the camera was housed inside a building behind a glass window overlooking the observed intersection. Figure 3-2 shows the camera level position and the camera setup at the signalized intersection.

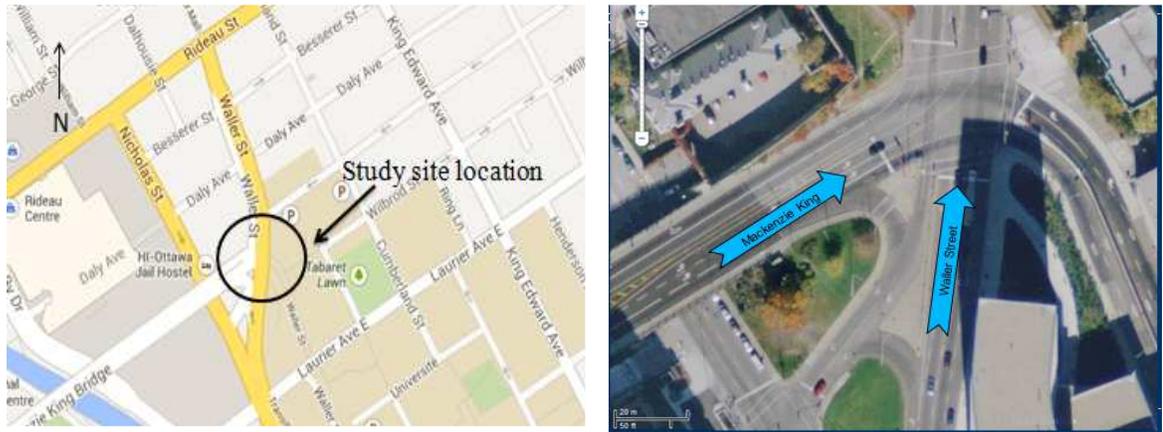


Figure 3-1: Site locations: signalized intersection of Stewart and Waller streets

(Google maps, 2011).

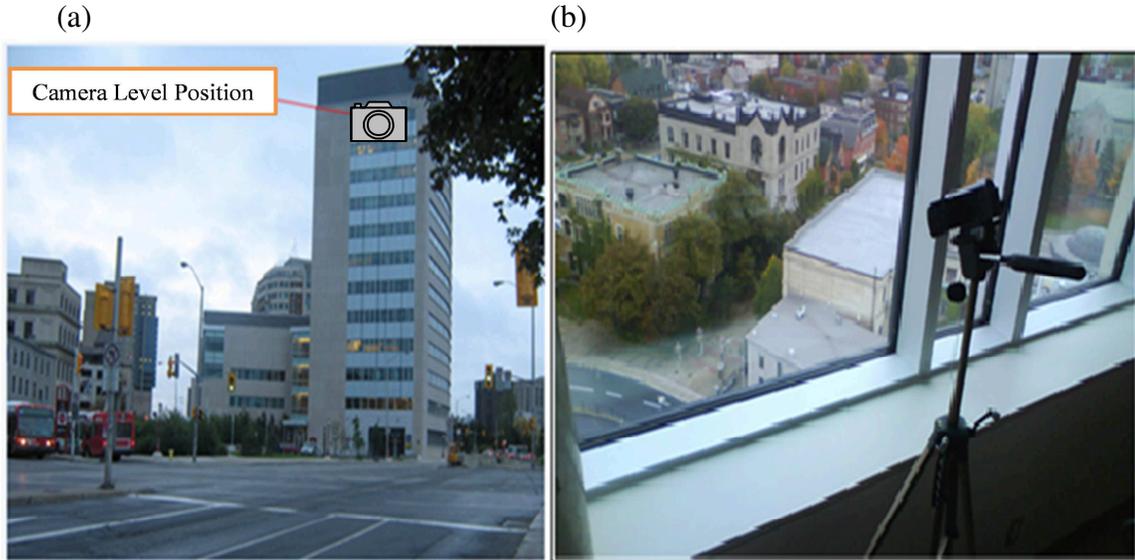


Figure 3-2: Camera position at the site location. (a) building and level at which the camera was housed. (b) video camera set-up.

3.1.2 Video data collection

Data was collected in 13 days in two different time periods to cover two different seasons: fall and summer. A total of 70 hours (6 days) of video data was collected in the months of October and November in 2010 and a total of 120 hours (7 days) in June 2011. Raw footage was subject to special processing in stages. First, the video format was changed from a proprietary format (MTS or MP4) to an open one (AVI) using a developed special script. In the second stage, video data was reviewed manually and accordingly video segments containing cyclist crossing movements were identified (video times for beginning and end). The crossing action was defined as the cyclist movement which extends in between the marked intersection lines (from stop line at the approach and to the pedestrian line). All information related to this stage including video path, name, format, start of the video subsequence containing the cyclist, and lengths of

those subsequences were recorded in an excel sheet as shown in Appendix (A). The third stage included dividing the main video into subsequences in order to reduce processing time on the basis of cyclist time spent at the intersection. For the purpose, a special script was developed and used the information from the excel sheet that was prepared in the second stage. The duration of these video subsequences ranged from 4.0 to 16.0 seconds with an average of 9.0 seconds. Cyclists crossing through the pedestrian crosswalk or using the sidewalk were excluded from the data. The total number of cyclist crossing events observed was 716.

3.1.3 Surveying works and camera calibration

In order to conduct metric measurements, the recording camera is calibrated by knowledge of geometric primitives that appear in the video and their corresponding appearance in an orthographic image of the intersection. Also, field measurements were made to establish the true lengths of segments that appear in the video in order to aid the calibration process.

3.1.3a Surveying works

A measuring wheel and measuring tape were used for the purpose of obtaining the true lengths of segments which appear in the video in order to aid the calibration process. The field measurements were taken from some of the distances at the signalized intersection that were also visible in the image plane, *e.g.*, stop line, straight line on curb side, distance between two manholes, street furniture. Figure 3-3 illustrates linear segments that were taken as required measurements to determine the camera calibration.

The accuracy of the measuring wheel was tested using a measuring tape. The measuring tape was placed on a flat level surface. The difference between a measured

distance (20 meter) by a measuring tape on the flat level surface and the measuring wheel was found to be 0.019 meter. This is an acceptable accuracy given that it is less than the precision of the measuring wheel (0.025 meter).

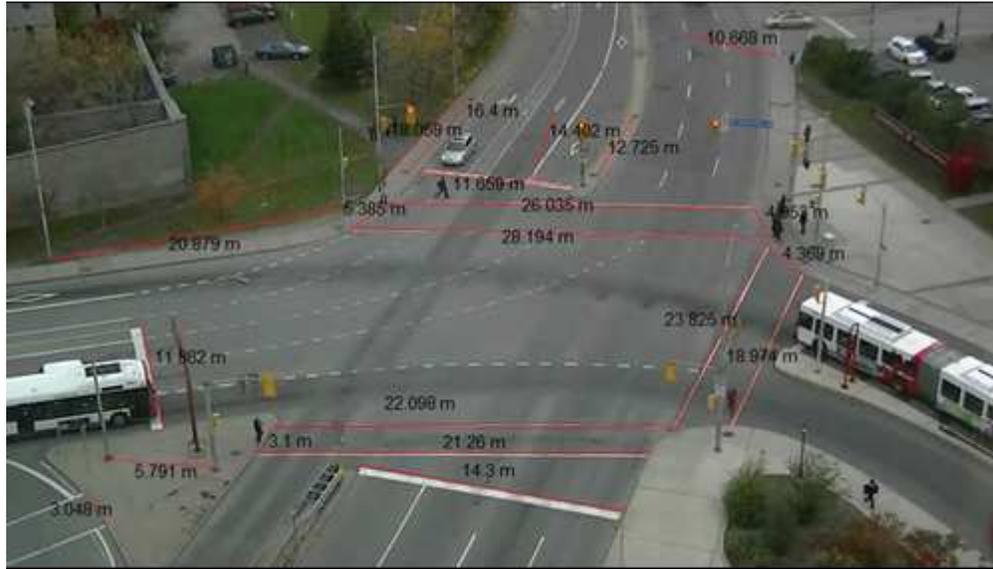


Figure 3-3: Field measurement distances at the signalized intersection.

3.1.3b Camera calibration

Video data show the position of individual cyclists relative to the road and to other users in the plane of the camera sensor. These metrics are affected by the position of the camera itself. Distances and angles determined from a video frame are *distorted* values from the *real* dimensions, due to parallax, perspective, and other geometric factors.

The main purpose of camera calibration is to estimate a set of camera parameters to project objects onto the image plane (road surface). Camera parameters include camera position and orientation, image format, focal length, skew angle and radial lens distortion. In order to transfer the pixel coordinates to world image coordinates, two snapshots should be prepared. The first snap shot is from the video itself as one frame, and the second one is the world image for the same location; in this thesis the location of interest

is the signalized intersection under study. This signalized intersection is located in an open area as shown in Figure 3-3 and non-high buildings surround the intersection, therefore, it is possible to obtain the world image for the location of the signalized intersection from the Google Maps.

The field measurements, parallel lines in real world image, camera position, camera height, and a frame taken from the video were considered as an input data in a special calibration script discussed by (Ismail, *et al.*, 2013), to produce a homography matrix. The homography matrix is a mathematical transformation matrix that effectively rotates and translates *raw* (video) measurements to *real world* values. This transformation is needed to track features in world coordinates (Coifman, *et al.*, 1998). Figure 3-4(a) shows a reference grid drawn in world-coordinates. Figure 3-4(b) shows the projection of the same reference grid from world space in (a) to image space using the calibrated parameters. The accuracy of the estimated parameters was tested using between 19- 23 line segments, whose true lengths were estimated from the orthographic image. Table 3-1 shows the absolute value of the difference between actual field measurements using a measuring wheel and distances estimated using calibrated camera parameters for a sample scene. This difference is interpreted as error in camera calibration. Table 3-2 shows the average percentage errors in linear measurement at different scenes (calculated by dividing the summation of all errors in each calibration linear segment by the total length of actual field distance measurements). The minimum and maximum values of the average percentage errors were found to be 1.03% and 1.90%.

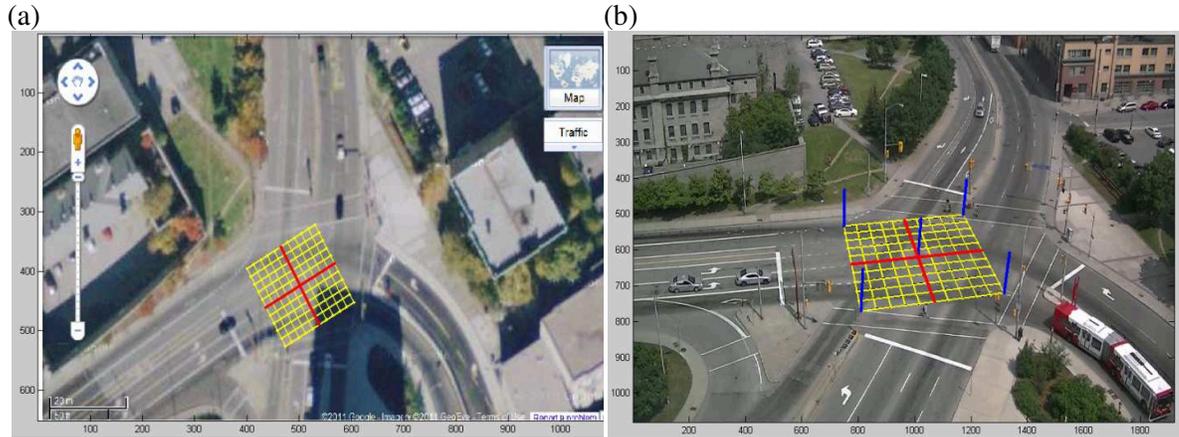


Figure 3-4: Visualization of camera calibration. (a) orthographic grid. (b) same grid projected on a sample video frame.

3.1.4 Selection of tracking parameters

The video analysis system which was used in this study is based on an open-source feature-based vehicle tracking system (Saunier & Sayed, 2006; Lucas & Kanade, 1981). A significantly improved version of an existing open-source platform (Google, 2011) was developed and used to conduct feature tracking and grouping. This tracking system was used to measure the crossing cyclist speed at signalized intersections (Sayed, *et al.*, 2013). A common tracking algorithm, Kanade-Lucas-Tomasi (KLT) feature tracker, was used to detect (locate first position) and track (locate subsequent positions in next frames) points or *features* as they move from frame to frame over time (Lucas & Kanade, 1981). A feature is defined as a salient point, or sub-window of the image frame, which can be tracked for an adequate number of frames while in motion. Not all tracked features are considered when identifying moving objects in the field of view. Only activated features are grouped together based on user-specified parameters which control the function of the tracker.

Table 3-1: Absolute value of the difference between actual field distances using measuring wheel and estimated distances using calibrated camera parameters.

Actual field distance measurements (m)	Homography distance measurements (m)	Absolute difference (m)	% Error in distances
14.30	14.02	0.28	1.96*
21.26	21.00	0.26	1.22
22.10	21.84	0.26	1.18
3.10	2.94	0.16	5.16
5.79	6.09	0.30	5.18
3.05	3.21	0.16	5.25
11.86	11.92	0.06	0.51
20.88	20.87	0.01	0.05
5.38	5.01	0.37	6.88
16.40	16.18	0.22	1.34
11.66	12.06	0.40	3.43
12.73	12.74	0.01	0.08
26.04	26.34	0.30	1.15
28.19	28.45	0.26	0.92
10.67	10.87	0.20	1.87
4.95	4.78	0.17	3.43
4.37	4.20	0.17	3.89
7.54	7.24	0.30	3.98
13.66	13.37	0.29	2.12
8.20	8.18	0.02	0.24
18.97	18.86	0.11	0.58
22.30	23.56	1.26	5.65

* (% Error is calculated by dividing the absolute difference between actual distance (shown in Figure 3-3) and those estimated using the camera calibration output (homography matrix) by the actual length. For example, $((0.28/14.3)*100=1.96)$.

These features are considered for grouping (or activated) if they survive and continue to move for long enough duration. Activated features are connected together if they meet the conditions for proximity and similarity in movement. This connection is

represented as an edge in a graph of connected features. An isolated graph of connected features forms an object. An object is defined as an apparently coherent moving road user in the video.

Table 3-2: Average overall errors between linear field measurement (shown in Figure 3-3) and estimated distances using camera calibration output in different days and seasons.

Scene	Season	Average overall errors (%)
Oct-19	FALL 2010	1.90
Oct-21		1.52
Oct-27		1.03
Oct-28		1.49
Nov-01		1.38
Nov-02		1.32
Jun-08	SUMMER 2011	1.90
Jun-09		1.90
Jun-13		1.90
Jun-14		1.75
Jun-15		1.90
Jun-17		1.70
Jun-24		1.73

The video tracking system is controlled by a configuration file which contains different tracking parameters which affect how the processes of feature detection, tracking, and grouping are conducted. The tracker system configuration file is a separate text file that the user supplies to the tracker in order to configure or assign values for specific parameters or conditions that the tracker system will depend on. A sample of the tracker system configuration file is displayed in Appendix (A). Multiple trials were conducted to change the value of these parameters in order to achieve adequate cyclist tracks. These parameters were found to depend on: field conditions, weather conditions,

video illumination, traffic density, and the clarity of visual identification of the cyclist. The parameters whose values were found more effective in achieving an adequate tracking in different conditions were: [i] minimum distance moved required, [ii] maximum feature distance to be activated, [iii] minimum distance between tracks, [iv] number of corners to track, and [v] tracker quality level. The descriptions of the parameters are shown in Table 3-3. Different configuration files were developed based on the different parameters. Table 3-4 shows a summary of the different tracking parameter values that can be used in different conditions based on the experience gained in this thesis. The terminology used to describe these parameters is adopted from (Saunier & Sayed, 2006). The adequacy of tracking is assessed according to the procedure in the upcoming subsection. Figure 3-5 shows a flow chart describing the steps of video tracker system and Figure 3-6 illustrates the video processing loop. Figure 3-7 shows frames of a video segment describing the temporal progress of detecting, grouping and tracking features in the image plan. Different sample frames are shown in Appendix (A).

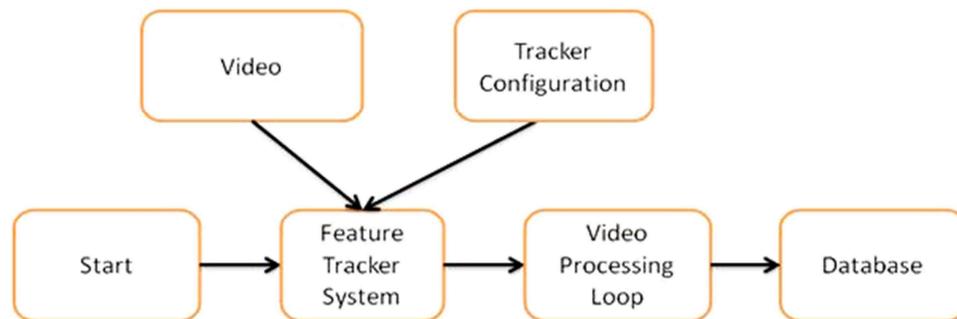


Figure 3-5: Flow chart diagram of cyclist video tracking system.

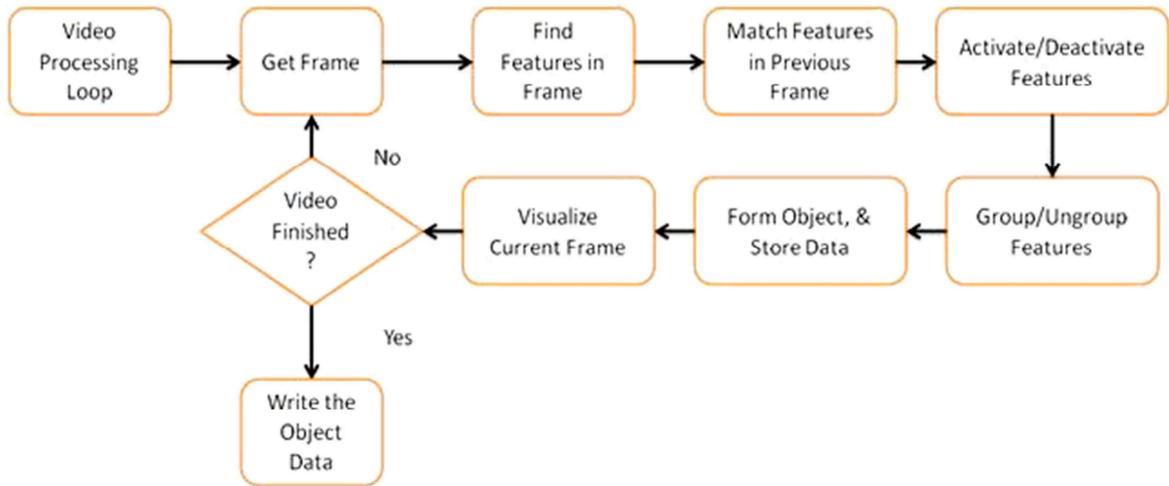


Figure 3-6: Video processing loop.

Table 3-3: Description of the configuration file tracking parameters *

Parameter Name**	Default Value	Description
min_distance_moved_required (m)	0.02	A feature has to move in minimum frames tracked before it is considered to be activated. If the distance moved by the feature is greater than or equal to the min_distance_moved_required, then the feature is activated, otherwise is ignored.
maximum_distance_activated (m)	1.5	The distance between two features before an edge is formed between them. If the distance between the two features exceeds this threshold then the edge is severed or deleted.
min_distance_between_tracks (m)	0.2	How far two tracks can be from each other before they are segmented. If the two tracks closer than this distance, is considered duplicate and one will be removed.
tracker_max_corners	1750	It is defined as the number of the healthiest and strongest features to return from all features detected in a frame. This parameter is part of the primary tracker mechanism, <i>good features to track</i> .
tracker_quality_level	0.05	Tracker_quality_level defines the quality level restriction on each feature. A feature quality is a value computed by the tracker mechanism, <i>good features to track</i> , to evaluate the strength of a feature.

* (OpenCV, 2011)

** names literally appear in the tracker system C++ code

Table 3-4: Summary of the tracking parameters.

Tracker parameter	Time the day		Weather condition		Traffic density		Cyclist identification	
	day-time	night-time	day	rain	low	high	easy	Hard*
min_distance_moved_required	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
maximum_distance_activated	1.50	2.00	1.50	2.00	1.00	0.90	1.00	2.00
min_distance_between_tracks	0.50	0.01	0.50	0.30	0.50	0.20	0.50	0.01
tracker_max_corners	900	1500	900	1750	900	750	900	1750
tracker_quality_level	0.07	0.04	0.07	0.01	0.07	0.05	0.07	0.02

*due to dark-clothed cyclist and shade area



Figure 3-7: Sample frames showing features identified by the tracker. (a) before the tracker system applied. (b) after the tracker system applied.

3.2 Speed Measurements

In this study two different techniques were used to measure cyclist crossing speed. The first technique is an automated computer vision technique and the set of observation is called video-measured speed. The second technique is manual measurement technique and the set of observations is called field-measured speed. The field-measured speed data was collected by tracing the actual path of the front wheel of the bike while crossing the

intersection. The field-measured speed was manually calculated based on the time required by moving the cyclist to traverse marked intersection lines. Tracking the actual path of the bike was achieved using two different approaches which will be described in the next sections.

3.2.1 Cyclist water trail (CWT)

The *first approach* is to trace a test bike's path by physically marking the path of the bike using water. This was achieved by tracking the water trail generated by water-wetted bike front tire on the pavement during a crossing maneuver. A revolving measurement wheel was used to measure the length of the cyclist path while immediately following the cyclist after the wet trail is created. The total number of the cyclist path measured in this approach was 16 paths. This process is illustrated in Figure 3-8. Figure 3-8(a) shows a video sample frame, where the left circle shows the person holding the measuring wheel (thesis author). The right circle shows the test bike while conducting a crossing maneuver. Figure 3-8(b) shows the test bike. All crossing manoeuvres were recorded on video and were conducted at Steward and Waller signalized intersection where all video data was collected. In order to compare the field-measured and video-measured speeds, video data was reviewed manually and the total number of the frames of each crossing manoeuvres was counted between the marked intersection (from stop line at the approach to pedestrian line). The crossing time of the cyclist of each crossing manoeuvres was calculated as follows: Crossing time = total number of frames / camera frame rate (approximately 30 frames/second)

Then the cyclist crossing speed for each crossing manoeuvres (field-measured speed) was calculated as follows: Field-measured speed = cyclist path length/crossing time

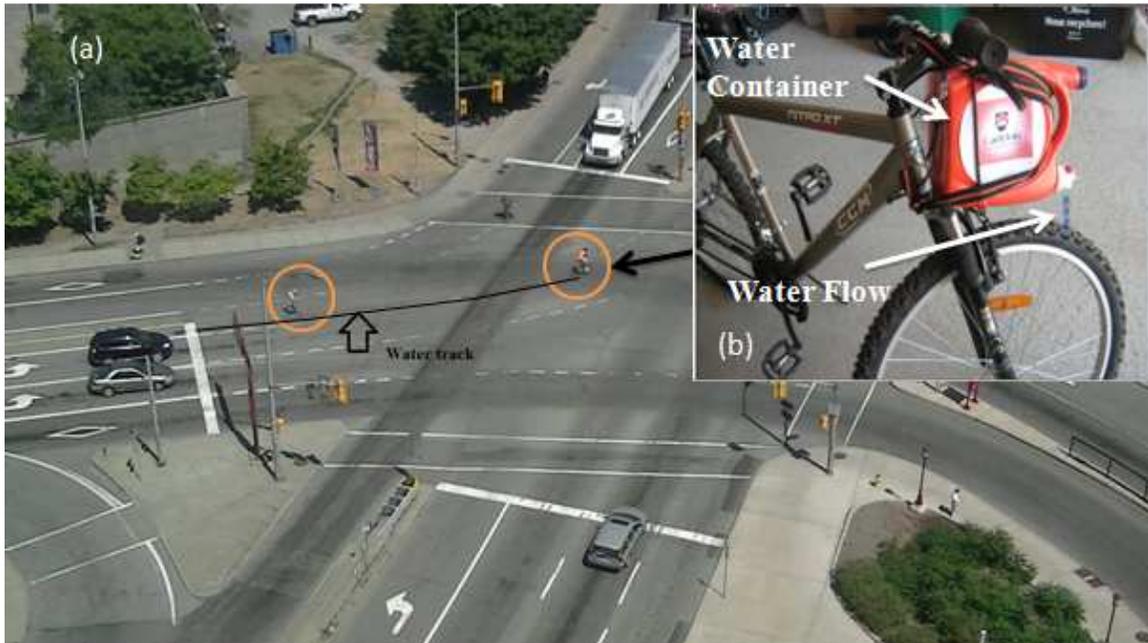


Figure 3-8: Cyclist track measurement using the CWT method during the validation experiment at study area. (a) video sample frame. (b) test bicycle.

3.2.2 Active crossing area (ACA)

The *second approach* was to obtain the field-measured speeds which follow pre-measured reference paths. These paths were defined and measured in the field prior to observing cyclists in the intersection. The reference paths were outlined within a frequently used part of the intersection, active crossing area (ACA) as shown in Figure 3-9. A total of 35 reference cyclist paths were marked on the pavement and measured in the

field using measuring wheel. Most of reference paths were located in the ACA as shown in Figure 3-10(a). Other reference paths were added to those within the ACA. These reference paths were for through and right-turn cyclist movements. In order to carry out field measurements, these paths were marked using different colored stickers (yellow and red) as shown in Figure 3-10(b) and Figure 3-10(c). The stickers helped in judging the similarity between cyclist tracks and the measured paths.

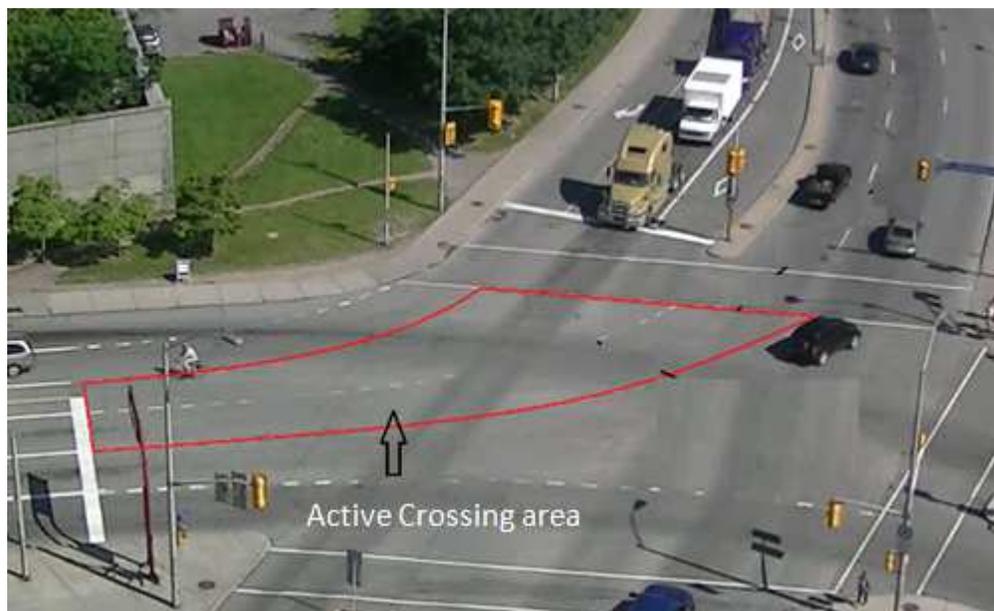


Figure 3-9: The active cyclist crossing area at the site location.

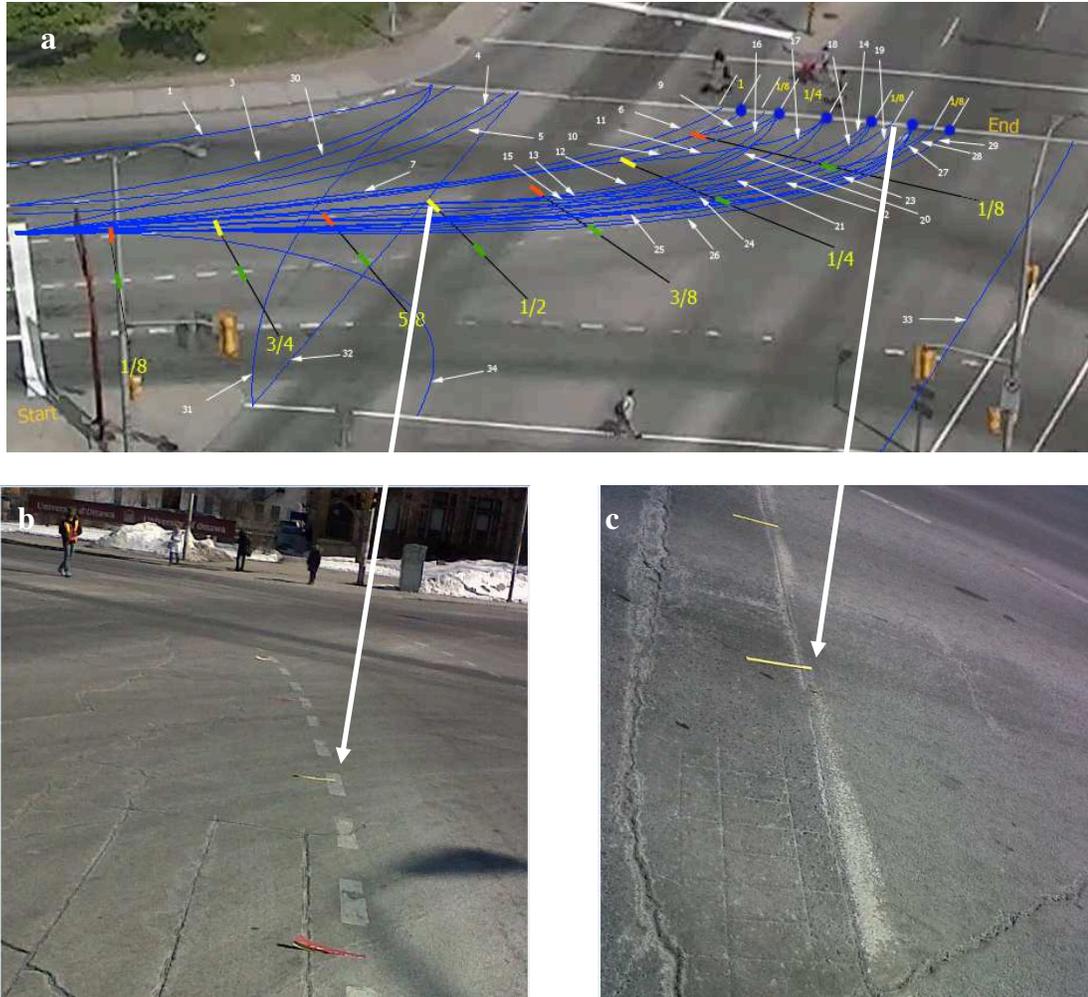


Figure 3-10: The field common cyclist paths measurements at the site location. (a) shows the reference paths. (b) and (c) show the color marks used in the field to delineate the paths.

The steps that were followed for the purpose of determining the path taken by the cyclist are: [i] each measured path was given an identification number, [ii] all videos were reviewed and cyclist tracks were compared with the paths measured in advance, [iii] the cyclist track was given the same number of the measured path which the cyclist was observed to follow or wobbling left and right not more than 0.25 meter [iv] if a cyclist

was travelling between two measured paths, this cyclist track distance was taken equal to the average of these two measured paths. The field-measured speed of 483 cyclists whose travel distance could be estimated following this approach was calculated using the same procedure as mentioned earlier.

3.2.3 Cyclist video-measured speed

Once the tracks were determined, speed of features can be calculated based on the displacement (dx and dy) of each feature between two successive frames. Those dx and dy represent the rates of change of any moving object feature (*e.g.*, vehicle and cyclist) being tracked. In order to calculate the dx and dy values of a moving object, the features (*i.e.*, vertices) that form this object must be considered as shown in Figure 3-11(a). Those features are detected and tracked by the tracker system. It is expected to keep the object active and move from frame to frame as shown in Figure 3-11(b). The red dots represent the new positions of moving features in a new frame. The tracker system keeps track of each feature and its coordinates (x and y) at each frame for which they exist. Using the coordinates (x_t, y_t) of each feature t , which are obtained using the tracker system, the change of coordinates (dxt, i, dyt, i) between frame i and the following frame is calculated as follows:

$$dxt, i = xt, i + 1 - xt, i; dyt, i = yt, i + 1 - yt, i \quad (3.1)$$

The average displacement of all features in a group is calculated between two successive frames. This average displacement of all constituent features becomes the displacement of the moving object. Thus, the Δx and Δy for the moving object can be

calculated as the average dxt and dyt of all features that form the object between two successive frames. The algorithm can be summarized in the following steps:

- 1- Let n be the number of features that form object O between two successive frames.
- 2- $\Delta xi = \frac{\sum_t dxt,i}{n}$ and $\Delta yi = \frac{\sum_t dyt,i}{n}$
- 3- Let N be the number of frames that form object O .
- 4- Let Dx be the sum of all Δxi values divided by N , let Dy be the sum of all Δyi values divided by N .
- 5- cyclist speed = $\sqrt{Dx^2 + Dy^2} \times \text{camera frame rate}$

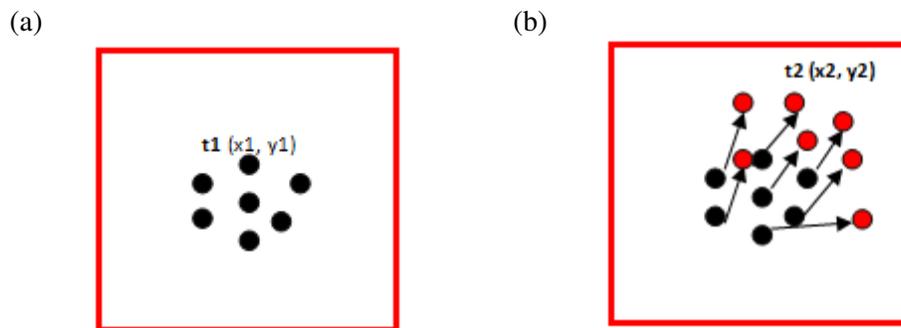


Figure 3-11: Schematic diagram illustrates object features: (a) features ($t1, t2, \dots, tn$) of the moving object in frame i . (b) new positions of moving features in frame $i+1$.

3.3 Validation of Cyclist Speed Measurements

In the context of this study, the validity of using computer vision techniques for speed measurement is determined by the accuracy of measured speed. In order to validate

the speed measurement technique, the accuracy is represented by the difference between average cyclist crossing speed measured using automated computer vision techniques and speed obtained using manual measurement techniques. In this study, cyclist movements with different geometry, including, straight and curve segments were included in the validation process.

3.3.1 CWT field-measured speed

As explained earlier the CWT is the cyclist path along a water trail generated by water-wetted bike tires on the pavement during a crossing maneuver. Table 3-5 shows the calculated error between the field-measured and video-measured speeds for the first approach (CWT) using 16 crossing events.

Table 3-5: Validation test between cyclist field-measured and video-measured speed for the first approach (16 crossing events).

Cyclist field-measured speed (m/s)	Cyclist video-measured speed (m/s)	Error (m/s)	% Error (Error divided by field-measured speed)
3.46	3.42	0.04	1.16
3.39	3.27	0.12	3.54
3.45	3.40	0.05	1.45
3.56	3.47	0.09	2.53
3.65	3.49	0.16	4.38
3.66	3.60	0.06	1.64
3.81	3.68	0.13	3.41
3.34	3.10	0.24	7.19
4.02	3.66	0.36	8.96
4.30	4.05	0.25	5.81
4.09	3.83	0.26	6.36
3.55	3.26	0.29	8.17
3.80	3.62	0.18	4.74
3.76	3.61	0.15	3.99
3.80	3.55	0.25	6.58
3.52	3.26	0.26	7.39

3.3.2 ACA field-measured speed

As explained earlier the active crossing area (ACA) is the part of the intersection that the cyclists frequently used to cross the intersection. This part includes 483 crossing events. Figure 3-12 shows the scatter plot of the cyclist field-measured speed, and the difference (error) between cyclist field-measured and video-measured speeds for the second approach (difference = field-measured – video-measured speeds). As shown in Figure 3.12, a very weak correlation was found between cyclist field-measured speed and the error between cyclist field-measured and video-measured. This provides evidence that the accuracy of the video analysis technique does not deteriorate at higher speeds. In addition, the majority of the error values between cyclist field-measured and video-measured speed were positive, *i.e.*, the video-measured speeds tend to be slightly underestimated. The likely explanation is that the video-measured speed tracks the average movement of different parts of the cyclist and the bike. The actual paths were measured at the active crossing area and represent the movement of the bike tires. It is likely due to lateral wobbling that the path traversed by the tires may be longer than the overall path traversed by the centeroid of the cyclist. The underestimation of video-measured speeds may have also resulted from the projection error of object features to the ground plan. Table 3-6 shows the summary of the results, it can be seen that the average field-measured speed is greater than the average video-measured speed. The average errors for both approaches were 0.18 m/s, and 0.16 m/s, respectively.

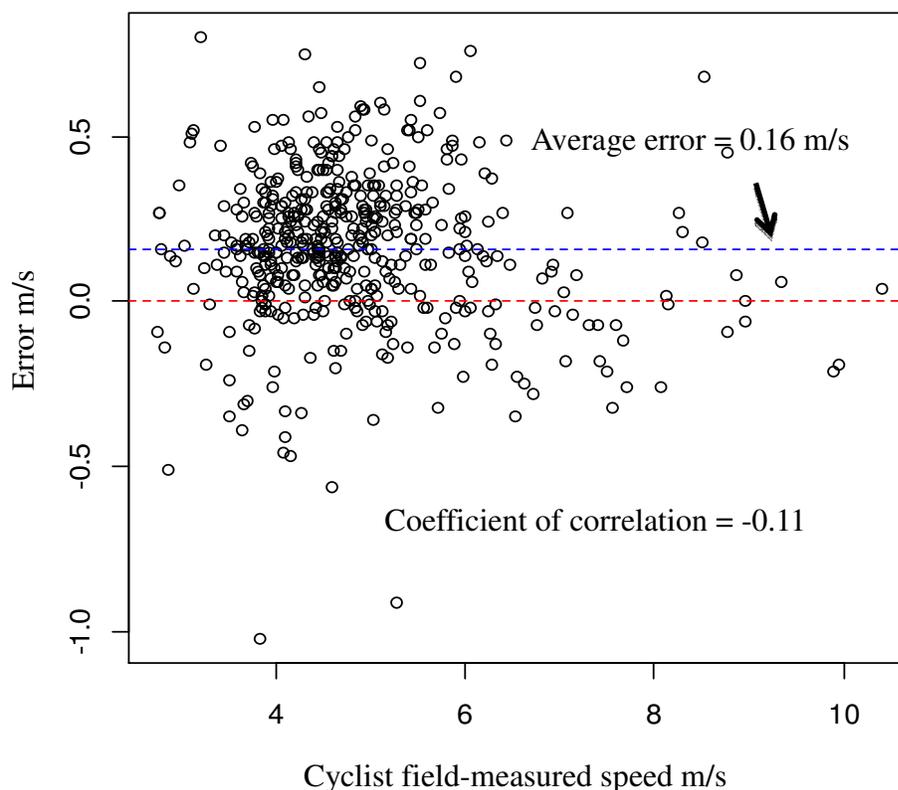


Figure 3-12: Relationship between cyclist field-measured speed and the error value between cyclist field-measured and video-measured speeds (483 crossing events).

Table 3-6: Summary results of cyclist field-measured and video-measured speeds.

Approach	Events	Cyclist speed (m/s) field/video		Average cyclist speed (m/s) field/video	Average error (m/s)	Overall* % absolute error (m/s)	Cyclist speed standard deviation (m/s) field/video
		Min.	Max.				
CWT	16	3.34/ 3.10	4.30/ 4.05	3.70/3.52	0.18	4.83	0.27/0.24
ACA	483	2.75/ 2.40	10.41/ 10.38	4.85/4.68	0.16	4.98	1.19/1.23

* Overall % absolute error = average (|error| / field cyclist speed)*100

The results of the validation showed an acceptable agreement between the field-measured and video-measured speeds. As shown in the Figure 3-13, the line of equality provides a very good fit for the two sets of measurement, with a high coefficient of determination, R^2 , of 0.97.

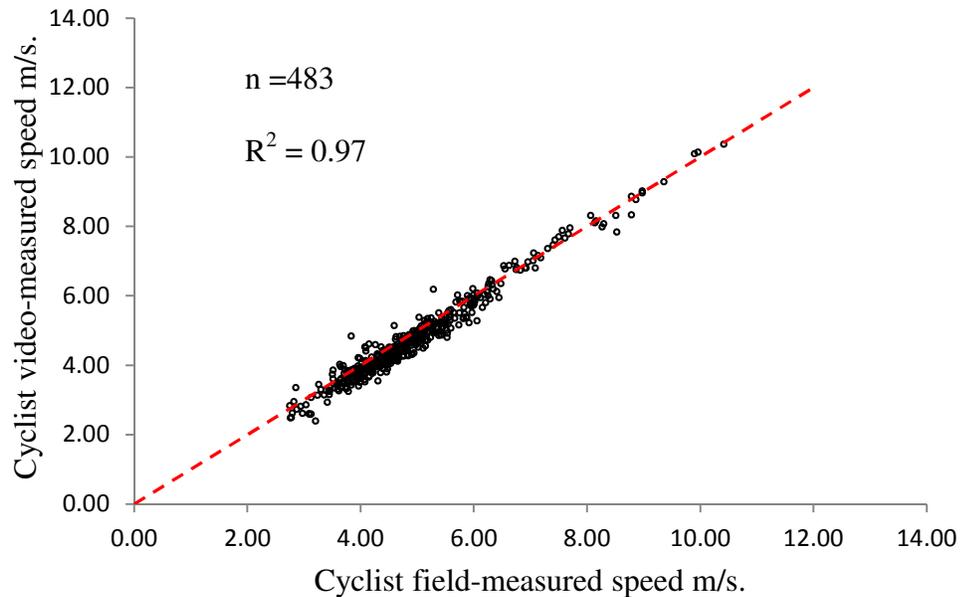


Figure 3-13: Relationship between cyclist field-measured and video-measured Speeds.

3.4 Data Analysis and Results

3.4.1 Descriptive statistics

Figure 3-14 and Figure 3-15 provide frequency distributions of cyclist field-measured and video-measured speed. As shown in these figures, the majority of the cyclist speeds are between 4.0 m/s and 6.0 m/s for both field-measured and video-measured speed.

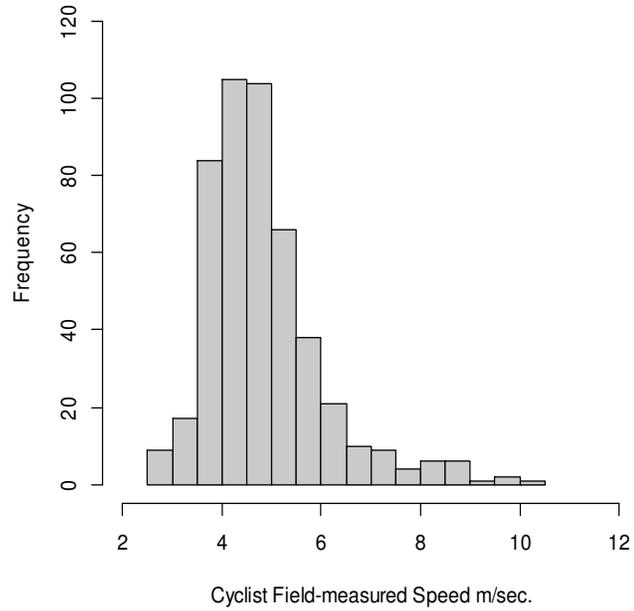


Figure 3-14: The frequency distribution of 483 cyclists' field-measured speed.

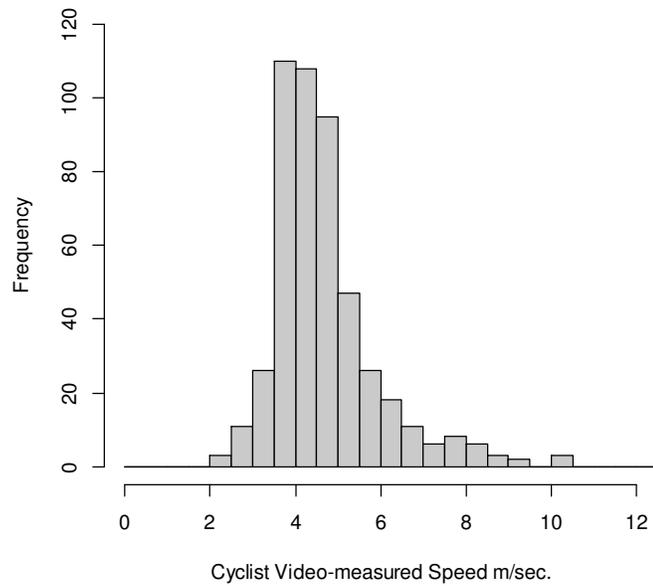


Figure 3-15: The frequency distribution of 483 cyclists' video-measured speed.

In order to identify the appropriate distribution of the field-measured and video-measured speeds, different fitting distributions were examined. Depending on the lowest

value of the log likelihood, the distribution was selected. Table 3-7 shows the log likelihood for each distribution. As shown in the Table 3-7, the Generalized Extreme Value was found to be the best fit distribution for both field-measured and video-measured speeds. This similarity in the distribution confirms the possibility of using the video-measured technique to measure the cyclist speed. In addition, this similarity can be considered as a validation of the video-measured technique to measure the cyclist speed.

Table 3-7: Fit Distribution of the field-measured and video-measured speeds.

Distributions	Field-measured speed			Video-measured speed		
	Log likelihood	Mean m/s	Standard deviation m/s	Log likelihood	Mean m/s	Standard deviation m/s
Normal	-767.82	4.84	1.18	-1,432.74	5.11	1.79
Lognormal	-710.29	4.84	1.09	-1,297.35	5.09	1.58
Gamma	-724.61	4.84	1.10	-1,328.87	5.11	1.60
Generalized Extreme Value	-696.81	4.85	1.15	-1,273.97	5.10	1.74

3.4.2 Factors affecting crossing speed

Further analysis was carried out on the data to investigate different factors affecting cyclist speed. This was achieved by conducting multi-way Analysis of Variance (ANOVA) in order to investigate the effect of these factors on cyclist speed. The factors studied were: [i] whether the crossing action is a traffic violation (against traffic, zigzag crossing, hit the signal, cyclist fails yielding to the vehicle, omitting/inadequate use of hand signal, and inadequate lighting), [ii] weather conditions (rain or dry as was visible in the videos), [iii] roadway condition (wet or dry as was visible in the videos), [iv] time of the day (day-time and dusk/night), [v] longitudinal grade (uphill and downhill), [vi] crosses without stopping (cyclist started from stop or continued to cross from motion),

and [vii] whether the cyclist was involved in a traffic interaction (only PET of 3 seconds or less was adopted to identify manually the interaction events involving a cyclist and a vehicle (Wang & Nihan, 2004)).

A summary of the results of different factors affecting cyclist speed is given in Table 3-8. In addition, ANOVA was conducted to investigate the effect of these factors on the error in speed measurements. This was performed to identify recurrent conditions which affect tracking performance. Table 3-8 shows the ANOVA results of the cyclist speed unsigned and signed error.

As shown in Table 3-8, the roadway conditions (wet or dry), longitudinal grade (uphill and downhill), and crosses without stopping were found to be significant in affecting cyclist speed (716 events). The same factors except for roadway conditions (wet or dry), were found to be significant based on analysis of the measured speed for the same validation events (483 events). This provides further evidence as to the reliability of the video analysis technique. On the other hand, the factors rain, grade and crosses without stopping were found to be significant in affecting cyclist speed (483 events) as shown in Table 3-9.

Table 3-8: ANOVA results for effect of different factors on cyclist crossing speed.

Factor	<i>p</i> -value		
	Video-measured speed		Field-measured speed
	(716 events)	(483 events)	(483 events)
Violation	0.92	1.90	0.23
Rain	0.32	0.65	0.77
Roadway	< 0.05	0.26	0.41
Day-time	0.99	0.55	0.91
Grade	< 0.05	< 0.05	< 0.05
Crosses without stopping	< 0.05	< 0.05	< 0.05
Interaction	0.38	0.27	0.35

Table 3-9: ANOVA results of the cyclist speed unsigned and signed error between field-measured and video-measured speed.

Factor	<i>p</i> -value	
	Unsigned (absolute) error (483 events)	Signed error (483 events)
Violation	0.09	0.65
Rain	0.001	< 0.05
Roadway	0.98	0.15
Day-time	0.68	0.017
Grade	< 0.05	< 0.05
Crosses without stopping	0.003	< 0.05
Interaction	0.97	0.48

3.4.3 Comparison of cyclist speed results

As described in Chapter Two, different cyclist speed results were obtained in previous studies. Table 3-10 summarizes the reported cyclist speeds along with the related information of the device utilized, observed movement, location, and sample size. Figure 3-16 provides a graphical comparison of these cyclist speed results. The figure shows that the range of the average speeds of the cyclists using the GPS devices was between 12 km/h and 22 km/h. While the range of the average speeds of the cyclists using video camera devices was between 11 km/h and 25 km/h.

Table 3-10: Summary of the cyclist speed measurements using different devices.

Study	Locations	Devices	Movement direction	Mean speed Km/h	Total number of speed observations
(Thompson, <i>et al.</i> , 1997)	Road section	Rader gun & Stopwatch	Through	14.80	152
(Virkler & Balasubramanian, 1998)	Two trail road sections	Stopwatch	Through	20.74 21.42	105 35
(Akira, <i>et al.</i> , 2012)	Public road	Experimental bicycle	Through	16.80	120
(Hunter, <i>et al.</i> , 2009)	Two traffic corridors	Pneumatic tube sensor	Through	19.30	Not given
(Parkin & Rotheram, 2010)	Different road (journey)	GPS	All movements ¹	21.60	518
(El-Geneidy, <i>et al.</i> , 2007)	On/off street	GPS	All movements	15.62 ² 16.25 ³	Not given
(Dill & Gliebe, 2008)	not given	GPS	All movements	17.28	164
(Witte & Wilson, 2004)	Running track	GPS & speedometer	Curved and Straight road	Not given	5,060
(Sarosh & Winai, 2001)	Road section	Video camera	Through	24.83	53
(Huang & Jianping, 2004)	Signalized intersection	Video camera	Through & Left	11.44 13.21	470
(Opiela, <i>et al.</i> , 1980)	Signalized intersection	Video camera	Through	Not given	Not given
(Wayne, 1997)	Signalized intersection	Video camera	Through	12.70	442
(Singh, 2012)	Road section	Video camera	Through	22.30 ⁴ 19.80 ⁵ 16.20 ⁶	187 ⁴ 194 ⁵ 20 ⁶
(Davies, <i>et al.</i> , 2003)	Road section	Video camera	Through	16.00	1,681

1 right, through, and left movements, 2 on street, 3 off street, 4 before cyclist lane marked, 5 after cyclist lane marked (vehicle zone), 6 after cyclist lane marked (pedestrian zone),

Table 3-10: continued

Study	Locations	Devices	Movement direction	Mean speed Km/h	Total number of speed observations
(Zaki, <i>et al.</i> , 2013)	Section in roundabout	Video camera	Through	16.78 ⁷ 14.26 ⁸	681 53
(Wachtel, <i>et al.</i> , 1995) ⁹	Six Signalized intersection	Not given	Through	14.00 11.75 11.27 8.37 10.14 11.59	131 61 113 123 83 39
(Wheeler, <i>et al.</i> , 2010) ⁹	Two Signalized intersection	Video camera	Through	11.54 ¹⁰ 11.92 ¹¹	189 223
Current study	Signalized intersection	Video camera	All movements & curvilinear	17.46 ¹² 16.85 ¹³ 18.40 ¹⁴	483 ¹² 483 ¹³ 716 ¹⁴

7 group size = one cyclist, 8 group size > one cyclist, 9 Cyclist speed has been studied at the signalized intersections by other researchers, as a part of their studies. 10 winter season, 11 summer season, 12 field-measured speed 483 cyclists, 13 vide-measured speed 483 cyclists, 14 video-measured speed 716 cyclists.

Based on reported results in the reviewed studies, it is possible to examine whether cyclist speed depends on whether the crossing action was conducted within an intersection or along a road segment. The histograms presented in Figure 3-17 show the distribution of the average crossing speed reported in reviewed studies measured at signalized intersections and road sections.

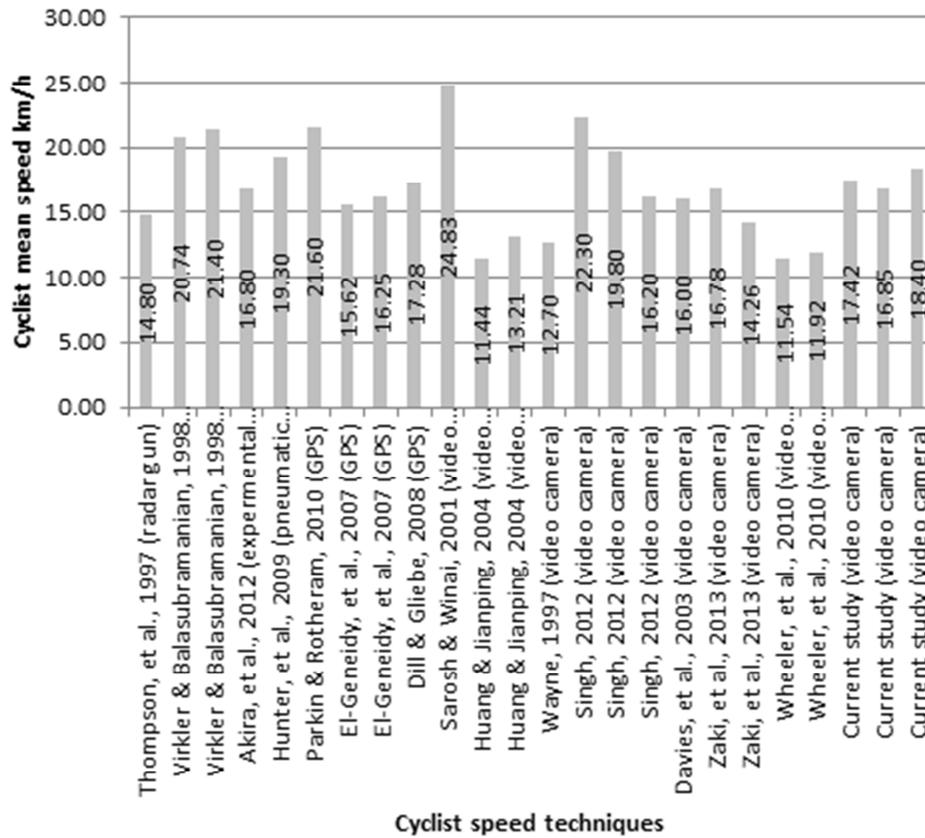


Figure 3-16: Average cyclists’ speed distribution used different measurement techniques.

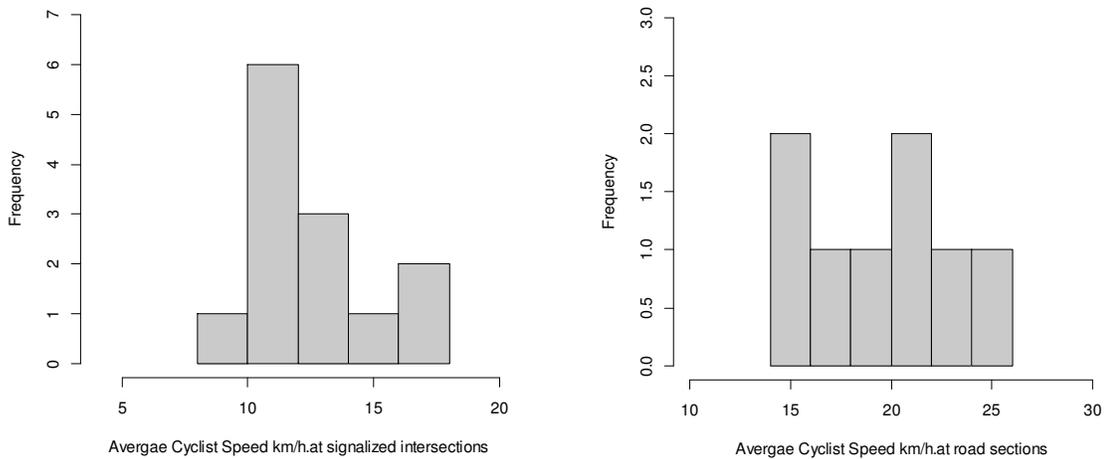


Figure 3-17: Average cyclist’s speed distribution at the road sections and signalized intersections.

Meta analysis was conducted to test whether reported cyclist speed in the reviewed studies depends on where they were recorded; intersections or road segments. The unpaired t-test is conducted to examine the significance of the difference in mean speed values. The null hypothesis (H_o) is defined such that there is no significant difference between the mean of the cyclist speed at the signalized intersections and the road sections. Table 3-11 summarizes Meta analysis results.

Table 3-11: Meta analysis results of the average cyclists' speed at the road sections and the signalized intersections.

Location	Number of observations (n)	Mean ¹ speed (km/h)	Standard deviation ¹ (km/h)
Road sections	2,427	17.19	16.98
Signalized intersections	3,101	13.81	13.21

¹ the reported mean and standard deviation is the weighted mean and standard deviation which was calculated by weighing every reported average speed by the sample size.

The t -statistic is calculated as follows:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \quad (3.1)$$

Where t is the t -statistic, \bar{x}_1 , and \bar{x}_2 are the average of all reported cyclist speed for the road sections and signalized intersections respectively, s_1^2 , and s_2^2 are the variances, n_1 , n_2 are their number of the observations. Since the calculated t -statistic (0.16) is less than the t -tabulated value (1.96), the null hypothesis cannot be rejected. It is concluded that there is no significant difference between the means of the cyclist speed at the 5% level of significance.

3.5 Summary

This chapter presented the technical details of comprehensive evaluation of computer vision technique to measure cyclist field-speed at signalized intersection. The average crossing speed calculated in this study is based on the average of all displacements of the constituent features (cyclist's features). The results of the video-measured speed of the cyclist are generally acceptable in terms of the accuracy (average percentage of error). This demonstrates that video-measured speed can be performed using computer vision techniques.

Two different validation approaches were performed to measure the accuracy of video-measured speeds. The differences in the average field-measured and video-measured speeds for both validation approaches were 0.18 m/s (16 tracks), and 0.16 m/s (483 tracks). In this study, cyclist movements with different geometry were included in the validation process. The overall average percentage in speed error (absolute speed difference divided by field-measured speed) for both validation approaches were 4.83% (16 tracks) and 4.98% (483 tracks).

A high coefficient of determination of 0.97 was found between field-measured and video-measured speeds. This high coefficient of determination can be considered as a validation of the video-measured technique to measure cyclist speed. A very weak correlation was found between cyclist field-measured speed and the error between cyclist field-measured and video-measured. This provides evidence that the accuracy of the video analysis technique does not deteriorate at higher speeds.

The Generalized Extreme Value was found to be the best fit distribution for both field-measured and video-measured speeds. This shows consistency in distribution

between field-measured and video-measured speeds. This similarity in the distribution confirms the possibility of using the video-measured technique to measure the cyclist speed. Statistical analysis was conducted to investigate the effect of different factors on cyclist speed.

Factors such as the longitudinal grade (uphill and downhill), and crosses without stopping significantly influence the cyclist field-measured speed (483 events). The same factors were found to be significant based on analysis of video-measured speed for the same validation events (483 events). This similarity in the outcome provides evidence as to the reliability of the video analysis technique.

This study is non-intrusive with regard to the position of the camera, which was placed on the 12th floor of the building. However, different shortcomings occurred due to the position of the video camera. The video camera was laterally far from the study location and the angle of the camera made it challenging to obtain the tracks of road users using the tracker system when there was traffic congestion. During the traffic congestion, some cyclists would disappear behind vehicles as well as, when cyclists crossed the intersection within a group of cyclists. These shortcomings can be attributed to using improper or suboptimal configuration tracking parameters. A careful consideration of this issue was conducted in this study, but the problem of over-grouping still occurred. Over-grouping is defined as two or more different objects are grouped or merged, incorrectly, by the tracker system and considered as one object. In addition, the study did not take into account characteristics that were not easy to determine in the videos such as the cyclist age and gender.

4. AUTOMATED MEASURING OF CYCLIST-MOTOR VEHICLE POST ENCROCHMENT TIME AT SIGNALIZED INTERSECTIONS

Proximity of cyclists to vehicles is a major source of safety hazard facing cyclists. Traffic interactions have been utilized to quantify proximity and estimate the long-term risk of collisions. Interactions between motor vehicles and cyclists at signalized intersections were characterized in this chapter using an objective conflict indicator; Post-Encroachment Time (PET) which offers a simple way to measure proximity without the need for the prediction/extrapolation of future positions or speeds of road users. This indicator can objectively measure the proximity of road users to collision in situations in which a collision was not observed (Tarko, *et al.*, 2009). PET can be defined as the time difference between the moment a first road user leaves an area of potential collision and the moment of arrival of a second road user to this area while possessing the right of way (Allen, *et al.*, 1978).

In this study, PET is defined as the time difference between the moment when a rear part of the vehicle leaves (or a front part arrives) at the area of potential collision and the moment when a bicycle's front wheel arrives at (or the rear wheel departs from) this area. Depending on the movements of both the cyclist and motor vehicle, different types of crossing actions may occur between the paths of the vehicle and the cyclist. These crossing actions will be referred to in this study as interactions. It should be noted that this study is not intended to count interactions or to suggest a PET threshold, but to describe a novel technique to automatically measure PET between cyclists and vehicles.

The objective of this chapter is to present an automated method to measure PET between cyclists and motor vehicles using video analysis techniques and to test the accuracy of the automated technique compared to manual methods. This chapter is organized in three main sections. In the first section, data collection is described. The chapter then proceeds to describe different PET measurement methods and the algorithm used for automated PET measurement in the second section. In the third section statistical analysis is presented for the validation of the PET measurement methods, descriptive statistics of PET, and PET agreements in three different measurement methods.

4.1 Data Collection

4.1.1 Study site and equipment

For the purpose of studying PET between a motor vehicle and a cyclist, a video camera was used to monitor cyclist and motor vehicle movements. The case study in this chapter uses two four-leg signalized intersections in the Downtown of Ottawa, Ontario. A map showing the location of the intersections is shown in Figure 4-1. The view of the intersections as viewed from the monitoring camera is shown in Figure 4-2. As shown in this figure, Laurier Avenue West, which is a two-lane two-way street with a segregated cyclist lane on each side of the street, is the common street in both intersections. The intersecting street at the first intersection, which is Lyon Street, is a three-lane one-way street with bike lane on the right side. The left most lane is used as a parking lane during off-peak. The intersecting street at the second intersection, which is Bay Street, is a two-lane one-way street with bike lane on the right side. The allowable movements within the intersections for motor vehicles on both directions of Laurier Avenue West are through/left-turn for one direction and through/right-turn for the other direction. As well,

the allowable movements within Laurier Avenue West for cyclists are through for both directions and right-turn for the direction turning into the traffic direction of the intersecting street. Left turns for cyclists in the relevant direction are allowed by using a bike box and crossing the intersection during the pedestrian signal.

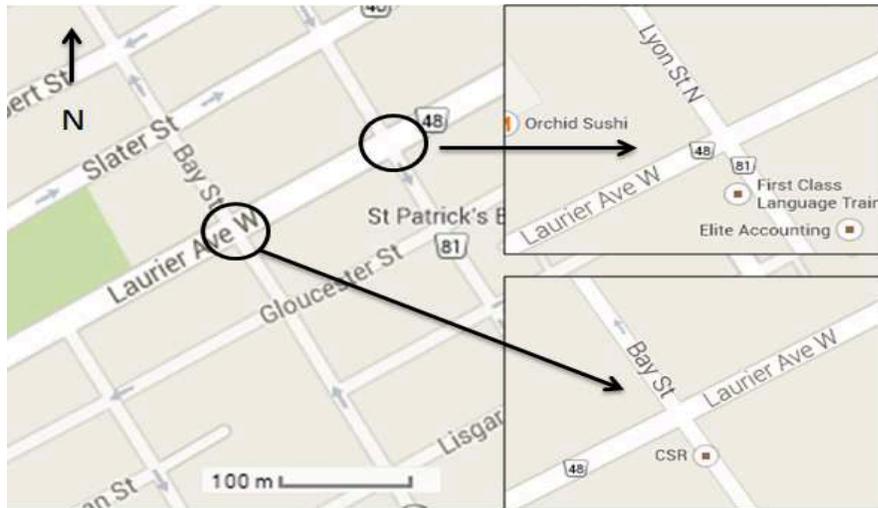


Figure 4-1: Location of the study intersections (Google, 2013).



----- Vehicle movements, — Cyclist movements

Figure 4-2: Camera views. (a) camera view of the Lyon-Laurier intersection. (b) camera view of the Bay-Laurier intersection.

4.1.2 Video data collection

In order to have a reasonable view of areas of interest in each intersection, a video camera was placed at elevations of the 15th and 17th floors of two buildings overlooking the intersections. The cameras were placed in those positions to cover all approaches where the paths of through cyclists and turning vehicles would be expected to intersect. Figure 4-3 shows the camera level position and the camera setup at the signalized intersections. Data was collected in two different time periods to cover two different seasons: fall and summer. The fall period covered 4 days in the months of September and October in 2011. The summer period covered 5 days in the months of June and July in 2012. The total hours of the data collection were 84 hours. A total of 2,736 cyclists and 11,804 vehicles were observed in this study. Out of these, a total of 1,901 through cyclists and 6,648 right-turn and 5,156 left-turn vehicles were observed.

4.1.3 Surveying works and camera calibration

In order to conduct metric measurements, the recording camera is calibrated by knowledge of geometric primitives (points and lines) that appear in the video and their corresponding appearance in an orthographic image of the intersection. Therefore, field measurements were made to establish the true lengths of segments that appear in the video in order to aid the calibration process.

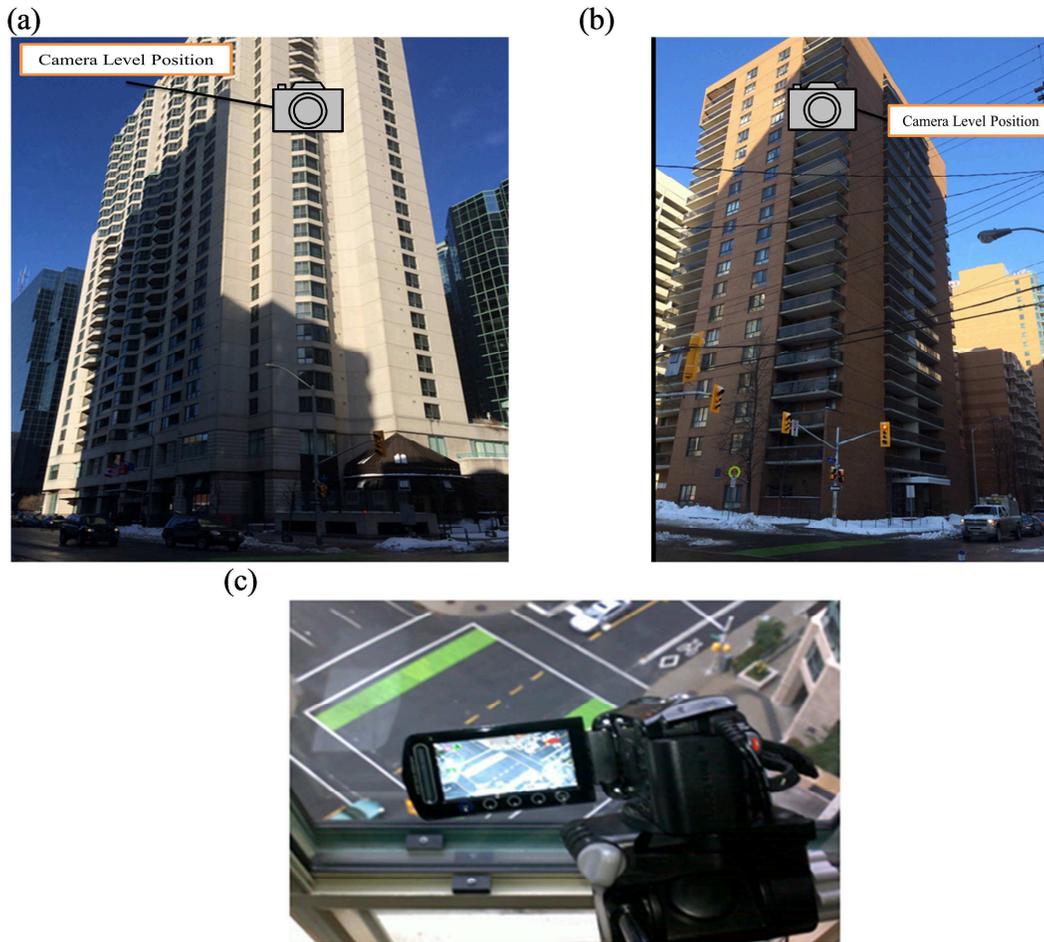


Figure 4-3: Camera level positions at the site location. (a) at Laurier-Lyon intersection. (b) at Laurier-Bay intersection. (c) camera setup.

4.1.3a Surveying works

In order to transfer the pixel coordinates to world image coordinates using the homography matrix, two snap shots should be prepared. The first snap shot is from the video itself as one frame, the second one is the world image for the same location (study area). It is possible to obtain the world image for the location from mapping websites directly. However, there are some obstacles that may make it difficult to obtain the world

image of the location. For example, high buildings will mask or overshadow the study area, thus making it difficult or impossible to locate points and lines in the work image. This problem was faced in this study. The intersections along Laurier Avenue were surrounded by high-rise buildings and therefore they made it difficult to obtain acceptable quality world image for the two intersections that were included in the study.

To overcome this problem, detailed surveying works were conducted using a Total Station for the purposes of obtaining the coordinates of the points that can be seen in both video frame and study area world image. Local coordinates were used and different points were surveyed, such as ends of a stop line, curb side, centers of manhole covers, and street furniture. At each intersection there were 25 – 30 points surveyed, and these points were identified in the video frame in advance. In addition the measuring wheel was used to measure several lines in the study area that appears in the video frame. The coordinates of the surveyed points were located and plans were constructed to represent the world image of the intersection as shown in Figure 4-4.

4.1.3b Camera calibration

Video data show the position of individual cyclists relative to the road and to other users, on the plane of the camera sensor. These metrics are affected by the position of the camera itself. Distances and angles determined from a video frame are *distorted* values from the *real* dimensions, due to parallax, perspective, and other geometric factors.

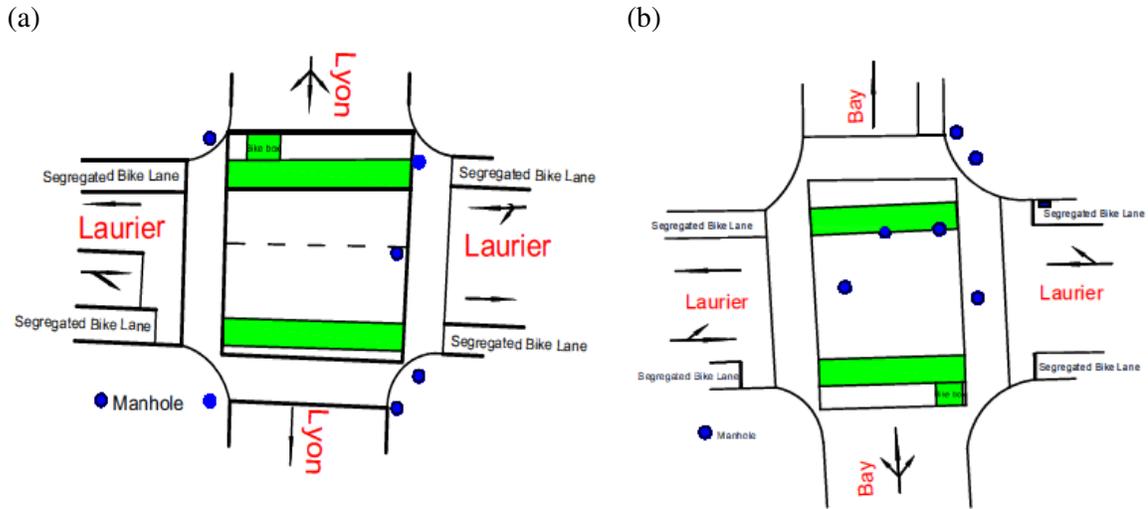


Figure 4-4: AutoCAD intersections drawing. (a) Laurier-Lyon intersection. (b) Laurier-Bay intersection.

To conduct metric measurements at a video frame, a mathematical transformation from photographic coordinates to real world coordinates is required. This coordinate transformation is accomplished by using a homographic transform that effectively translates and rotates raw (video) measurements to real world values. Camera parameters were estimated using a special calibration script discussed by (Ismail, *et al.*, 2013) in order to produce a homography matrix. Figure 4-5 and Figure 4-6 illustrate linear segments that were taken as measurements required for camera calibration, orthographic grid, and world grid.

4.2 PET Measurement Methods

In addition to the automated method investigated in this study, two more PET measurement methods were used in order to examine the accuracy and validity of the automated method. Therefore, three measurement methods are used in this study, and can be classified as: [i] manual video timer measurement (MVTM) method, [ii] manual frame

count measurement (MFCM) method and [iii] automated measurement (AM) method. The three methods are described in the following subsections.

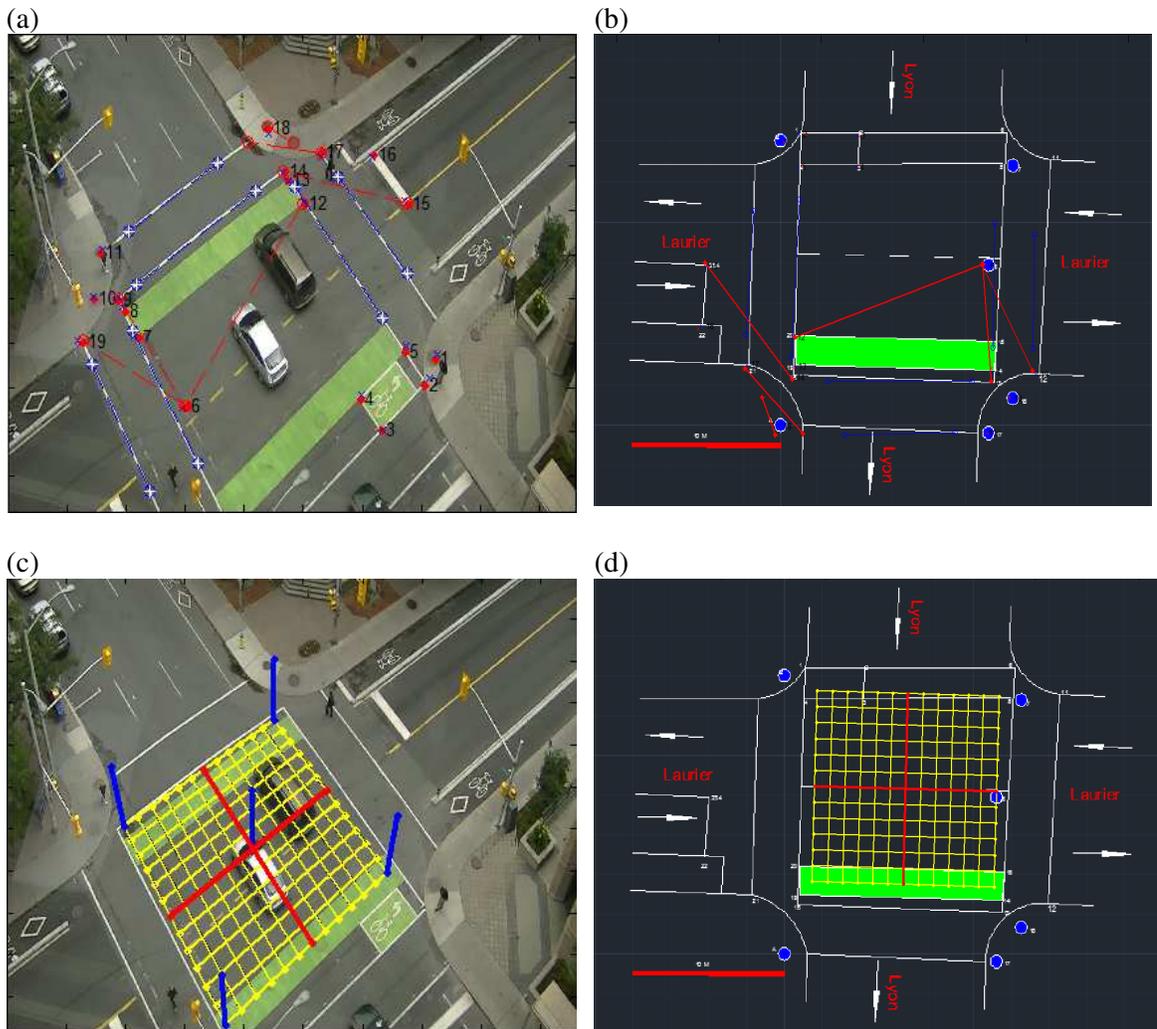


Figure 4-5: Illustration of camera calibration. (a) field distance measurements. (b) projected distance using camera calibration. (c) orthographic grid. (d) world grid at the area of study (signalized intersection at Laurier and Lyon streets).

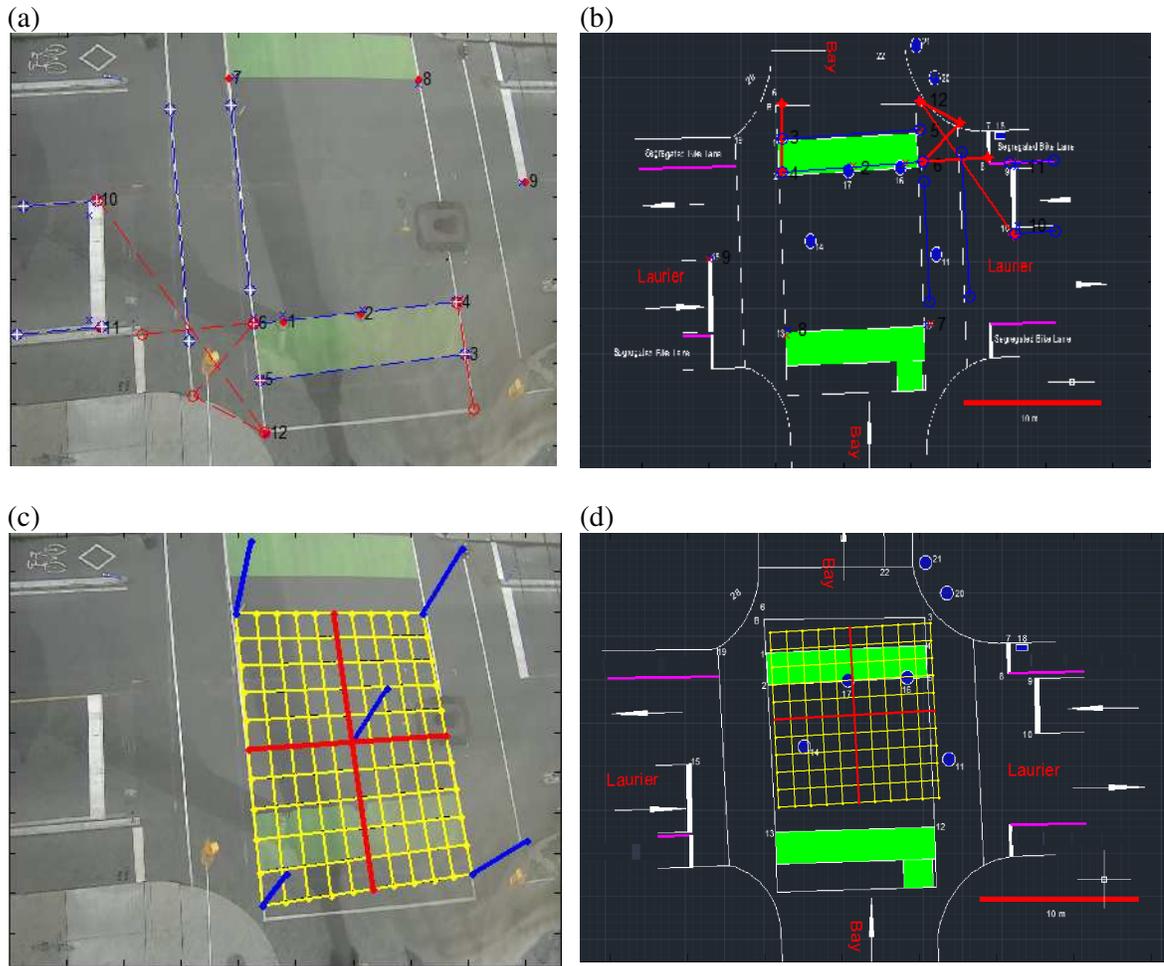


Figure 4-6: Illustration of camera calibration. (a) field distance measurements. (b) projected distance using camera calibration. (c) orthographic grid. (d) world grid at the area of study (signalized intersection at Laurier and Bay streets).

4.2.1 Manual video timer measurement (MVTM) method

This method was used to identify the behaviour of the cyclists and motor vehicles while crossing the signalized intersection and determine PET. Videos were manually observed to identify possible types of interaction. Two types of interactions observed at the study area were: right-turn and left-turn interactions. These types of interactions were defined by the turning direction of motor vehicles at an intersection. As shown in Figure 4-7, a right-turn interaction happens when a motor vehicle is making a right turn, and crossing the cyclist path while the cyclist is moving straight. The figure also shows that left-turning vehicles may collide with a through-moving cyclist while crossing the intersection and coming from the opposite direction. Furthermore, four different scenarios were observed under these two interaction types depending on whether the turning vehicle or the cyclist arrives first at the possible collision area as shown in Figure 4-8.

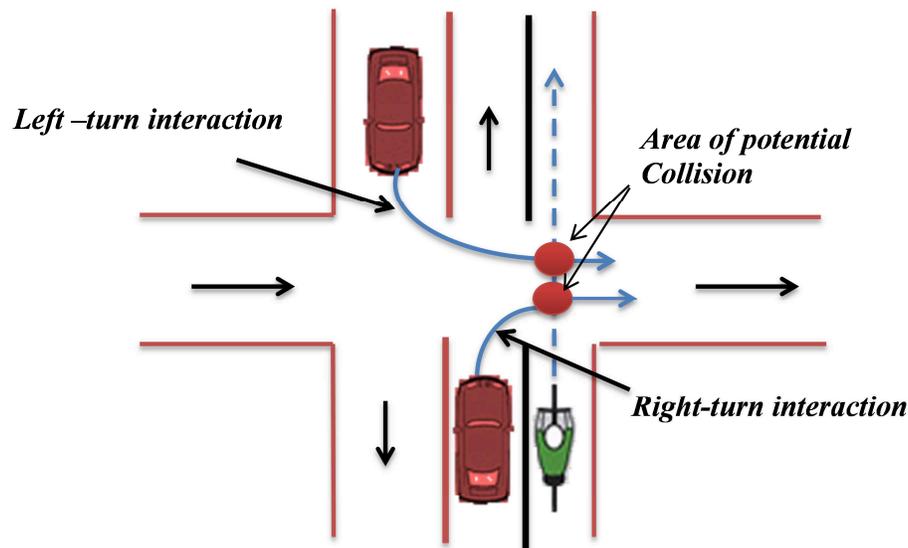


Figure 4-7: Schematic diagram: right-turn and left-turn interactions.

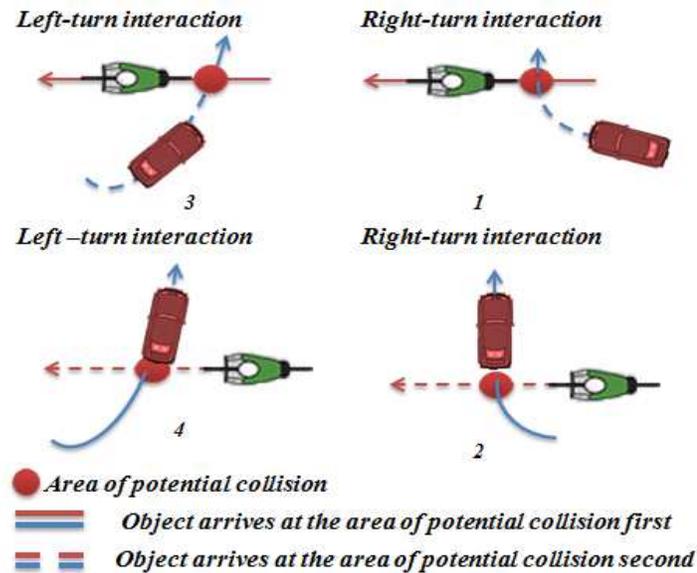


Figure 4-8: Four different cyclist-motor vehicle interaction scenarios.

In order to count the number of interactions, a threshold needs to be selected in order to count observations below it. The PET values which are too long may not be relevant to safety and may be more related to regular non-hazardous events. Several studies used a critical threshold of 3.0 seconds for PET between pedestrians and vehicles (Lord, 1996) and between vehicles (Gettman, *et al.*, 2008), while other recent studies investigated thresholds between left turning and opposing through vehicles below 3.0 seconds with good correlation to crash frequency for 1 second threshold (coefficient of determination above 0.6) (Peesapati, *et al.*, 2013).

A PET threshold of 3.0 seconds was used in this study to identify events involving a cyclist and a vehicle based on the following reasons; [i] (Wang & Nihan, 2004) found that the average proximity between vehicle and cyclist at signalized intersection was 2.93

seconds. [ii] an average cyclist speed of 4.59 m/s was found within PET (0, 3] seconds in this study by applying the tracker system to the study area (Lyon-Laurier and Bay-Laurier intersections), and the cyclist path across the intersection was measured to be 17.80 meter, then the crossing time is estimated to be 3.88 seconds. Furthermore, the proximity between cyclists and vehicles in these interactions was classified to three different levels based on the PET value. These PET classification levels are (0, 1], (1, 2], and (2, 3] seconds. The precision of this MVTM method is the same as the precision of video time, which is 1.0 second. In this MVTM method, the video time was used to measure the PET between the cyclist and motor vehicle. The PET is the time difference between the moment at which the motor vehicle part that caused the interaction leaves (or arrives at) the area of potential collision and the moment at which the cyclist part involved in this interaction arrives at (or leaves) the area of potential collision. The part of the motor vehicle can be any of the four corners of the vehicle, and the part of the cyclist is either the edge of the rear or front wheels.

4.2.2 Manual frame count measurement (MFCM) method

This method was applied to the same events that were observed in the previous method. This method is virtually error-free and therefore will be used as a benchmark to evaluate the accuracy of other methods. In this method, the area of potential collision is defined as the location of the intersection point between the paths of motor vehicle and cyclist according to each scenario as shown in Figure 4-8. This intersection point was identified visually by running the video and marking the intersection point using a pointing device on the computer screen. The accuracy of the observer affects the accuracy of identifying this point of intersection. Each event was independently reviewed

by two different observers in different sessions. The first observer was the thesis author who reviewed all events. The second observer was a group of undergraduate and postgraduate students, all events distributed to the group members to perform the review. If the difference between the two observers was more than two frames, the event is reviewed for a third time by the thesis author. For each event, the frame count was recorded for the two frames corresponding to the arrival of the interacting motor vehicle and cyclist. The first frame was recorded when the rear end of the vehicle (or the cyclist) leaves an area of potential collision. The second frame was recorded when the front end of the cyclist (or the vehicle) arrives at the area of potential collision. PET was calculated as the difference in frame count of these two frames divided by the camera frame rate. The camera frame rate is approximately 30 frames/second, and therefore the precision of the estimated PET is 0.033 second.

4.2.3 Automated measurement (AM) method

For the automated measurement method, the PET was measured using a tracker system that uses computer vision techniques as described in Chapter 3. Many improvements to the tracker system that was used in Chapter 3 were conducted in order to overcome the shortcomings that were mentioned in Chapter 3. Details about the tracker system improvements illustrated in the Appendix (B). In this tracking system, each object of interest, a motor vehicle or cyclist, was represented by a set of features. A bounding box, for each object, is formed as the smallest upright rectangle containing all object features as shown in Figure 4-9. As shown in Figure 4-9 some areas are masked out. These areas may contain objects that are considered irrelevant or unimportant to the interaction analysis such as pedestrians or other moving objects.

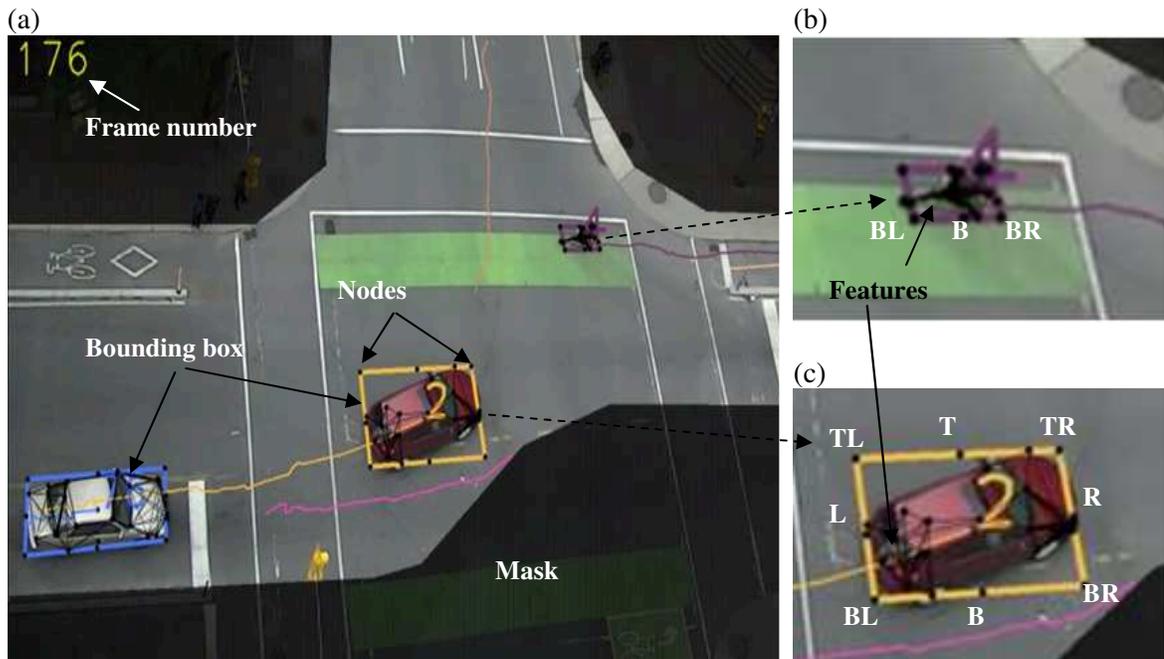


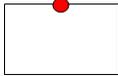
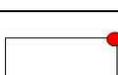
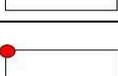
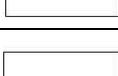
Figure 4-9: Object features, bounding box, nodes and mask in the tracker system.

(a) the masked area, and the bounding box. (b) the cyclist features and bounding box nodes. (c) the motor vehicle features and the bounding box. The coordinates in world and the x-axis increases to the right, y-axis increase to the bottom.

As shown in Figure 4-9, eight different nodes were identified on the bounding box of a motor vehicle. These nodes are: Top-Right, Right, Bottom-Right, Bottom, Bottom-Left, Left, Top-Left, and Top (as shorthand will be denoted *TR*, *R*, *BR*, *B*, *BL*, *L*, *TL*, and *T* respectively). Similarly, the bounding box of a cyclist with eight nodes was defined, and the three nodes with lowest elevation were used: Bottom-Right, Bottom and Bottom-Left (or *BR*, *B* and *BL*, respectively) because those positions should ideally represent the projection of the cyclist's position on the pavement plane. The majority of the bounding boxes fully contained the outline of road users as shown in Figure 4-9(a). However, there

are some bounding boxes which partially covered road users. In case of partial coverage, the part of road users covered by bounding box is used to calculate PET. Table 4-1, shows how the bounding box nodes coordinates were calculated. The following section describes in detail the computation technique to estimate PET using the AM method. Similar to the MVTM method, the precision of AM method should be 0.033 second.

Table 4-1: Coordinates of the bounding box nodes.

Object Position	Symbol	Bounding box nodes X, Y coordinates	Node position location
Top	T	X = The x coordinate of the object centre *, Y = The smallest y coordinate of all features.	
Bottom	B	X = The x coordinate of the object centre *, Y = The largest y coordinate of all features.	
Right	R	X = The largest x coordinate of all features, Y = The y coordinate of the object centre *.	
Left	L	X = The smallest x coordinate of all features, Y = The y coordinate of the object centre *.	
Top-Right	TR	X = The largest x coordinate of all features, Y = The smallest y coordinate of all features.	
Top-Left	TL	X = The smallest x coordinate of all features, Y = The smallest y coordinate of all features.	
Bottom-Right	BR	X = The largest x coordinate of all features, Y = The largest y coordinate of all features.	
Bottom-Left	BL	X = The smallest x coordinate of all features, Y = The largest y coordinate of all features.	

* The object centre is the center of the best fitting ellipsoid (Bradski & Kaehler, 2008) that encloses the majority of object features according to a least-squares fitness functions.

4.3 Automated PET Measurement

The section presents the algorithm used to compute the value of PET between the motor vehicle and cyclist in the AM method and detailed description of the steps in this algorithm.

4.3.1 PET measurement algorithm

This algorithm uses the tracker system output to calculate the value of PET. The output of the tracker system is a collection of objects' tracks, where an object can be a vehicle, cyclist, or another road user. An object's track is the successive positions of the nodes of this object at each video frame. Positions are given as x and y coordinates in the local world coordinate system. The algorithm estimates the PET between each node in the interacting vehicle and each node in the interacting cyclist. The critical PET is then found and adjusted based on the vehicle's direction of movement. Following is a textual representation of the algorithm.

Input:

- 1- Tracks of cyclist's nodes: BR , B and BL nodes (Figure 4-9b).
- 2- Tracks of vehicle's nodes: TR , R , BR , B , BL , L , TL and T nodes (Figure 4-9c).
- 3- Frame rate.

Output:

- PET in seconds.

Procedure:

- 1- Form conflict node pairs: a node pair refers to a possible combination of one vehicle node and one cyclist node for PET measurement. Given the eight nodes defining the vehicle's geometry and three nodes defining the cyclist's geometry,

there are 24 possible conflict node pairs between each interacting vehicle and cyclist.

- 2- For each of the 24 conflict node pairs a PET value is estimated:
 - Calculate the distance from each cyclist node position to each vehicle node position in their respective tracks. These positions would typically correspond to different video frames.
 - Determine the shortest distance out of all calculated distances for the conflict node pair.
 - Determine the two frame numbers corresponding to the positions of the vehicle node and cyclist node at the shortest distance. Those two frames, *vehicle conflict frame* and *cyclist conflict frame*, are collectively called *conflict frames*.
 - Determine PET for this conflict node pair, which will be equal to the difference between the two frame numbers divided by the frame rate.
- 3- Find the critical conflict node pair and initial PET: Select the conflict node pair with the lowest PET out of all 24 values corresponding to the 24 possible conflict node pairs. This conflict node pair would represent the critical node pair and the corresponding PET value would be the initial PET for this vehicle and cyclist.
- 4- Adjust the initial PET depending on the vehicle's direction at the *vehicle conflict frame*.
 - Determine the change in the x and y coordinates for each feature t in the vehicle between two frames, the *vehicle conflict frame* and the previous vehicle frame. This change will be referred to as dxt and $d yt$.

- Determine the change in the x and y coordinates for the whole vehicle (dx and dy) as the average change for all features that are tracked in the vehicle between two frames, the vehicle *conflict frame* and the previous vehicle frame.
- Determine the direction of the vehicle movement (θ , at the vehicle *conflict frame*, using the values of dx and dy . The direction is the angle between the vehicle path respect to the x -axis calculated as: $\theta = \tan^{-1} (|dy|/|dx|)$.

Depending on the direction of vehicle movement, apply an adjustment to the initial PET as explained later.

4.3.2 Detailed procedures in the PET measurement algorithm

Finding the Initial PET Position:

As shown in the previous section, an initial value of PET is first determined as the critical or lowest PET value for all possible conflict node pairs between the interacting vehicle and cyclist. Using the vehicle's eight nodes (TR , R , BR , B , BL , L , TL and T) and the cyclist's three nodes (BR , B and BL) as shown earlier in Figure 4-9, 24 possible conflict pairs are formed combining each vehicle node against each cyclist node. For each conflict node pair, a value of PET is calculated.

In real life, the conflict would happen at the point where the paths of vehicle and cyclist intersect with each other. However, since the video recording is a discrete sequence of images, both the vehicle and cyclist paths are discrete sequences of positions that may not exactly coincide on each other. Therefore, the procedure for calculating the initial value of PET for a specific conflict pair depends on finding the frames corresponding to the shortest distance between the vehicle and cyclist nodes in the pair.

First, the sequence of vehicle positions h is defined by the sequences of positions for all eight vehicle nodes. Mathematically, the sequence can be defined as:

$$h_{i,f} = (x_{i,f}, y_{i,f}) \forall i \in \{1,2, \dots, 8\}; f \in \{n_1, n_2\} \quad (4.1)$$

where n_1 and n_2 = first and last frames between which the vehicle is automatically tracked, $x_{i,f}$ and $y_{i,f}$ = the position of node i in frame f defined by the x and y coordinates on the local world coordinate system.

Similarly, the sequence of cyclist positions c is defined by sequence of positions for all three cyclist nodes. Mathematically, the sequence can be defined as:

$$c_{j,m} = (x_{j,m}, y_{j,m}) \forall j \in \{1, \dots, 3\}; m \in \{n_3, n_4\} \quad (4.2)$$

where n_3 and n_4 = first and last frames between which the cyclist is automatically tracked; $x_{j,m}$ and $y_{j,m}$ = the position of node j in frame m defined by the x and y coordinates on the local world coordinate system.

Having defined the sequences of vehicle and cyclist positions, the procedure to compute the initial value of PET involves finding the *conflict frames*, *vehicle's conflict frame* and *cyclist's conflict frame*, within which the interaction event takes place. At those two frames, the Euclidean distance between the interacting motor vehicle position and cyclist position is the shortest. Therefore, in the AM method, the area of potential collision for every pair can be identified by the shortest distance between this pair of nodes. For example, the *vehicle's conflict frame* (f^*) and *cyclist's conflict frame* (m^*) for nodes i and j are found by finding the shortest distance between the two nodes for all vehicle frames f and all cyclist frames m using the following formula:

$$\text{shortest distance} = \|h_{i,f^*} - c_{j,m^*}\| = \min_f \{ \min_m \{ \|h_{i,f} - c_{j,m}\| \} \} \quad (4.3)$$

Figure 4-10 shows a schematic diagram for a specific conflict node pair and the distances between the vehicle node at one position and the different positions of the cyclist node. The shortest distance for this conflict pair is also shown, which defines the location of the area of potential collision. The number of frames which separates the arrivals of these two nodes, i and j , at the area of potential collision represents the duration that separates the passages of the vehicle and the cyclist at positions h_{i,f^*} and c_{j,m^*} . For example, Figure 4-10 shows that frame number 156 corresponds to the cyclist node frame number at the shortest distance (m^*), while frame number 177 corresponds to the vehicle node frame number (f^*). Therefore; the corresponding PET value for this node pair is calculated using the following formula:

$$PET = \frac{|f^* - m^*|}{\text{frame rate}} \quad (4.4)$$

As mentioned in the algorithm, the critical conflict node pair is the one that has the shortest PET out of the 24 values for all conflict node pairs. This shortest PET value is the initial PET for this vehicle-cyclist interaction.

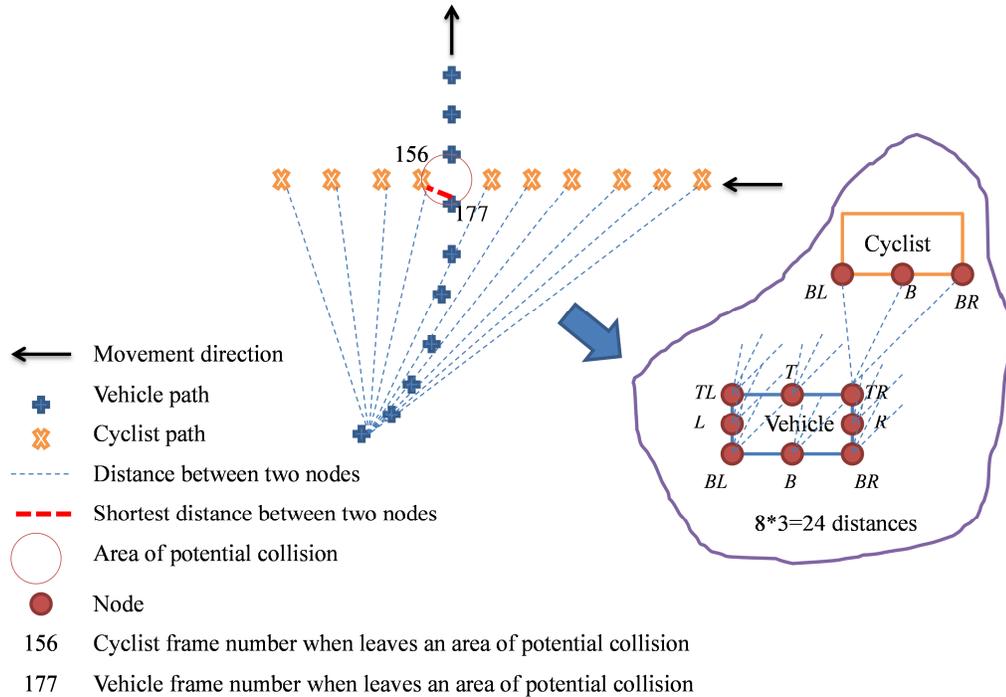


Figure 4-10: Schematic of the distance between different positions of vehicle's nodes and different positions of cyclist's nodes.

4.3.3 Adjusting the initial PET

The initial PET computed as previously described is based on the assumption that the cyclist path is conflicting with one of the eight nodes representing the bounding box of the vehicle. However, close examination of the observed interactions showed that the correct part of the interacting motor vehicle, which enters the area of potential collision first or leaves it last, can be between two nodes. Thus, an automated technique was developed to adjust the value of PET based on the vehicle direction of movement. The direction is defined in the two-dimensional (X, Y) local world-space based on the change of the coordinates of the vehicle's features. Using the coordinates (x_t, y_t) of each feature t ,

which are obtained using the tracker system, the change of coordinates (d_{xt} , d_{yt}) is calculated as follows:

$$d_{xt} = x_{t,i+1} - x_{t,i}; d_{yt} = y_{t,i+1} - y_{t,i} \quad (4.5)$$

The average change in direction (dx , dy) is then calculated for all features as follows:

$$d_x = \frac{\sum_t d_{xt}}{n_t}; d_y = \frac{\sum_t d_{yt}}{n_t} \quad (4.6)$$

where n_t = number of vehicle's features.

As shown in Figure 4-11, the direction of vehicle movement is defined using the angle θ , which is calculated as follows.

$$\theta = \tan^{-1} \left| \frac{d_y}{d_x} \right| \quad (4.7)$$

Because the angle θ will always be in the range $[0, 90^\circ]$, another piece of information that is required to define the vehicle's direction of movement is the quadrant for angle θ . Depending on the signs of (d_x , d_y), the movement direction would correspond to one of four possible quadrants as shown in Table 4-2. Several interaction events were observed and many trials were made to identify the correct part of the vehicle at the potential interaction area and relate it to the direction of the vehicle. Depending on the direction of vehicle movement, PET can be adjusted to represent the correct conflict part of the motor vehicle at the area of potential collision as shown in Table 4-2.

Table 4-2: Directions of vehicle movement and corresponding adjustments to initial PET.

<i>Sign of d_y, d_x</i>	<i>Bounds</i>	<i>Adjusted PET (vehicle arrives an area of potential collision second)</i>	<i>Adjusted PET (vehicle arrives an area of potential collision first)</i>
$d_x > 0, d_y > 0$	$0.0 \leq \theta < 75.0$	$(\text{PET at } B + \text{PET at } BL) / 2$	$(\text{PET at } T + \text{PET at } TL) / 2$
	$75.0 \leq \theta \leq 90.0$	$(\text{PET at } T + \text{PET at } TL) / 2$	$(\text{PET at } B + \text{PET at } BR) / 2$
$d_x > 0, d_y < 0$	$0.0 \leq \theta < 90.0$	$(\text{PET at } T + \text{PET at } TL) / 2$	$(\text{PET at } R + \text{PET at } BR) / 2$
$d_x < 0, d_y < 0$	$0.0 \leq \theta < 12.0$	PET at R	PET at L
	$12.0 \leq \theta < 23.0$	PET at TR	PET at BL
	$23.0 \leq \theta < 65.0$	$(\text{PET at } T + \text{PET at } TR) / 2$	$(\text{PET at } B + \text{PET at } BL) / 2$
	$65.0 \leq \theta < 80.0$	PET at T	PET at B
	$80.0 \leq \theta \leq 90.0$	PET at TL	PET at BR
$d_x < 0, d_y > 0$	$0.0 \leq \theta < 40.0$	PET at BR	$(\text{PET at } L + \text{PET at } TL) / 2$
	$40.0 \leq \theta \leq 90.0$	$(\text{PET at } B + \text{PET at } BR) / 2$	$(\text{PET at } L + \text{PET at } TL) / 2$

An example of this procedure is shown in Figure 4-11, where the node *TL* was selected as the initial critical conflict node. However, as shown in the figure, this node does not represent the correct conflicting part of the interacting motor vehicle. Rather, it is point *y* that refers to the right conflicting part of the vehicle as it is the first point in the vehicle to cross the cyclist's path. Therefore, the value of PET should be adjusted based on the relative location of *y*. In this case the adjusted PET is evaluated as the average of PET for node *TL* and the PET for node *T*. This will give an equivalent point of *y* at the midpoint between *TL* and *T*, as denoted point \bar{y} in Figure 4-11(b).

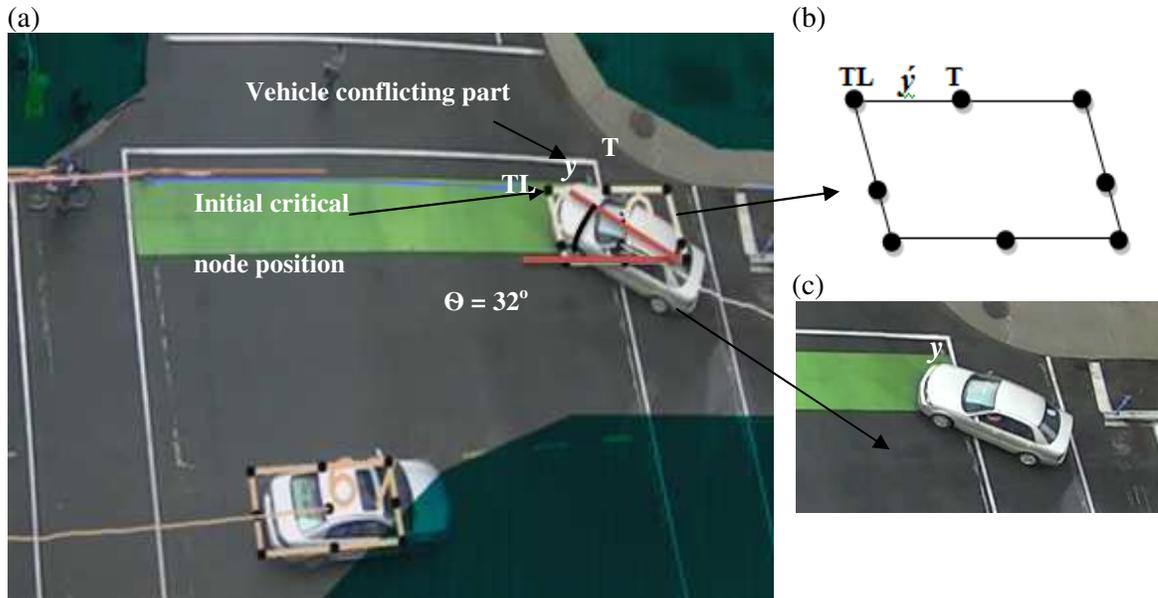


Figure 4-11: Example critical conflict node and correct conflicting part of an interacting vehicle. (a) motor vehicle direction. (b) equivalent point of “y”, as denoted point “y”. (c) right conflicting part of the interacting motor vehicle.

4.4 Results and Discussions

4.4.1 Descriptive statistics of PET

A total of 384 interaction events were identified manually after reviewing a total of 1,901 through cyclist movements. All these events were right-turn and left-turn interactions with PET values less than or equal to 3.0 seconds as determined using the MVTM method. This threshold is used only to select events for analysis. However, as mentioned earlier, this study does not attempt to establish a PET threshold for interactions. On the other hand, there were no interactions observed between left-turn cyclists and through motor vehicles. The reason for this was that cyclists used a bike box

to perform the left turn manoeuvres, thus dividing the maneuver into two crossing actions. The analysis results of PET using the three methods are shown in Table 4-3.

Table 4-3: Analysis results of PET using MVTM, MFCM and AM methods.

Analysis PET method	Precision second	PET Min./Max second	Average PET second	Standard deviation second
MVTM	1.00	1.00/3.00	2.53	0.64
MFCM	0.033	0.07/3.00	1.77	0.60
AM	0.033	0.03/3.27	1.83	0.60

Given the relatively large sample in this study, it was possible to compare different distributions to characterize the measured PET values using different measurement methods. In addition to knowing the mean and standard deviation, the whole distribution of any parameter can be useful in reliability analysis, statistical analysis of a phenomenon including future studies of potential relationship between collision frequency and interaction frequency. Different candidate distributions were examined based on the Log-likelihood. For example, Normal, Gamma, and Lognormal distributions are commonly used models to fit positive observations in statistical analysis. Depending on the lowest value of the Log-likelihood, the distribution is identified. Table 4-4 shows the Log-likelihood for each candidate distribution. As shown in the Table 4-4, the best distribution was found to be the normal distribution.

Table 4-4: Fit distribution of the PET measurement methods.

Distributions	Log likelihood		
	MVTM	MFCM	AM
Lognormal	-443.98	-404.87	-411.66
Gamma	-415.15	-368.60	-369.12
Normal	-371.46	-345.37	-348.52

4.4.2 Validation of AM method

Given the precision of these methods, both MFCM and AM methods are expected to be more accurate than MVTM method. On the other hand, both MFCM and MVTM methods are not required to be validated as opposed to the AM method. Therefore, the three methods are compared in this following subsections for the purposes of validating the AM method by comparing it to the MFCM method and evaluating reliability of the MVTM method by comparing it to the MFCM and AM methods.

Since both MFCM and AM methods estimate PET with the same precision of 0.033 seconds, the estimated values of PET in both methods can be compared to each other. Figure 4-12 shows the relationship between the individual PET measurements, which are the estimated PET values for the 384 observed interaction events, using MFCM and AM. As shown in the figure, the line of equality provides a very good fit for the two sets of measurement, with a high coefficient of determination, R^2 , of 0.94. Furthermore, the points seem to be closely and consistently scattered around the line of equality which suggests that the AM method does not tend to overestimate or underestimate the PET value relative to MFCM method.

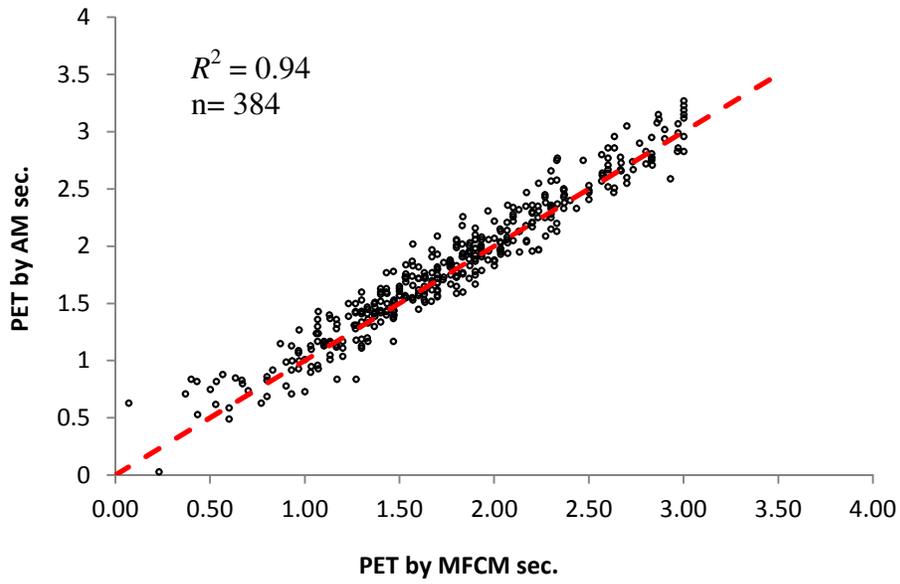


Figure 4-12: Scatter plot of PET measurements using the AM and MFCM methods.

The average value of the difference in PET between AM and MFCM was found to be -0.06 second with the minimum and maximum ones of -0.56 and 0.43 seconds. The Shapiro-Wilk test was used to test normality. The null hypothesis (H_0) is defined as H_0 : the observed differences in PET from the two methods come from a normally distributed population. The p -value was found to be 0.493 (>0.05) and therefore the null hypothesis cannot be rejected. This provides an evidence of the normality of the differences between AM and MFCM. With the mean value of the difference almost equal to zero, it can be stated that the AM method produces accurate, unbiased estimates of PET. This statement can be supported by noting that the average absolute difference was 0.12 second and the standard deviation was 0.10 second.

4.4.3 Accuracy of MVTM method

As the PET values in the MVTM method are obtained using a classification scheme based on one-second increment of video time, a direct comparison between MVTM with

AM or MFCM methods would not yield appropriate conclusions. The PET values resulting from the AM method were compared to the MFCM results for three interaction levels corresponding to PET values of (0, 1], (1, 2] and (2, 3] seconds as estimated using the MFCM method. The frequency of estimating of the interaction levels using MVTM, MFCM and AM methods are shown in Table 4-5. Kappa coefficient (K) was used to evaluate the level of agreement between the three methods of PET measurement by comparing two methods at a time. Kappa coefficient is defined as “*a statistical measure of the inter-agreement or inter-annotator agreement*” (Cohen, 1960), and is calculated for any two measurement methods as follows:

$$K = \frac{P_a - P_e}{1 - P_e} \quad (4.8)$$

where K = Kappa coefficient, P_a = Proportion of observed agreement where both measurement techniques classify an event as PET (0, 1], (1, 2] or (2, 3] seconds. P_e = Proportion of expected agreement.

Table 4-5: Frequency of the PET levels by MVTM, MFCM and AM methods.

Analysis PET method	(0, 1] second	(1, 2] second	(2, 3] second	Total
MVTM	30	120	234	384
MFCM	35	224	125	384
AM	35	212	126	373

Depending on the K value between two sets of observation, different levels of agreement between the two sets can exist (Viera & Garrett, 2005): [i] “*almost perfect agreement*” for K value between 0.81 and 0.99, [ii] “*substantial agreement*” for K value between 0.61 and 0.80, [iii] “*moderate agreement*” for K value between 0.41 and 0.60,

[iv] “*fair agreement*” for K value between 0.21 and 0.40. Table 4-6 shows the agreement results of PET values of (0, 1], (1, 2] and (2, 3] seconds as obtained by comparing the AM versus MFCM and MVTM methods, and MFCM versus MVTM methods.

Table 4-6: Agreement evaluation using kappa values under the PET levels of (0, 1], (1, 2] and (2, 3] seconds.

Measurement Methods	Kappa coefficient (K)		
	PET (0, 1] second	PET (1, 2] seconds	PET (2, 3] seconds
AM versus MFCM	0.81	0.76	0.81
AM versus MVTM	0.52	0.22	0.32
MFCM versus MVTM	0.47	0.20	0.31

As shown in Table 4-6, the Kappa results show that the agreement between the AM and MFCM methods is *almost perfect agreement* under the interaction level of PET of (2, 3] and (0, 1] seconds. The agreement between the two methods under the interaction level of PET of (1, 2] seconds was found to be *substantial agreement*. Moderate agreement was found between the MVTM and both MA and MFCM under the interaction level of PET of (0, 1] second. A *fair agreement* was found between the MVTM and both MA and MFCM under the severity levels of interaction for PET (1, 2] and (2, 3] seconds.

The agreement analysis is another evidence of the reliability of the AM method as compared to MFCM. On the other hand, the agreement analysis shows low reliability in using the MVTM method to measure PET.

4.5 Summary

Interactions between motor vehicle and cyclists at a signalized intersection were characterized in this study using an objective conflict indicator; Post-Encroachment Time (PET). This chapter presented an automated video analysis technique to measure PET between cyclists and motor vehicles. The developed automated method (AM) was compared to highly reliable manual frame count measurement (MFCM) method using video data on two signalized intersections in the City of Ottawa, Ontario.

Three measurement methods were used to measure PET: [i] manual video timer measurement (MVTM) method, [ii] manual frame count measurement (MFCM) method and [iii] automated measurement (AM) method. A total of 384 interaction events of PET (0, 3] seconds between cyclists and vehicles were analyzed in this study. The results showed that the average PET value using the MVTM method was 2.53 seconds with a standard deviation of 0.64 second. The average PET value using MFCM method was 1.77 seconds with a standard deviation of 0.60 second. The average PET value using the AM method was 1.83 seconds with a standard deviation of 0.60 second. Depending on the lowest value of the Log-likelihood, the Generalized Extreme Value was identified as the best-fit distribution of measured PET values for both MFCM and AM methods.

The AM method was evaluated based on the difference between its measurements and those obtained from the MFCM method. The results showed that the PET values obtained from both methods were very close. A high coefficient of determination, R^2 , was found between the AM and MFCM methods. The average value of the difference (signed) in PET was -0.06 second over a range of PET difference from -0.56 to 0.43 seconds. The average absolute difference (unsigned) between AM and MFCM methods

was 0.12 second. This is a very promising finding given that the precision is only 0.033 second. Furthermore, there was a very good agreement between AM and MFCM methods in PET count at three different severity levels: levels of (0, 1], (1, 2] and (2, 3] seconds.

It can be concluded that the AM method was proven to be accurate and reliable as demonstrated by comparing the results obtained using three different measurement methods. This technique has a promising practical benefit as it can be used to measure the PET with high accuracy while taking into account the physical extremities (rear and front parts) of the analyzed road users. The proposed technique also objectively defines the potential area of collision as the point where the paths of the interacting vehicle and cyclist intersect. Thus, it removes the need to subjectively define the potential area of collision or visually determine the location where the paths of the cyclist and the vehicle intersect for the purpose of measuring PET. The proposed technique requires very little human intervention and therefore can substantially help users save time and effort and reduce human errors. Overall, the time difference required for calculating the PET for processing all events using MFCM and AM methods in this study is estimated to be (18 hours). Therefore, the process time is decreased by 67% of the time when using the AM method. Practically, the proposed technique can support large-scale data collection for investigating the relationship between PET counts and collisions.

Different shortcomings occurred while measuring the PET between cyclists and motor vehicles using the tracker system. For example, there were some bounding boxes which partially covered road users. However, in case of partial coverage, the part of road users covered by the bounding box is used to calculate PET and this may explain why there is the residual error in an automated PET calculation. Although the tracker system

is improved and most trajectories are detected by the tracker system, over-segmentation still occurred especially on long vehicles. Over-segmentation, which means that more than one object can be seen representing real road users. In addition, other factors were found to impact the quality of the tracks, such as, shades and other disruptive textures.

The presented techniques can benefit from automated road user classification (cyclist or vehicle) thus improving the level of automation. Furthermore, the presented techniques offer the potential and flexibility to be expanded to consider other road users.

5. DEVELOPMENT AND APPLICATION OF STATISTICAL MODELING TO MEASURE A CYCLIST SAFETY

The goal of this chapter is to develop and test statistical models which can be used to measure the cyclist safety. In particular, the main task concerned in this chapter is proposing a scheme for ranking of intersections based on expected safety. Achieving this task using collision data can be very challenging due to inherent quantitative and qualitative issues in collision data as discussed in Chapter Two. This chapter provides statistical models which can assist in achieving this task using non-collision, or surrogate measures of safety. The surrogate measures of safety used in this chapter is PET measurement, as described in Chapter Four, and non-conformal traffic behaviour represented by traffic violations.

The data used in this chapter was collected from a total of nine intersections located at the downtown of the City of Ottawa. The focus of the data collection is to observe cyclist and vehicle behaviour at the extension of bike lanes through signalized intersections. The two different types of bike lanes leading to these intersections are segregated bike lanes (seven intersections) and marked bike lanes (two intersections). For objective comparison, a number of performance parameters were defined in this chapter and observed at these traffic facilities. The performance parameters can be grouped into two main categories:

1. Count of events involving close proximity between cyclists and motor vehicles.
2. Non-conformal behaviour committed by cyclists, and non-conformal behaviour committed by vehicles.

This chapter is organized in three main sections. In the first section, the types of data collected and the involved process are described. The chapter then proceeds to describe the surrogate safety measures as well as the observation methodology in the second section. Third section includes a descriptive and comparative data statistical analysis.

5.1 Data Collection

5.1.1 Site Description

The study area in this chapter consists of intersections in the downtown of the City of Ottawa. These intersections have bike lanes leading to their major approaches. These intersections are divided into two main groups depending on the type of bike lane. The first group of intersections have segregated bike lanes leading to their two major approaches. The second group of intersections has marked bike lanes leading to their major approaches as shown in Figure 5-1. The first group (segregated) consists of seven intersections that are located along Laurier Avenue (major street) in downtown Ottawa as shown in Figure 5-1 (Google, 2013). Laurier Avenue is the site of extended segregated bike lanes intended to channelize the cyclist traffic along this street and separate cyclists from motor vehicles. Six of these seven intersections are signalized and the seventh intersection is uncontrolled T-shaped intersection. Figure 5-2 shows a photo and layout of one of these intersections. The layout of the intersection was based on the site surveying by the thesis author. The other intersections are shown in Appendix (C). As shorthand, the first group intersections with segregated bike lanes are called treatment sites or intersections and the second group with marked bike lanes are called control sites or

intersections. The terminology is used because the second group of intersections was selected to represent the first group of intersections before segregation.

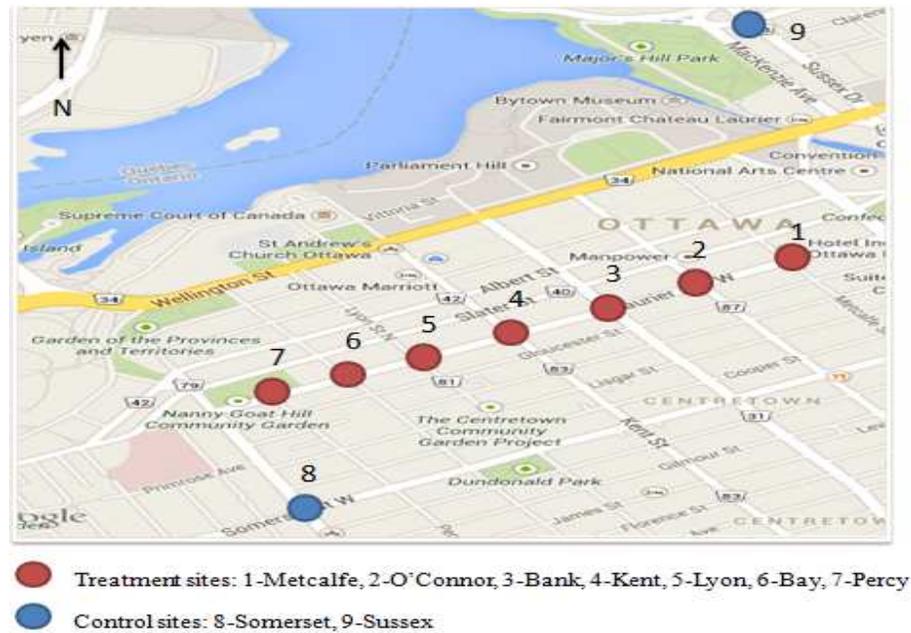


Figure 5-1: Location of the study intersections (Google, 2013).

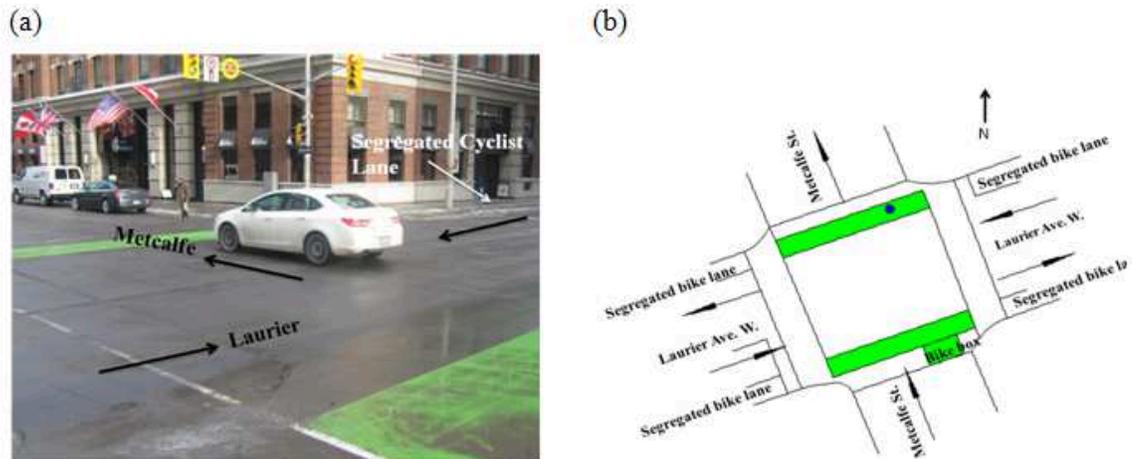


Figure 5-2: Laurier- Metcalfe signalized intersections. (a) photo taken at the intersection. (b) AutoCAD intersections drawing.

As shown in Figure 5-2, Laurier Avenue, which is a two-lane two-way street with a segregated cyclist lane on each side of the street, is the common street in all treatment

intersections. All intersections are signalized four-leg except for one intersection, which is un-signalized T-shaped (three-leg) intersection. All intersecting streets (cross-streets) are one-way street except for one (Bank street) which is a two-way cross-street. The allowable movements within the intersections for cyclists are through (move forward within the intersection) for both directions and right-turn in the appropriate direction turning into the cross-street traffic direction. Left turns for cyclists in the relevant direction are allowed by using a marked turning location (bike box) and they are allowed to cross the intersection and complete the left-turn movement during the pedestrian signal.

The second group of intersections (control) is located within the downtown area as shown in Figure 5-3 (Google, 2013). They can provide estimate of the conditions in the treatment intersections before segregation. Both control intersections, shown in Figure 5-4 and Figure 5-5 is signalized intersections.

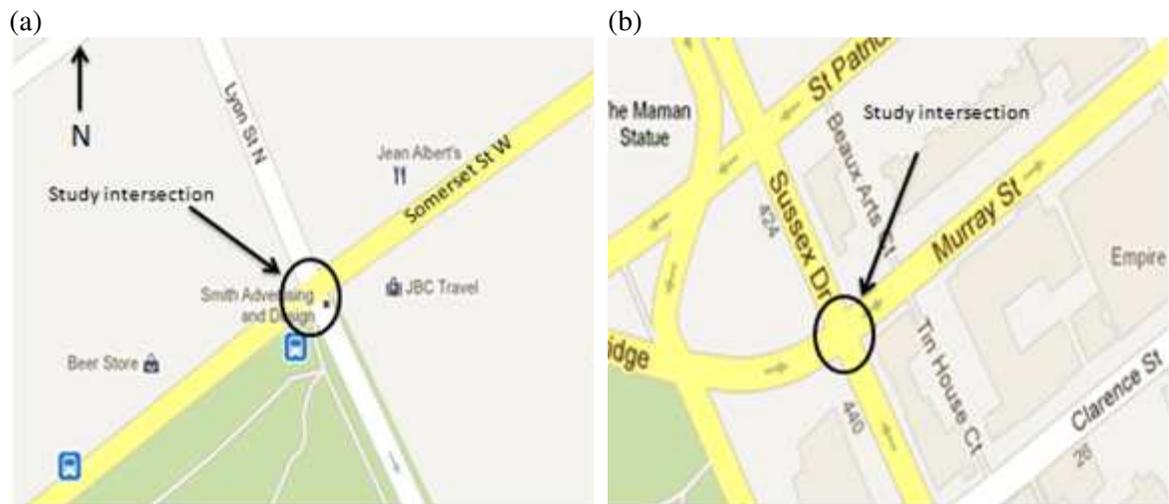


Figure 5-3: Location of the control (un-segregated) signalized intersections. (a) Lyon- Somerset intersection. (b) Murray-Sussex intersection (Google, 2013).

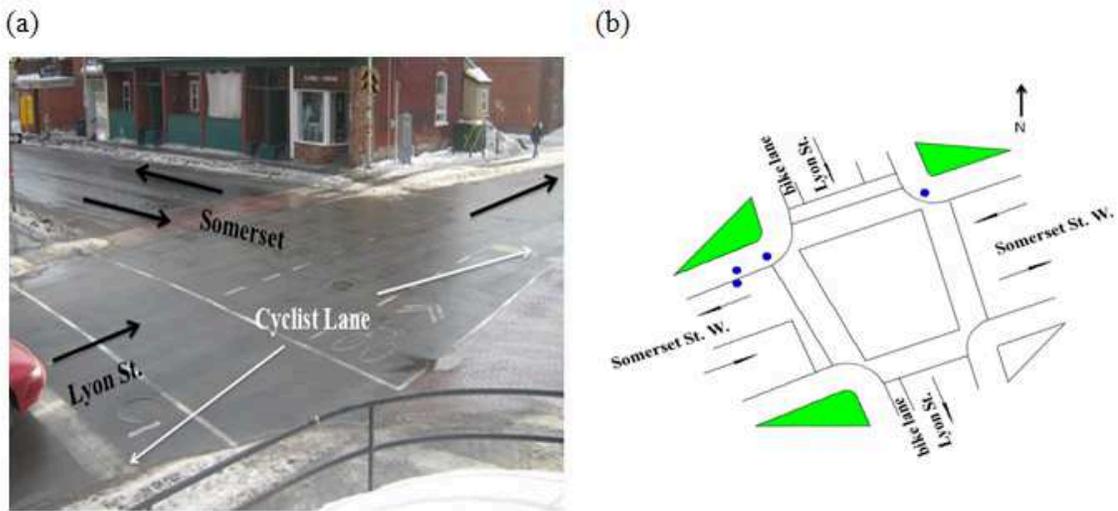


Figure 5-4: Lyon- Somerset signaled intersection. (a) photo taken at the intersection. (b) AutoCAD intersection drawing.

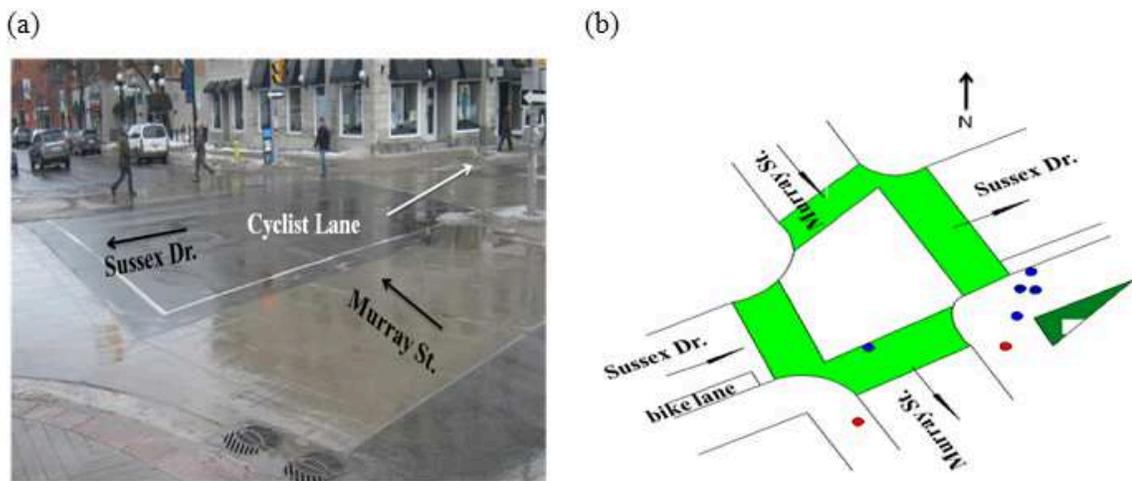


Figure 5-5: Murray- Sussex signaled intersection. (a) photo taken at the intersection. (b) AutoCAD intersection drawing.

The first control intersection; Somerset, which is a two-lane two-way street, intersecting with Lyon, which is a three-lane one-way street, has a marked bike lane on one side of Lyon street. Both Sussex and Murray Streets are two-lane one-way street. As shown in the previous figure the control cyclist lane is located on one side of Sussex Street.

5.1.2 Video Data Collection Equipment

In order to monitor cyclists and motor vehicles movement at all intersections, video data was collected from different sources: cameras operated by Carleton University (video data collection from 15th – 21th floor of a building), mobile camera (video data collection which can be erected from 3.5 – 5.0 meter height) and stationary camera (video data collection 6.0 – 7.5 meter altitude). Both mobile and stationary camera units operated by the City of Ottawa as shown in Figures 5-6, 5-7 and 5-8.



Figure 5-6: A high definition camera (image resolution is 1920 x 1080).



Figure 5-7: Mobile camera (image resolution is 320 x 240).



Figure 5-8: Stationary camera (image resolution is 640 x 480).

5.1.3 Video Data Collection

Video data was collected at the intersections in the study area in order to observe movements of road users within these intersections. The video data covers a common observation period from 8:30 AM to 4:00 PM for all intersections. Whenever possible, data collection was performed in two seasons (Fall of 2011 and Summer of 2012) and observation period was extended before and after this common window. Table 5-1 shows a summary of the data collection duration and locations. As shown in Table 5-1, video data collected at O'Connor intersection in the fall season showed a limited view which was not consistent with other intersections. Therefore, performance parameters related to vehicular observations for this intersection were excluded from the study. Bank and Sussex intersections video data was not collected in the fall season due to difficulties in installing the camera.

As shown in Table 5-1, there is no vehicular left-turn movement at Bank, Somerset and Sussex intersections. Therefore, only intersections with common turn movements are included in the analysis of intersections.

Table 5-1: Video data collection during Fall -2011 and Summer -2012 seasons.

Intersection id	Intersection name	Shorthand intersection name	Number of days		Approximate duration (hours)		Total cyclist volume (through, right and left)		Total vehicle volume		Vehicle turning movements	
			fall	summer	fall	summer	fall	summer	fall	summer	right	left
1	Metcalf & Laurier Westbound	Metcalf	2	2	19	26	417	1,511	1,813	4,039	✓	✓
2	Kent and Laurier Westbound	Kent	2	4	35	44	1,078	2,649	2,772	5,851	✓	✓
3	Lyon and Laurier Eastbound	Lyon	5	3	66	28	664	1,034	5,606	3,655	✓	✓
4	Bay and Laurier Westbound	Bay	4	2	62	24	761	1,040	5,848	3,105	✓	✓
5	Percy and Laurier Eastbound	Percy	5	2	50	32	337	535	1,195	924	✓	✓
6	Somerset and Lyon Southbound	Somerset	2	1	35	13	760	406	1,804	745	✓	✗
7	O'Connor and Laurier Eastbound	O'Connor	4	2	78	26	1,173	1,918	-	5,570	✓	✓
8	Bank and Laurier Eastbound and	Bank	-	2	-	52	-	2,958	-	2,259	✓	✗
9	Sussex and Murray Northbound	Sussex	-	2	-	26	-	1,111	-	3,446	✓	✗

5.2 Performance Parameters

5.2.1 Background

A number of performance parameters were identified in this chapter. The performance parameters are related to interactions between cyclists and vehicles and individual cyclist behaviour. The main interaction pattern is the temporal proximity of cyclist and right/left-turn vehicles while crossing each other's paths. A number of categories of non-conformal traffic movements were identified and observed for cyclists at treatment sites and control. The following sections provide a description of these performance parameters.

5.2.2 Movements

Cyclists crossing the treatment and control intersections were observed at the corners of these intersections where vehicular right-turn and left-turn movements were permitted. Movements of cyclists were categorized into left-turn, through, and right-turn. Movements of vehicles were categorized into through, right-turn, and left-turn. Only permitted right-turn and left-turn vehicle movements were observed. Right-turn movements on red in the control intersections were not included in order to conduct a valid comparison with the treatment intersections where right-turn on green only.

5.2.3 Surrogate Measures of Safety

The concept of surrogate measures of safety aims at developing safety indicators which are related to safety but are not based on collision observations (Hydén, 1987). The close-proximity interactions or surrogate events were observed at the treatment and control intersections. These surrogate events, or interactions, can reflect safety

shortcomings in the form of near-missed but are not themselves collisions. They may be placed near the top of the safety hierarchy shown in Figure 5-9.

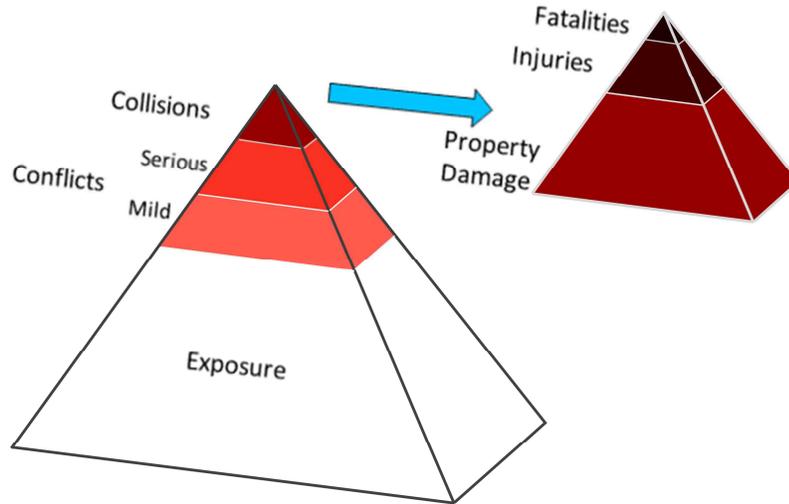


Figure 5-9: Illustration of the safety hierarchy (reproduced from (Hydén, 1987)).

In this chapter, a close-proximity interaction between a cyclist and a vehicle is identified based on measurement of Post Encroachment Time (PET). The measurement of PET used in this chapter is described in detail in Chapter Four. PET can be related to the observation of near-misses or critical proximity of road users while traversing an intersection. The shorter the PET, the closer the road users were to collision, with a value of 0 denoting a collision. PET is an objective measure of proximity which does not rely on extrapolation of road users' positions and is based on actual observed positions.

As mentioned in Chapter Four, PET is defined as the time difference between the moment when a rear part of the vehicle leaves (or a front part arrives) at the area of potential collision and the moment when a bicycle's front wheel arrives at (or the rear wheel departs from) this area. Depending on the movements of both the cyclist and motor

vehicle, different types of crossing actions may occur between the paths of the vehicle and the cyclist as described in Chapter Four.

5.2.4 Observation Methodology

The observation methodology consists of reviewing of videos and identification of events which involve the presence of cyclists in the intersection. These records of events are then assigned various attributes, depending on the occurrence of any anticipated behaviour. The observation matrix is composed of a set of rows, each one representing an event, and a set of columns, each one representing a particular observation (numeric value) or an attribute (0 or 1) based on whether the occurrence represented by this attribute was observed as shown in Figure 5-10. Following is a description of the attributes observed for each event.

Cyclist number	Start			Duration			Movement direction		Interaction types	PET			Non- conformal movements or violations				
	hh:mm:ss	hh:mm:ss	hh:mm:ss	cyclist		vehicle		by cyclist		by vehicle							
				Left					Right-turning				Against traffic				
				Through					Left-turning	(0, 3] second			Stopping a after stop-line				
				Right						(0, 2] second			Riding in a cross walk				
				Left						(0, 1] second			Out-of-path stop				
				Through									Riding out-of-path				
				Right									Cross stop-line on red signal				
													Veh. turns left and stops on cyclist path				
													Veh. turns right and stops on cyclist path				
													veh. Fails to yield the cyclist				

Figure 5-10: Observation methodology attributes.

Cyclist number: Number of cyclists involved in the event.

Start: The time at which cyclist starts crossing the intersection in hh:mm:ss.

End: The time at which cyclist ends crossing the intersection in hh:mm:ss.

Duration: The duration of the cyclist crossing the intersection.

Movement direction of cyclist: Direction was classified into left, through and right.

Movement direction of vehicle: Direction was classified into right-turn, through, and left-turn.

Quantifying Proximity of Cyclist and Vehicles:

Interaction detection and severity rating is based on the value of PET between the cyclist and vehicle. PET values (0, 3] seconds were considered to expose the cyclist to an elevated level of collision threat as described in Chapter Four. The interactions are further rated along three thresholds of PET within the range from 0.0 second to 3.0 seconds at increments of 1.0 second.

Exclusion of PET measurement:

Certain events were excluded from PET observations based on any of the following conditions:

- 1- Cyclist crosses the intersection with no interference with any vehicles (intersection is unoccupied).
- 2- Cyclist crosses the intersection with measured PET greater than 3.0 seconds.

PET measurement for interactions:

- 1- Cyclist exposed to one vehicle: two types of interactions may happen between the cyclist and right-turn or left-turn vehicles when the interaction time within 3.0 seconds.

2- Cyclist exposed to multiple vehicles:

- Two types of interactions may occur between the cyclist and two vehicles when the interaction time within 3.0 seconds. The occurrence of interaction does not depend on whether a cyclist/vehicle arrives first to the area of potential collision.

- Two interactions could occur with right-turn and one cyclist. When the rear of the vehicle leaves an area of potential collision and a cyclist arrives to this area within 3.0 seconds. This situation covers one interaction. When the cyclist leaves this area of potential collision, at that moment, the front of the second vehicle will arrive to a different area of potential collision and this will occur within 3.0 seconds. This is considered as a second interaction.

- Two interactions could occur with left-turn vehicles and one cyclist. When the rear of the vehicle leaves an area of potential collision and a cyclist arrives to this area within 3.0 seconds. This situation covers one interaction. When the cyclist leaves this area of potential collision, at that moment, the front of the second vehicle will arrive to a different area of potential collision and this will occur within 3.0 seconds. This is considered a second interaction.

- Three interactions, two Right-turn vehicles and one left-turn vehicle, may occur between three vehicles and a cyclist. When the rear of the right-turn vehicle leaves an area of potential collision and a cyclist arrives to this area within 3.0 seconds. This situation covers one interaction. When the cyclist leaves this area of potential collision, at that moment, the front of the second vehicle will arrive to a different area of potential collision while making a right turn and this will occur within 3.0 seconds with the cyclist. This is considered a second interaction. When the cyclist

leaves the second area of potential collision, at that moment, the front of the third vehicle will arrive, while making a left turn, to another different area of potential collision within 3.0 seconds. This is considered a third interaction.

- 3- Many cyclists in a group could be exposed to a number of interactions depending on the PET between each cyclist with the vehicle. If the PET value within 3.0 seconds, an interaction will be recorded for each cyclist. This case could occur when a group of cyclists is positioned between two vehicles. The rear of the first vehicle leaves an area of potential collision, while making a right turn, before the group of cyclists enters this area of potential collision. After the group of cyclists leave this area of potential collision, at that moment, the front of the second vehicle arrives, while making a right/or left turn, to another different area of potential collision.

Classification of PET cut-off thresholds:

As mentioned earlier, three different types of recording cameras were used to collect the video data. After checking the videos, it was found that, videos collected using the mobile and stationary cameras operated by the City of Ottawa were not appropriate to measure PET between cyclists and motor vehicles using the automated measurement method presented in Chapter Four. This can be attributed to the lower height of these cameras and the video resolution that was not suitable to identify the interaction between road users or to separate road users from each other and track each road user. For example, the low height of camera led to the disappearance of cyclists behind vehicles. This in turn would lead to failure to track the cyclist over the whole subsequence or overgrouping of road users. Therefore, both the Manual Video Timer Measurement (MVTM)

and Manual Frame Count Measurement (MFCM) methods were used in this chapter to accurately estimate the PET for each interaction.

A PET threshold of 3.0 seconds was used to identify events involving a cyclist and a vehicle which will be analyzed further. First, the MVTM method was used to identify the cyclists who were involved in an interaction with a motor vehicle where PET was 3 seconds or less and to identify cyclists who committed a violation. Then, the MFCM method was applied to the same events that were observed in the previous step. The proximity between cyclists and vehicles in these interactions was classified to three different levels based on the PET value. These PET classification levels are (0, 1], (0, 2], and (0, 3] seconds.

Non-conformal movements or violations: Non-conformal events or traffic violations may also be regarded as surrogate measures of safety. In order to expand the scope of this study, the events that listed following this paragraph were observed for cyclists and vehicles in the study area. The following sections describe the type of violations observed in this study.

5.3.5a Cycling against traffic

The cyclist committed a violation by cycling against the flow of traffic. Figure 5-11 shows a sample of this violation. As shown in the figure the cyclist riding the bike against the flow of traffic.



Figure 5-11: Cycling against traffic violation.

5.3.5b Stopping after stop-line

The cyclist committed a violation by stopping after the stop-line on red light. If the front wheel of the bike is outside the cyclist stop-line on a full stop, the cyclist has committed a violation. Anywhere the cyclist stops beyond the cyclist stop line, but still in the cyclist path, is still considered a violation. Figure 5-12 shows a sample of this violation.



Figure 5-12: Cyclists stopping beyond stop-line.

5.3.5c Ride in a crosswalk

Figure 5-13 refers to a cyclist committing a violation by riding in within the crosswalk marking.



Figure 5-13: Cyclists using the crosswalk.

5.3.5d Out-of-path stop

Figure 5-14 refers to two cyclists committing a violation by stopping out of the cyclist path. This could happen if the cyclist stops out of the cyclist path on red light.



Figure 5-14: Cyclists stopping out of path.

5.3.5e Riding out-of-path

Figure 5-15 refers to the cyclist committing a violation by riding out of the cyclist path.



Figure 5-15: Two examples of two cyclists riding out of path.

5.3.5f Cross stop-line on red signal

Figure 5-16 shows a cyclist committing a violation by crossing on a red light. As shown in the figure, the vehicle labelled number 1 stopped, while the vehicle labelled number 2 crossed the intersection from the adjacent approach. This means that the signal turned red in the direction of east to west. In addition, there were two cyclists crossing the intersection from west to east. The cyclist labelled number 1 stopped at the same time as the vehicle labelled number 1. The cyclist labelled number 2 should have stopped (but did not) at the stop line because the traffic signal changed to red.

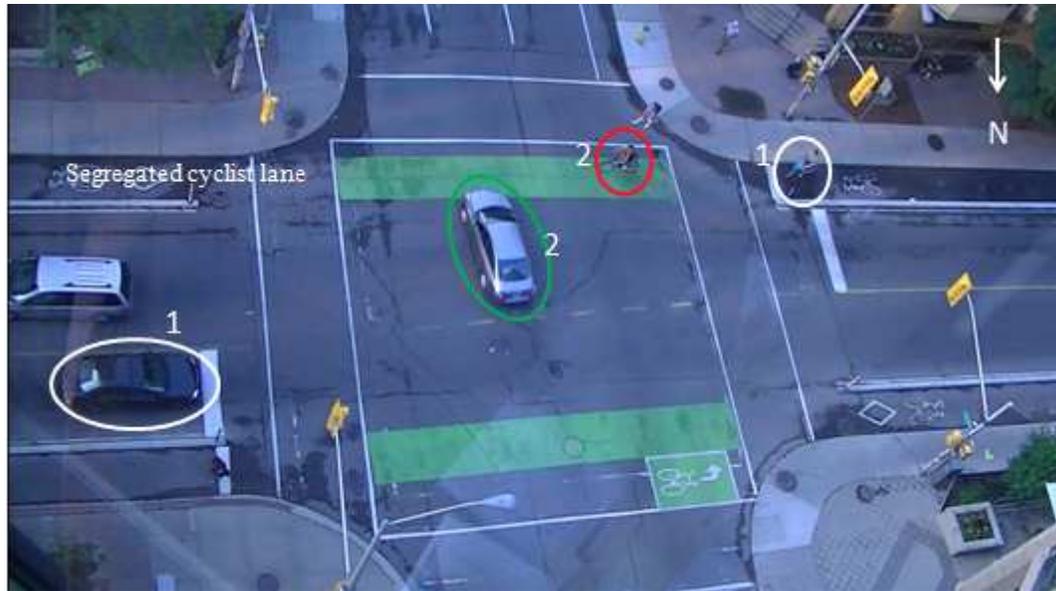
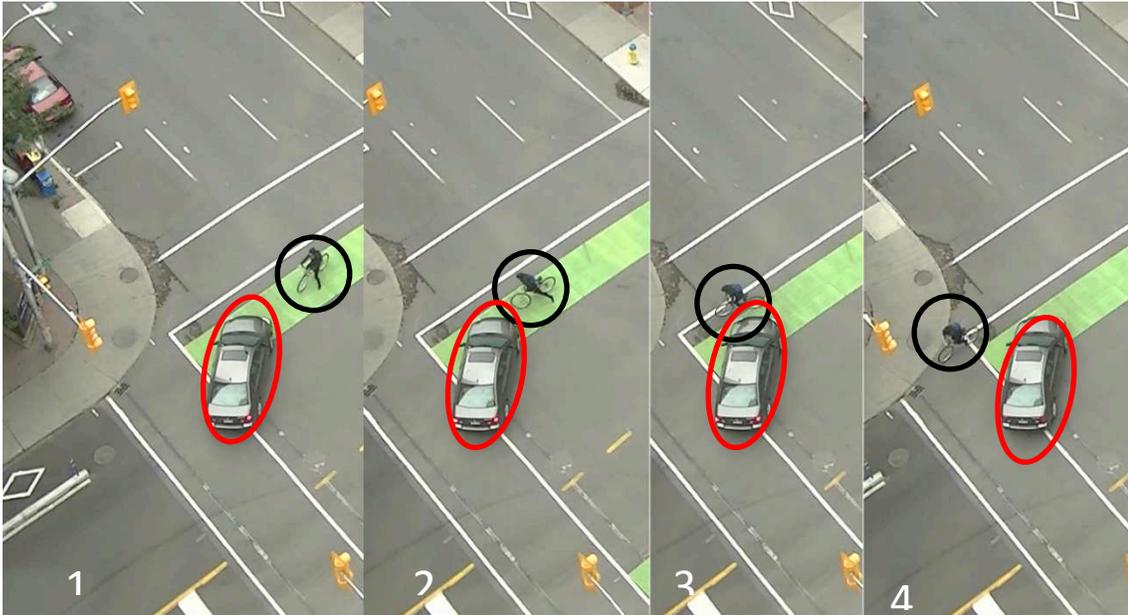


Figure 5-16: Cyclist crossing on red signal.

5.3.5g Vehicle turns left or right and stops on cyclist path

Figure 5-17 refers to the vehicle committing a violation by stopping on cyclist path while turns left or right. As shown in the figure, the turning vehicle yields the cyclist late and stopped on cyclist path. Therefore, the cyclist manoeuvred to avoid any possible collision and to return to the cyclist path.

(a)



(b)



Figure 5-17: Two examples of vehicle turn violation. (a) vehicle turns to the left and stopped on cyclist's path. (b) vehicle turns to the right and stopped on cyclist's path.

5.3.5h Vehicle fails to yield to the cyclist

The vehicle committed a violation by failing to yield to the cyclist during an interaction event as shown in Figure 5-18.

(a)



(b)



Figure 5-18: Two examples of vehicle fail to yield to the cyclist. (a) vehicle turns to the left and fails to yield to the cyclist. (b) vehicle turns to the right and fails to yield to the cyclist.

Other violations were initially observed, *e.g.*, yield hesitation and failure to use hand signals, but the reliability and consistency of their observation in all videos was not adequate and also their frequency was limited.

Types of evasive action: The exposed cyclists may have taken evasive action to avoid a collision in response to an interaction situation. The evasive actions observed were deceleration, stopping, acceleration, and swerving. Possible swerving actions by cyclists are shown in Figure 5-19.

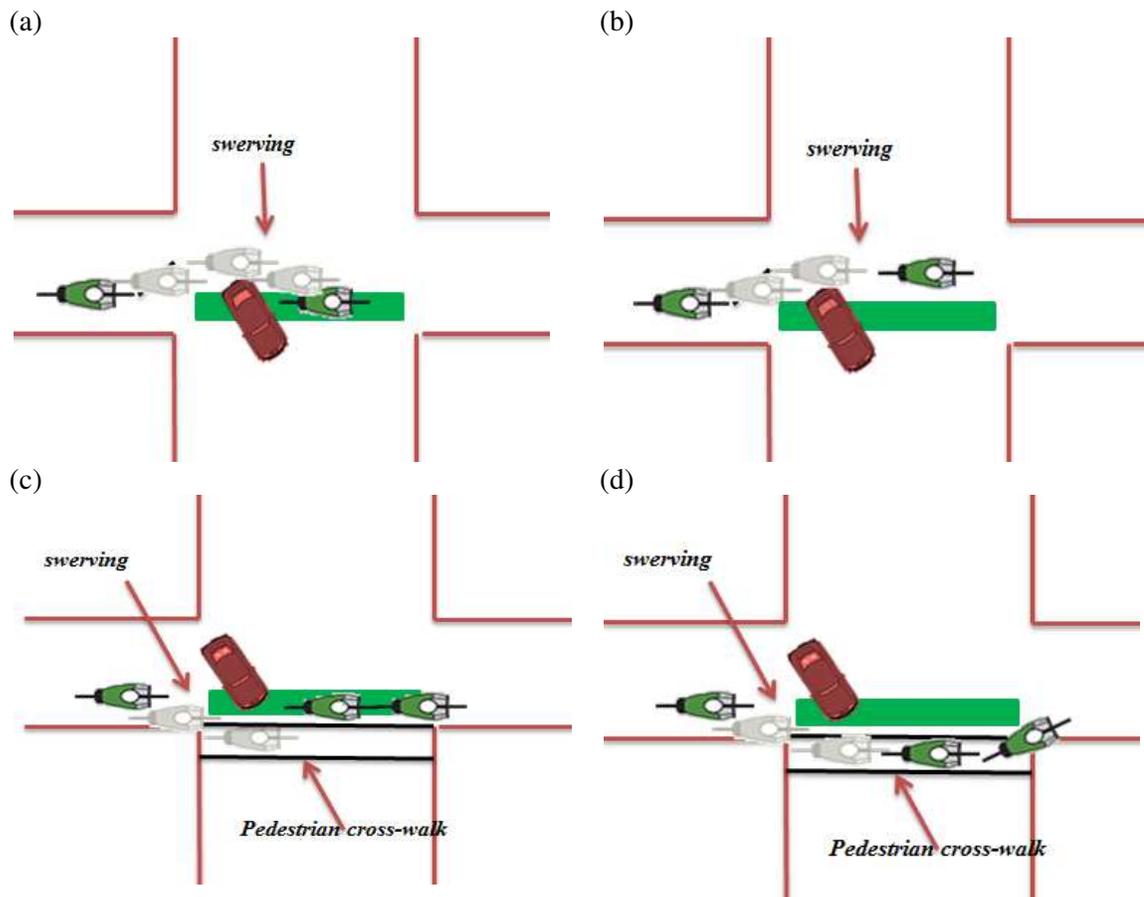


Figure 5-19: Schematic diagrams: Evasive actions by cyclists in a interaction situation. (a) and (c) swerving without violation. (b) swerving with riding out-of-path violation. (d) swerving with ride in cross-walk violation.

In this chapter, the performance parameters were measured based on video observations collected at treatment (segregated) and control (un-segregated) intersections. A total of 18,352 cyclists and 48,632 vehicles were observed in this study. Out of these, a total of 15,150 through cyclists and 31,604 right-turn and 17,028 left-turn vehicles were observed. The observations were made within a total period of 44 days and include 9 intersections.

5.3 Data Analysis

The observations in this chapter are mainly count events and the statistical models were selected accordingly. However, some counts are created based on threshold selection, *e.g.*, count of interactions. All PET measurements less than or equal to 3.0 seconds are counted. It would be advantageous if there is a statistical model which allows the use of the continuum of all PET measurements. This issue was investigated and the application of extreme value theory analysis was recommended to achieve this objective. The following sections present the statistical techniques along with their application to assess the measured performance parameters and support meaningful conclusions.

5.3.1 Different measure of exposure

It is difficult to compare the number of violation or interaction events without reference to exposure, *i.e.*, the total number of possible events which could turn into an interaction. An important variable which can help predict this number of interactions is traffic volume; cyclists and vehicles. This study follows an empirical approach to investigate the relationship between cyclist and vehicle volumes and interaction with PET less or equal 3.0 seconds. This is performed by comparing statistical models for candidate covariates which represent exposure.

The following figures show the relationship between different candidate exposure measures and the observed number of interactions. Figures 5-20 shows the relationship between the total number of through cyclist and the observed number of interactions of PET less or equal to 3.0 seconds in every observed video. Figure 5-21 shows the relationship between the total number of right-turn vehicles and the observed number of interactions of PET less than or equal to 3.0 seconds in every observed video. Figure 5-22 shows the relationship between the total number cyclist-vehicle events (CVE) and the observed number of interactions of PET less or equal to 3.0 seconds in every observed video. Upon close examination, it is noticeable that the exposure, which correlates the best with observed interaction count is the product of the right-turn vehicle volume in every observed video subsequence and the through cyclist volume in the same video subsequence. This is in fact the count of the maximum theoretical number of interactions which could happen in a video. This product is called in this study the total number of cyclist-vehicle events (CVE). The right-turn vehicle movements were taken here for the analysis due to the reason that was mentioned in Section 1.2.3. As shown in Figure 5-22, the line of equality provides a very good fit for the two sets of measurement, CVE and the observed number of interactions of PET less or equal to 3.0 seconds, with a high coefficient of determination, R^2 , of 0.63. For later parts of this chapter, candidate models are developed and statistical measures will be drawn to select the best model of PET counts in terms of goodness of fit.

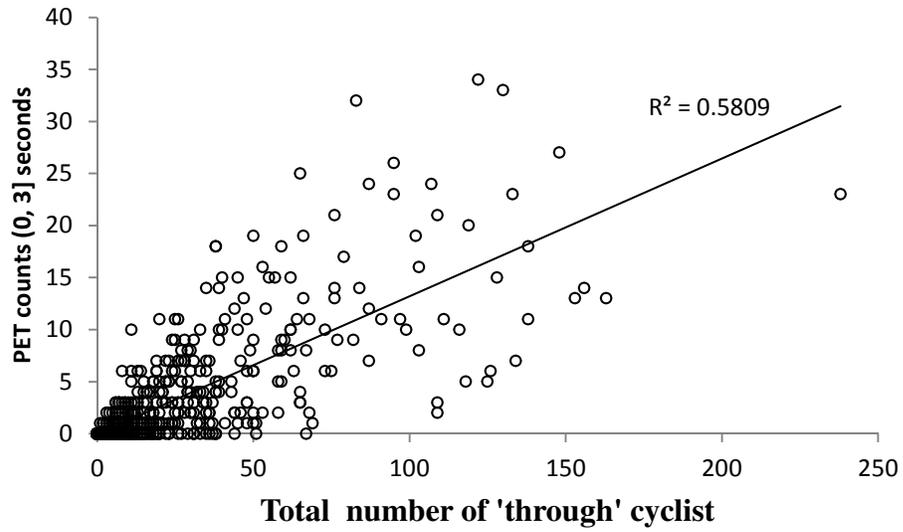


Figure 5-20: Cyclist volume as a candidate measure of exposure.

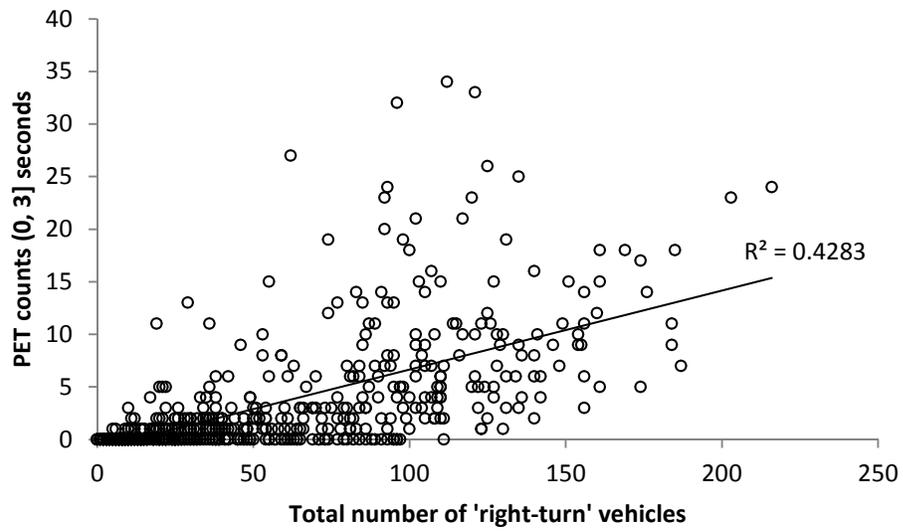


Figure 5-21: Turning vehicles as a candidate measure of exposure.

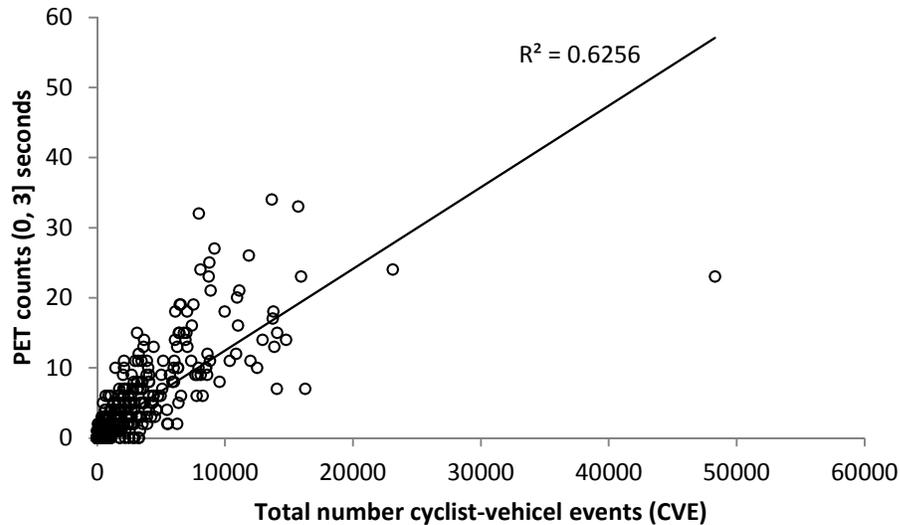


Figure 5-22: The product of turning vehicle volume and through cyclist volume as a candidate measure of exposure.

5.3.2 Comparative statistical analysis

The presentation of the statistical analysis is structured as a sequence of questions. The question is stated, data is presented, statistical tests are conducted, and comments are provided. The Two-tail t-test and Chi-square tests were conducted using the Excel statistical Add-on.

Does the cyclist volume change due to different seasons of observations?

Cyclist volume was collected at different seasons, at different intersections, for a common observation window during the day. Every intersection is monitored for a common observation window at the fall and summer seasons. Intersection data used to answer this question is limited to Intersections 1 to 7 from Table 5-1 because they have common data of cyclist volume for both seasons as shown in Table 5-2. The average hourly volume is calculated and paired two-tail t-test was conducted in order to test the

statistical significance of the difference in cyclist volume. The null hypothesis (H_0) is defined as ($H_0: M_f = M_s$), where, M_f : average hourly volume of cyclists in the fall season, M_s : mean average hourly volume of cyclists in the summer season. The p -value was found to be 0.02 (<0.05) and therefore the null hypothesis is rejected and the alternative hypothesis is accepted ($H_a: M_s \neq M_f$). The statistical results are summarized in the Table 5-3.

Table 5-2: Average hourly volume during the Fall and Summer 2011 seasons.

Cyclist Facility Type	Intersection Name	Hours	Cyclist Volume			
			Fall-2011		Summer-2012	
			Volume	Average cyclists/hour	Volume	Average cyclists/hour
Treatment	Metcalf	13	211	16	895	69
	O'Connor	26	546	21	1,918	74
	Kent	26	1,036	40	1395	54
	Lyon	11	159	14	364	33
	Bay	23	591	26	978	43
	Percy	16	107	7	359	22
Control	Somerset	13	363	28	406	31
Total		128	3,013	152	6,315	326

Table 5-3: Statistical results of the average hourly volume during Fall -2011 and Summer -2011 seasons.

Mean	21.69	46.51
Variance	115.37	383.65
Observations	7	7
df	6	
t Stat	-3.34	
P(T<=t) two-tail	0.02	
t Critical two-tail	2.44	

Does the average of cyclist right-turn interactions change with the different seasons at the study area?

Intersection data used to answer this question is limited to Intersections 1 to 6 from Table 5-1 because they have common data of cyclist right-turn interactions for both seasons as shown in Table 5-4. The average of cyclist right-turn interactions is calculated for each video using the formula as shown below. The paired two-tail t-test was conducted in order to test the statistical significance of the difference in average of cyclist right-turn interaction. The null hypothesis (H_0) is defined as ($H_0: R_f=R_s$), where, R_f : is the rate of the average of cyclist right-turn interaction at the fall season, R_s : is the rate of the average of cyclist right-turn interaction at the summer season. The p -value was found to be 0.96 (>0.05) and therefore the null hypothesis cannot be rejected. The statistical results are summarized in the Table 5-5.

Average of cyclist right-turn interactions per video (5.1)

$$= \left(\frac{\text{right-turn interactions}}{CVE} \right) \times 1000$$

Table 5-4: Data summary of the average of cyclist right-turn interactions for Fall - 2011 and Summer -2012 seasons.

Intersection name	Rate of average right-turn interactions	
	Fall-2011	Summer-2012
Metcalfe	1.95	1.45
Kent	2.00	1.52
Lyon	4.06	3.43
Bay	0.97	2.23
Percy	*	*
Somerset	2.14	2.57

*no interactions were observed

Table 5-5: Statistical results of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.

	Fall-2011	Summer-2012
Mean	1.85	1.87
Variance	1.84	1.37
Observations	6	6
df	5	
t Stat	-0.05	
P(T<=t) two-tail	0.96	
t Critical two-tail	2.57	

Furthermore, the average of cyclist right-turn interactions for each intersection was calculated using the formula as shown below. Intersection data used in this case is limited to Intersections 1 to 6 from Table 5-1 because they have common data of cyclist right-turn interactions for both seasons as shown in Table 5-6. The paired two-tail t-test was conducted in order to test the statistical significance of the difference in average of cyclist right-turn interaction. The null hypothesis (H_0) is defined as ($H_0: R_f=R_s$), where, R_f : is the average of cyclist right-turn interaction at the fall season, R_s : is the average of cyclist right-turn interaction at the summer season. The p -value was found to be 0.22 (>0.05)

and therefore the null hypothesis cannot be rejected. The statistical results are summarized in the Table 5-7.

Average of cyclist right-turn interactions per intersection (5.2)

$$= \left(\frac{\text{total right-turn interactions}}{\text{total } CVE} \right) \times 1000$$

Table 5-6: Data summary of the average of cyclist right-turn interactions for Fall - 2011 and Summer -2012 seasons.

Intersection Name	Average right-turn interaction	
	Fall-2011	Summer-2012
Metcalfe	1.46	1.07
Kent	2.10	1.59
Lyon	2.59	2.84
Bay	1.95	1.77
Percy	*	*
Somerset	2.28	2.14

*no interactions were observed

Table 5-7: Statistical results of the average of cyclist right-turn interactions for Fall -2011 and Summer -2012 seasons.

	Fall-2011	Summer-2012
Mean	1.73	1.57
Variance	0.86	0.94
Observations	6	6
df	5	
t Stat	1.41	
p(T<=t) two-tail	0.22	
t Critical two-tail	2.57	

Does the average of cyclist left-turn interactions change with the different seasons at the study area?

Intersection data used to answer this question is limited to Intersections 1 to 5 from Table 5-1 because they have common data of cyclist left-turn interactions for both seasons as shown in Table 5-8. The average of cyclist left-turn interactions is calculated for each video using the formula as shown below. The paired two-tail t-test was conducted in order to test the statistical significance of the difference in average of cyclist left-turn interaction. The null hypothesis (H_0) is defined as ($H_0: R_f=R_s$), where, R_f : is the rate of the average of cyclist left-turn interaction at the fall season, R_s : is the rate of the average of cyclist left-turn interaction at the summer season. The p -value was found to be 0.12 (>0.05) and therefore the null hypothesis cannot be rejected. The statistical results are summarized in the Table 5-9.

Average of cyclist left-turn interactions per video (5.3)

$$= \left(\frac{\text{left-turn interactions}}{CVE} \right) \times 1000$$

Table 5-8: Data summary of the average of cyclist left-turn interaction for Fall - 2011 and Summer -2012 seasons.

Intersection Name	Rate of average left-turn interactions	
	Fall-2011	Summer-2012
Metcalfe	0.49	1.20
Kent	1.61	1.36
Lyon	1.35	2.71
Bay	0.19	1.65
Percy	0.22	0.28

Table 5-9: Statistical results of the average of cyclist left-turn interactions for Fall - 2011 and Summer -2012 seasons.

	Fall-2011	Summer-2012
Mean	0.77	1.44
Variance	0.44	0.76
Observations	5	5
df	4	
t Stat	-1.96	
P(T<=t) two-tail	0.12	
t Critical two-tail	2.78	

Does the proportion of cyclists involved in a right-turn or left-turn interactions change with the different seasons at each intersection?

The proportion is the number of cyclists involved in a right-turn or left-turn interactions out of the total through cyclist movement. Intersection data used to answer this question is limited to Intersections 1 to 5 from Table 5-1 because they have common data for both seasons as shown in Table 5-10. The proportion was calculated using the formula as shown below and the chi-square test was conducted in order to test the statistical significance of the difference in cyclist proportion. The null hypothesis (H_0) is

defined as ($H_0: P_f=P_s$), where, P_f : interaction proportion of cyclist in fall season, P_s : interaction proportion of cyclist in summer season. The p -value was found to be 0.01 (<0.05) and therefore the null hypothesis is rejected and the alternative hypothesis is accepted ($H_a: P_f \neq P_s$). A sample of statistical results for Metcalfe intersection is summarized in the Table 5-11. Summary of findings seasonal change in proportion of cyclists involved in right-turn or left-turn interactions for all intersections is shown in Table 5-12.

proportion of cyclist involved in interactions (5.3)

$$= \left(\frac{\text{total right-turn or left-turn interactions}}{\text{total through cyclist movement}} \right) \times 100$$

Table 5-10: Summary of right-turn and left-turn interactions rate for Fall -2011 and Summer -2012 seasons.

Intersection Name	Cyclists' proportion of right-turn interactions		Cyclists' proportion of left-turn interactions	
	Fall-2011	Summer-2012	Fall-2011	Summer-2012
Metcalfe	20.79	14.82	1.32	3.60
Kent	18.04	15.58	2.99	3.75
Lyon	6.12	5.62	8.67	14.05
Bay	18.24	21.97	1.35	0.95
Percy	-*	-*	0.31	0.76

*no interactions were observed

Table 5-11: Statistical results of right-turn interaction rate for different seasons at Metcalfe intersection.

Metcalfe	Right-turn interaction	Non-interaction	Total
Fall	79	301	380
Summer	202	1161	1,363
Total	281	1,462	1,743
<i>Expected value</i>			
Fall	61.26	318.74	380
Summer	219.74	1,143.27	1,363
Total	281	1,462	1,743
<i>p-value</i>	0.01		

Table 5-12: Summary of findings seasonal change in proportion of cyclists involved in right-turn or left-turn interactions.

Intersection Name	<i>p-value</i>	
	Right-turn interaction	Left-turn interaction
Metcalfe	0.01	0.02
Kent	0.09	0.29
Lyon	0.68	<0.05 (0.002)
Bay	0.15	0.55
Percy	-*	0.43

*no interactions were observed

As shown in Table 5-12, there is no strong evidence of change in the proportion of cyclists involved in right-turn interaction across the study area, except at Metcalfe and Lyon. The findings at Metcalfe are consistent for both types of interactions and are interpreted as a reduction in the interaction rate from fall to summer.

Does the rate of cyclist violation change with different seasons at the study area?

As explained in Section 5.3, different types of violations were studied. The rate is the number of cyclists involved in a violation out of the total cyclist volume. The rate was calculated for each intersection using the formula as shown below. Intersection data used

to answer this question is limited to Intersections 1 to 7 from Table 5-1 because they have common data of cyclist volume for both seasons as shown in Table 5-13. The paired two-tail t-test was conducted in order to test the statistical significance of the difference in violation rate. The null hypothesis (H_0) is defined as ($H_0: R_f=R_s$), where, R_f : is the proportion of cyclist violations out of the total cyclist volume in the fall season, R_s : is the proportion of cyclist violations out of the total cyclist volume in the summer season. The p -value was found to be 0.07 (>0.05) and therefore the null hypothesis cannot be rejected. The statistical results are summarized in the Table 5-14.

$$\text{Violation rate} = \left(\frac{\text{total violation}}{\text{total cyclist volume}} \right) \times 100 \quad (5.4)$$

In addition, the rate of the number of motor vehicles involved in a violation out of the total vehicle volume was studied. The rate was calculated for each intersection using the formula as shown below. Intersection data used to answer this question is limited to Intersections 1 to 4 and Intersection number 6 from Table 5-1 because they have common data of right-turn motor vehicle volume for both seasons as shown in Table 5-15. The paired two-tail t-test was conducted in order to test the statistical significance of the difference in violation rate. The null hypothesis (H_0) is defined as ($H_0: R_f=R_s$), where, R_f : is the proportion of motor vehicle violation out of the total right-turn motor vehicle volume in the fall season, R_s : is the proportion of motor vehicle violation out of the total right-turn motor vehicle volume in the summer season. The p -value was found to be 0.18 (>0.05) and therefore the null hypothesis cannot be rejected. The statistical results are summarized in the Table 5-16.

$$\text{Violation rate} = \left(\frac{\text{total violation}}{\text{right-turn vehicle volume}} \right) \times 100 \quad (5.5)$$

Table 5-13: Summary of the cyclist violation rate for Fall -2011 and Summer -2012 seasons.

Intersection Name	Percentage of cyclists committing violation	
	Fall-2011	Summer-2012
Metcalfe	11.03	13.63
O'Connor	40.32	20.59
Kent	16.23	12.12
Lyon	28.92	22.73
Bay	23.78	19.42
Percy	4.15	3.93
Somerset	17.63	4.43

Table 5-14: Statistical results of the cyclist violation rate for Fall -2011 and Summer -2012 seasons.

	Fall-2011	Summer-2012
Mean	20.30	13.84
Variance	142.93	57.60
Observations	7	7
df	6	
t Stat	2.23	
P(T<=t) two-tail	0.07	
t Critical two-tail	2.45	

Table 5-15: Summary of the motor vehicle violation rate for Fall -2011 and Summer -2012 seasons.

Intersection Name	Percentage of motor vehicles committing violation	
	Fall-2011	Summer-2012
Metcalfe	0.41	1.37
Kent	1.21	1.53
Lyon	0.28	0.00
Bay	0.05	0.25
Somerset	0.11	0.54

Table 5-16: Statistical results of the motor vehicle violation rate for Fall -2011 and Summer -2012 seasons.

	Fall-2011	Summer-2012
Mean	0.41	0.74
Variance	0.22	0.46
Observations	5	5
df	4	
t Stat	-1.63	
P(T<=t) two-tail	0.18	
t Critical two-tail	2.78	

5.4 Regression Analysis of Count Data

5.4.1 Counts models

There are a number of methods that can be used to properly model data counts (interactions). The most popular regression that can be used to model data counts is Poisson and Negative Binomial regression (Hunter, *et al.*, 2009; Lord & Mannering, 2010; Washington, *et al.*, 2011). Negative Binomial regression is a classic model for modelling collision counts (Lord & Mannering, 2010), with applications in the prediction of PET counts (Songchitrukksa & Tarko, 2007). Regression analysis for Negative Binomial regression takes place within the framework of generalized linear regression as follows:

$$m_i = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n} \quad (5.6)$$

where m_i is the mean value of the PET count given the values of covariates x_1, x_2, \dots, x_n .

The typical purpose of regression analysis is to estimate the parameters $\beta_0, \beta_1, \beta_2, \dots, \beta_n$. The covariates can potentially affect the frequency of PET counts and this is tested during the parameter estimation process. The regression outcome shows the statistical significance of the variables, how likely to generate the data with no true effect, and also

their signs. Because Negative Binomial regression is in theory an expansion of the Poisson regression (Wang & Nihan, 2004), the count models in this analysis were developed for both distributions.

The limitation of Poisson regression is that it restricts the mean and variance to be equal. If this quality does not hold, the data can be under-dispersed ($E(y_i) > var(y_i)$) or overdispersed ($E(y_i) < var(y_i)$). The Negative Binomial model aims to address this issue (Washington, *et al.*, 2011). In addition, for the case when the number of zero counts is excessive, the Zero-Inflated Negative Binomial regression model is found to be a suitable model to overcome this issue (Famoye & Singh, 2006).

5.4.2 Effect of cyclist volume, vehicle volume, treatment and season on PET counts

The generalized relationship between the dependent variable (PET counts) and the independent variables (cyclist volume, vehicle volume, CVE, log cyclist volume, log vehicle volume, log CVE, treatment, and season) using Poisson regression, Negative Binomial regression and Zero-Inflated Negative Binomial regression, the regression analysis described in previous section is applied.

Two factors are used to describe the season and the treatment. The factor representing treatment takes the value of 1 for videos collected at treatment intersections and 0 at videos collected at control intersections. The factor representing the season is 1 for videos collected in the fall and 0 for videos collected in the summer. Positive signs of these coefficients reflect a tendency to generate more interactions (PET counts) while a negative sign represents the opposite effect. Poisson regression, Negative Binomial

regression, and goodness of fit measures were conducted using relevant packages in the *R* statistical computing language (R_Core_Team, 2012).

Several fitted models were generated and the comparison in model fit is conducted using the Akaike Information Criterion (AIC). Lower values of AIC indicate improved model fit without using too many parameters. The AIC is defined as (R_Core_Team, 2012);

$$AIC = -2\log\text{-likelihood} + k * n_{par} \quad (5.7)$$

where *AIC* is the Akaike Information Criterion, $k = 2$ for the usual *AIC* or $k = \log(n)$ (n is the number of the observation), and n_{par} is represent to the number of parameters in the fitted model.

Table 5-17 shows the results of the Poisson and Negative Binomial regression of PET (0, 3] seconds. Depending on the AIC value (1,726.7), the best model was the model which uses the natural logarithm of CVE as the measure of exposure, treatment and season. Table 5-18 shows the results of the Poisson and Negative Binomial regression of PET (0, 2] seconds. The lowest value of the AIC was found to be (1,448.2) and the best model was the model which uses the natural logarithm of CVE as the measure of exposure and the treatment. Table 5-19 shows the results of Poisson and Negative Binomial regression of PET (0, 1] second. The lowest value of the AIC was found to be (731.6) and the best model was the model which uses the natural logarithm of CVE as the measure of exposure and the treatment. For all models, the classic Negative Binomial regression proved to be better than the Poisson model. This is consistent with collision prediction theory.

In order to mitigate any estimation issues which may occur due to potential correlation between log(CVE) and season, a model containing only log(CVE) and treatment was estimated. For all proximity thresholds, this model was found to have lower AIC values, providing another evidence that season is not significant factor.

The level of significance for the estimated parameters, covariate multipliers (coefficients), is provided in parentheses in Tables 5-17, 5-18, and 5-19. As explained earlier positive signs of these parameters (multipliers of variables and factors) reflect a tendency to generate more interactions while a negative sign represents the opposite effect. That is, positive sign for treatment parameter indicates a harmful effect and a negative sign for treatment parameter indicates a safety benefit. Positive sign for season indicates an elevated number of interactions in the fall as compared to the summer. Exposure measures were found all to be positive; as expected for all selected models. Table 5-20 shows the best selected three models with the significant level for each parameter and with AIC value for each model.

Table 5-17: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 3] seconds.

Model No.	Distribution	Intercept (p-value)	Cyclist volume [*]	Vehicle volume ^{**}	CVE	Log (cyclist volume)	Log (vehicle volume)	Log (CVE)	Treatment	Season	AIC
		β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	
1	Poisson	1.56 (<0.001)	-	-	8.05e-05 (<0.001)	-	-	-	-5.05e-01 (<0.001)	-0.74 (<0.001)	3,557.0
2	Poisson	0.05 (0.54)	0.01 (<0.001)	0.01 (<0.001)	-	-	-	-	-0.24 (<0.001)	-0.25 (<0.001)	2,539.8
3	Poisson	-5.08 (<0.001)	-	-	-	-	-	0.86 (<0.001)	-0.22 (<0.001)	0.16 (<0.001)	1,844.2
4	Poisson	-4.99 (<0.001)	-	-	-	0.88 (<0.001)	0.83 (<0.001)	-	-0.24 (<0.001)	0.16 (0.0031)	1,845.3
5	Negative Binomial	0.60 (<0.001)	-	-	3.31e-04 (<0.001)	-	-	-	-0.57 (<0.001)	-0.53 (<0.001)	2,108.9
6	Negative Binomial	-0.89 (<0.001)	0.02 (<0.001)	0.02 (<0.001)	-	-	-	-	-0.28 (0.02)	-0.18 (0.08)	1,902.9
7	Negative Binomial	-5.41 (<0.001)	-	-	-	-	-	0.90 (<0.001)	-0.23 (0.01)	0.15 (0.08)	1,726.7
8	Negative Binomial	-5.39 (<0.001)	-	-	-	0.91 (<0.001)	0.89 (<0.001)	-	-0.24 (0.01)	0.15 (0.08)	1,728.7
9	Negative Binomial	-5.26 (<0.001)	-	-	-	-	-	0.89 (<0.001)	-0.24 (0.01)	-	1,727.8

*Through cyclist movement, ** Right-turn vehicle movement

Table 5-18: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 2] seconds.

Model No.	Distribution	Intercept (p-value)	Cyclist volume [*]	Vehicle volume ^{**}	CVE	Log (cyclist volume)	Log (vehicle volume)	Log (CVE)	Treatment	Season	AIC
		β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	
1	Poisson	1.28 (<0.001)	-	-	8.22E-05 (<0.001)	-	-	-	-0.55 (<0.001)	-0.88 (<0.001)	2,759.9
2	Poisson	-0.23 (0.01)	0.01 (<0.001)	0.01 (<0.001)	-	-	-	-	-0.30 (<0.001)	-0.37 (<0.001)	2,020.8
3	Poisson	-5.72 (<0.001)	-	-	-	-	-	0.91 (<0.001)	-0.28 (<0.001)	0.02 (0.764)	1,524.8
4	Poisson	-5.58 (<0.001)	-	-	-	0.94 (<0.001)	0.85 (<0.001)	-	-0.31 (<0.001)	0.031 (0.646)	1,525.3
5	Negative Binomial	0.32 (0.0394)	-	-	3.22E-04 (<0.001)	-	-	-	-0.61 (<0.001)	-0.64 (<0.001)	1,780.3
6	Negative Binomial	-1.2 (<0.001)	0.02 (<0.001)	0.02 (<0.001)	-	-	-	-	-0.33 (0.012)	-0.28 (0.011)	1,594.6
7	Negative Binomial	-6.03 (<0.001)	-	-	-	-	-	0.94 (<0.001)	-0.28 (0.005)	0.0174 (0.850)	1,458.2
8	Negative Binomial	-5.96 (<0.001)	-	-	-	0.96 (<0.001)	0.92 (<0.001)	-	-0.30 (0.004)	0.025 (0.791)	1,459.9
9	Negative Binomial	-6.01 (<0.001)	-	-	-	-	-	0.94 (<0.001)	-0.29 (0.004)	-	1,448.2

*Through cyclist movement, ** Right-turn vehicle movement

Table 5-19: Summary of the results of the Poisson and Negative Binomial regression of PET (0, 1] second.

Model No.	Distribution	Intercept (p-value)	Cyclist volume [*]	Vehicle volume ^{**}	CVE	Log (cyclist volume)	Log (vehicle volume)	Log (CVE)	Treatment	Season	AIC
		β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	
1	Poisson	-0.36 (0.020)	-	-	8.52E-05 (<0.001)	-	-	-	-0.21 (0.195)	-1.11 (<0.001)	1,102.0
2	Poisson	-1.75 (<0.001)	0.01 (<0.001)	0.01 (<0.001)	-	-	-	-	-0.07 (0.686)	-0.53 (<0.001)	896.7
3	Poisson	-8.41 (<0.001)	-	-	-	-	-	1.03 (<0.001)	-0.01 (0.936)	-0.09 (0.508)	740.2
4	Poisson	-7.98 (<0.001)	-	-	-	1.13 (<0.001)	0.87 (<0.001)	-	-0.07 (0.698)	-0.05 (0.702)	739.2
5	Negative Binomial	-1.24 (<0.001)	-	-	2.85E-04 (<0.001)	-	-	-	-0.30 (0.195)	-0.72 (<0.001)	884.0
6	Negative Binomial	-2.83 (<0.001)	0.02 (<0.001)	0.02 (<0.001)	-	-	-	-	-0.04 (0.856)	-0.34 (0.064)	799.2
7	Negative Binomial	-8.77 (<0.001)	-	-	-	-	-	1.08 (<0.001)	-0.02 (0.898)	-0.08 (0.617)	733.4
8	Negative Binomial	-8.43 (<0.001)	-	-	-	1.14 (<0.001)	0.95 (<0.001)	-	-0.04 (0.846)	-0.05 (0.746)	734.1
9	Negative Binomial	-8.84 (<0.001)	-	-	-	-	-	1.08 (<0.001)	-0.03 (0.870)	-	731.6

*Through cyclist movement, ** Right-turn vehicle movement

Table 5-20: Suggested the best three Negative Binomial models to accommodate the PET counts.

Model No. (Table)	$Y = \exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3)$				
	PET counts	Parameters			AIC
		Variable	Estimate	P- value	
7/ (5-15)	(0, 3]	β_0	-5.41	< 0.001	1,726.7
		β_1	0.90	< 0.001	
		β_2	-0.23	0.013	
		β_3	0.15	0.08	
9/ (5-16)	(0, 2]	β_0	-6.01	< 0.001	1,448.2
		β_1	0.94	< 0.001	
		β_2	-0.29	0.004	
9/ (5-17)	(0, 1]	β_0	-8.84	< 0.001	731.6
		β_1	1.08	< 0.001	
		β_2	-0.03	0.870	

In addition, the same three Negative Binomial models were applied in order to study the individual effect of treatment, season, and the exposure measures on PET counts, in the case of the vehicle falls to yield (negative) to the cyclist when the vehicle making a right turn and cross the cycle path as described in Section 5.3.4. The summary of regression analysis for PET counts (0, 3], (0, 2] and (0, 1] seconds are shown in Tables 5-21, 5-22 and 5-23.

Table 5-21: Summary of regression analysis of PET (0, 3] seconds when vehicle yields to the cyclist (positive).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-5.44 (<0.001)	0.85 (<0.001)	-0.26 (0.01)	0.20 (0.04)	1,479.2
2	NB	-5.22 (<0.001)	0.83 (<0.001)	-0.28 (0.01)	-	1,481.4

Table 5-22: Summary of regression analysis of PET (0, 2] seconds when vehicle yields to the cyclist (positive).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-6.17 (<0.001)	0.90 (<0.001)	-0.40 (<0.001)	0.061 (0.582)	1,187.8
2	NB	-6.11 (<0.001)	0.90 (<0.001)	-0.41 (<0.001)	-	1,186.1

Table 5-23: Summary of regression analysis of PET (0, 1] second when vehicle yields to the cyclist (positive).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-10.47 (<0.001)	1.18 (<0.001)	-0.82 (0.004)	-0.10 (0.752)	317.85
2	NB	-10.54 (<0.001)	1.18 (<0.001)	-0.81 (0.004)	-	315.95
3	ZEROINF NB	-8.24 (<0.001)	0.95 (<0.001)	-0.71 (0.08)	-0.51 (0.212)	320.94
4	ZEROINF NB	-3.36 (0.905)	-0.51 (0.440)	5.59 (0.841)	-	321.83

Tables 5-24, 5-25, and 5-26 summarize the findings of Negative Binomial and Zero-inflated regression, in the case of the vehicle yields (positive) to the cyclist when the vehicle making a right turn and cross the cyclist path. In the case of negative and

positive cases of PET counts (0, 1] second has too many zeros (observed proportion of zeros of PET counts negative and positive cases were 69.0 % and 88.5% respectively). Therefore, the Zero-Inflated Negative Binomial regression model was applied as described in Section 5.5.1.

The regression analysis reveals the following observations:

- 1- All signs of the treatment effect in all models are found to be negative and statistically significant for PET thresholds of (0, 3] (interaction between cyclist and motor vehicle less or equal 3 seconds) and (0, 2] seconds. For threshold of (0, 1] second, the effect is not statistically significant.
- 2- The signs of the season factor are positive with models containing the logarithm of CVE for PET thresholds of (0, 3] and (0, 2] seconds and negative for PET threshold of (0, 1] second. Positive sign for season indicates an elevated number of interactions in the fall as compared to the summer. However, they are not statistically significant.
- 3- The models containing the logarithm of CVE were found to provide the best model.

Table 5-24: Summary of regression analysis of PET (0, 3] seconds when vehicle fails to yield to the cyclist (negative).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-7.14 (<0.001)	0.99 (<0.001)	-0.17 (0.200)	0.06 (0.606)	1,128.2
2	NB	-7.08 (<0.001)	0.99 (<0.001)	-0.18 (0.189)	-	1,126.4

Table 5-25: Summary of regression analysis of PET (0, 2] seconds when vehicle fails to yield to the cyclist (negative).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-7.39 (<0.001)	0.99 (<0.001)	-0.11 (0.446)	-0.04 (0.731)	1,004.2
2	NB	-7.44 (<0.001)	1.00 (<0.001)	-0.11 (0.457)	-	1,002.4

Table 5-26: Summary of regression analysis of PET (0, 1] second when vehicle fails to yield to the cyclist (negative).

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Treatment β_2	Season β_3	AIC
1	NB	-9.18 (<0.001)	1.06 (<0.001)	0.38 (0.104)	-0.09 (0.641)	659.2
2	NB	-9.25 (<0.001)	1.06 (<0.001)	0.39 (0.097)	-	657.4
3	ZEROINF NB	-7.07 (<0.001)	0.84 (<0.001)	0.32 (0.242)	-0.02 (0.941)	663.6
4	ZEROINF NB	-8.72 (<0.001)	-1.01 (<0.001)	0.29 (0.216)	-	656.9

Another test of the Negative Binomial regression model that contains the logarithm of CVE was applied in order to study the effect of each intersection on PET count. This intersection-specific variable can be used subsequently to rank these intersections. Two

factors are used to describe the intersection. When videos are being collected from a specific intersection, the factor that is representing the intersection takes the value of 1 and 0 at the videos collected at the rest of the intersections. Table 5-27 summarizes the findings of Negative Binomial regression model for the PET counts (0, 3], (0, 2] and (0, 1] seconds. As shown in Table 5-27 the effect and the significance of each intersection on PET counts (0, 3], (0, 2] and (0, 3] seconds is considered negligible (not significant).

Table 5-27: Summary of negative binomial regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds.

PET seconds	Intercept (p-value)	Log (CVE)	Metcalfe	O'Connor	Bank*	Bank**	Kent	Lyon	Bay	Somerset	Sussex	AIC
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	
(0, 3]	-6.36 (<0.001)	0.90 (<0.001)	0.62 (0.563)	0.28 (0.794)	1.04 (0.332)	0.67 (0.537)	0.80 (0.453)	1.10 (0.309)	0.85 (0.430)	1.02 (0.344)	0.97 (0.365)	1,684.5
(0, 2]	-23.29 (0.992)	0.97 (<0.001)	16.49 (0.994)	16.38 (0.994)	17.09 (0.994)	16.67 (0.994)	16.88 (0.994)	17.19 (0.994)	16.89 (0.994)	17.00 (0.994)	17.10 (0.994)	1421.1
(0, 1]	-33.29 (1.000)	1.07 (<0.001)	24.24 (1.000)	24.55 (1.000)	25.05 (1.000)	24.69 (1.000)	24.83 (1.000)	24.86 (1.000)	23.89 (1.000)	24.44 (1.000)	24.64 (1.000)	722.9

*Eastbound

**Westbound

5.4.3 Effect of pedestrian volume on PET Counts

One of the questions which can be answered is whether the variability in safety measures is due to variability in pedestrian volume (parallel to cyclist volume). Count models can enable a straightforward answer to this question. In order to study the effect of pedestrian volume on PET counts, video data was reviewed for one day (The video data covers a common observation period from 8:30 AM to 4:00 PM) in six intersections. The intersections that were selected are; Metcalfe, O'Connor, Bank, Kent, Lyon and Bay for the following reasons; [i] physically segregated bike lanes were present in all these intersections, [ii] variability in pedestrian volume across intersections and in different times of the day was evident in this study, and [iii] the intersections are reasonably homogeneous in terms of land use: [1] all intersections are sequentially arranged along the major street and [2] all intersections are in the Downtown area and time of the year (all video data was collected in the summer of 2012). The pedestrians' volume at the crosswalks where the interactions were under investigation was calculated at each intersection. The pedestrian volume observed at each intersection is shown in Table 5-28. The Negative Binomial regression model that contains the logarithm of CVE was applied as it was proved to be better than the Poisson model. The significance level of the estimated pedestrian volume, covariate multipliers, is shown in Table 5-29 for the levels of PET counts (0, 3], (0, 2], and (0, 1] seconds.

Table 5-28: Pedestrian volume at six signalized intersections (8:30 AM to 4:00 PM).

Intersection	Pedestrian volume
Metcalfe	5,296
O'Connor	5,222
Bank	6,405
Kent	4,781
Lyon	1,281
Bay	527

Table 5-29: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds.

(a) (0, 3] seconds

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
Negative Binomial	-5.42 (<0.001)	0.88 (<0.001)	-0.0003 (0.283)	367.04

(b) (0, 2] seconds

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
Negative Binomial	-6.26 (<0.001)	0.95 (<0.001)	-0.0005 (0.200)	329.23

(c) (0, 1] second

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
Negative Binomial	-8.63 (<0.001)	1.01 (<0.001)	0.0007 (0.300)	189.42

The reviewed data set, in the case of negative and positive cases of PET count (0, 1] second has too many zeros (observed proportion of zeros of PET counts negative and

positive cases were 44 % and 50% respectively). Table 5-30 summarize the findings of Negative Binomial and Zero-inflated Negative Binomial regression, in the case of the vehicle yields (positive) the cyclist when the vehicle making a right turn and cross the cyclist path. While Table 5-31 summarize the findings of Negative Binomial and Zero-inflated Negative Binomial regression, in the case of the vehicle falls to yield (negative) to the cyclist when the vehicle making a right turn and cross the cycle path. The effect and the significance of the pedestrian volume on PET counts are summarized in Table 5-32.

Table 5-30: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds when vehicle yields to the cyclist (positive).

(a) (0, 3] seconds

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
NB	-5.60 (<0.001)	0.87 (<0.001)	-0.0009 (0.028)	308.79

(b) (0, 2] seconds

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
NB	-6.53 (<0.001)	0.95 (<0.001)	-0.001 (0.035)	270.2

(c) (0,1] second

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
1	NB	-11.31 (<0.001)	1.39 (<0.001)	-9.12E-05 (0.943)	88.14
2	ZEROINF NB	-11.77 (<0.001)	1.19 (<0.001)	0.001 (0.441)	91.54

Table 5-31: Summary of regression analysis of PET counts (0, 3], (0, 2] and (0, 1] seconds when vehicle fails to yield to the cyclist (negative).

(a) (0, 3]

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
NB	-7.03 (<0.001)	0.91 (<0.001)	0.0007 (0.172)	259.04

(b) (0, 2]

Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
NB	-7.52 (<0.001)	0.95 (<0.001)	0.0004 (0.507)	237.48

(c) (0,1]

Model No.	Distribution	Intercept (<i>p</i> -value) β_0	Log (CVE) β_1	Pedestrian volume β_2	AIC
1	NB	-7.91 (<0.001)	0.88 (<0.001)	0.0001 (0.177)	169.29
2	ZEROINF NB	-7.46 (0.184)	0.30 (0.546)	0.007 (0.057)	169.66

Table 5-32: Summary of the pedestrian effects on PET counts.

PET (second)	Covariate multiplier sign	p-value	Result
(0, 3]	negative*	0.28	Not significant
(0, 2]	negative	0.20	Not significant
(0, 1]	positive	0.30	Not significant
Positive case**			
(0, 3]	negative	0.02	significant
(0, 2]	negative	0.03	significant
(0, 1]	negative	0.94	Not significant
ZEROINF NB	positive	0.44	Not significant
Negative case			
(0, 3]	positive	0.17	Not significant
(0, 2]	positive	0.50	Not significant
(0, 1]	positive	0.17	Not significant
ZEROINF NB	positive	0.05	Not significant

*Positive signs reflect a tendency to generate more interactions while a negative sign represents the opposite effect.

** Vehicle yield the cyclist when vehicle making a right turn and cross the cyclist path while a negative case represents the opposite effect.

As shown in the Table 5-32, the negative sign of the covariate multiplier refers to a desirable effect of the pedestrian volume on cyclist's safety. This was clear on the PET proximity level (0, 2] and (0, 3] seconds. In the positive PET scenario, the pedestrian volume has had a positive impact on cyclist safety making the vehicle yields to the cyclist when the vehicle is making a right turn and crosses the cyclist path. However, in the negative PET scenario, the pedestrian volume was not found to have a significant effect.

That is, it does not appear that pedestrian volume affect motorist behaviour for vehicles failing to yield to the cyclist when vehicle is making a right turn and crosses the cyclist path. This was shown in the results of the PET proximity level (0, 1] second.

In order to examine other aspects of the effect of pedestrian volume, non-conformal vehicle movements (traffic violations) were identified. For example, vehicle turns left or right and stops on cyclist path as shown in Figure 5-17. The total number of vehicle violation was found to be 80 out of 6,522 right-turn vehicles. This violation occurred by the vehicle due to the presence of pedestrians, it may cause an interaction with cyclists or may lead the cyclist to commit a violation. As shown in the figure, the turning vehicle yielded late to the cyclist and stopped on cyclist path. Therefore, the cyclist manoeuvred to avoid any possible collision and to return to the cyclist path. The correlation between the pedestrian volume and the average vehicle violation was studied per video and the result was 0.20. Thus, it does not appear that pedestrian volume has an effect on this pattern of motorist violations.

5.5 Statistical Modeling Based on Extreme Value Theory

5.5.1 Background

Extreme value theory (EVT) is the theory of modeling and measuring events which occurs with extremely low probabilities (Coles, 2001). Collisions are much more rare events than interactions. Capability of EVT is to estimate the probability of extreme rare events from frequent events. Therefore, the collision related information can be obtained from interactions by using EVT, it was used to validate surrogate safety measures such as interactions (Zheng, *et al.*, 2014).

There are two approaches in EVT, and they are Block Maxima (BM) approach and Peak Over Threshold (POT) approach. Studies have shown that the POT approach performance is better than the BM approach in the case of limited data (Carries, 2009). In this study, data collected within limited time at each intersection, therefore the POT is used. The POT approach takes the observations exceeding predetermined threshold as extreme events. (Pickands, 1975), proved that when the threshold is high enough the extremes will follow the Generalized Pareto distribution (GPD), as shown in Equation 5.8.

$$H(y) = H(x - u) = 1 - \left(1 + \frac{\xi(x - u)}{\sigma}\right)^{-1/\xi} \quad (5.8)$$

where y is exceedance over the threshold, x is the extreme events, u is the threshold, $\sigma > 0$ is the scale parameter, and $-\infty < \xi < \infty$ is the shape parameter.

5.5.2 Return level value

In EVT, the return level value is the standard interpretation of Generalized Pareto Distribution (GPD) model. In common terminology, the return level (z_p) is associated with the return period $N = 1/p$, this can be obtained in Equation 5.9.

$$z_p = u + \frac{\sigma}{\xi} [(Nn\lambda_u)^\xi - 1] \quad (5.9)$$

where, n is the number of the observations during the observation time, λ_u is the proportion of observations exceeding the threshold.

Equation 5.9, presents a time based description of return level, and the safety analysis prefers an event based description. In this study CVE (as described in Section 5.4.2) is used instead of time to calculate the return level. In this case, N is the total number of CVE. For example, if $N=1,000,000$ CVE, the return level z_p can be interpreted

as: when observing PET (seconds) values, a PET value of z_p (seconds) is expected to be exceeded once every 1,000,000 CVE. This can also be interpreted as: there is a probability of $1/1,000,000$ that the return level, z_p , is exceeded for every CVE.

PET values are obtained in seconds, with tail distribution pointing to extreme values close to zero. Analysis is conducted based on negated values of PET, *i.e.*, every PET value is multiplied by -1. This is done in order to create a tail distribution pointing to positive real numbers. Thus, this distribution will be consistent with GPD definition (extremes are on the right side or pointing to the positive direction of number axis). In the case of negated PET, which is used in this chapter to calibrate the GPD model, $z_p > 0$ means that positive collision frequency is expected for every 1,000,000 CVE, and $z_p \leq 0$ means that there is no collision expected.

5.5.3 Threshold selection

To fit the GPD, the first step is to select a proper threshold and then the other two parameters (σ and ξ) can be estimated. Graphically, the threshold u can be selected by looking at the mean residual life plot and the threshold stability plot (Coles, 2001). The basic steps are as follows:

Step 1: Through the mean residual life plot, an initial range of threshold is selected in which the plot is almost linear.

Step 2: The stability plots are generated from the initial range of threshold. Through the stability plots, a new range of threshold is selected in which the modified scale and shape parameters are constant.

Step 3: Form the new threshold range both the minimum and maximum threshold are used to fit the GPD model. The one that gives the best fit depending on the following criteria is selected as the final threshold;

The criteria's are (Coles, 2001) when:

- $\xi > -0.5$, maximum likelihood estimators are regular,
- $-1.0 < \xi < -0.5$, , maximum likelihood estimators are generally obtainable,
- $\xi < -1.0$, maximum likelihood estimators are unlikely to be obtainable.

5.5.4 Modeling results

The statistical R –package is used in order to generate the two plots as described in the previous steps and to estimate the model parameters. For example, Figures 5-23 and 5-24 show the procedure of the threshold selection and model estimation for the Metcalfe intersection. The summary of the results for all intersections are shown in Table 5-33. As shown in Table 5-33, the return level is calculated based on Equation 5.9.

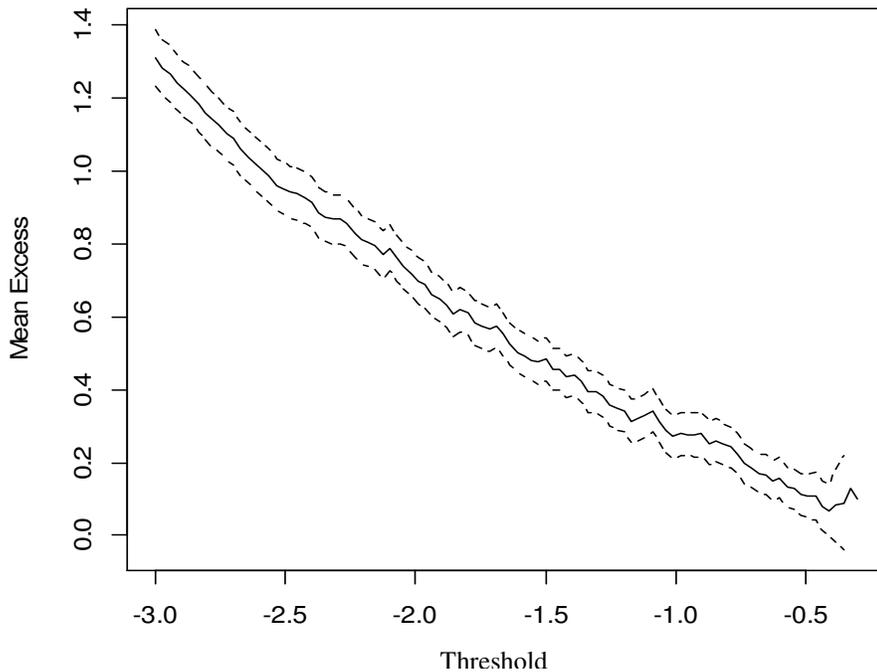


Figure 5-23: Mean residual life plot (Metcalf intersection).

As shown in Figure 5-23, based on visual inspection the mean residual life is substantial curvature in the range -0.7 to -1.2 seconds. The plot is reasonably linear when threshold is less than -1.2 seconds.

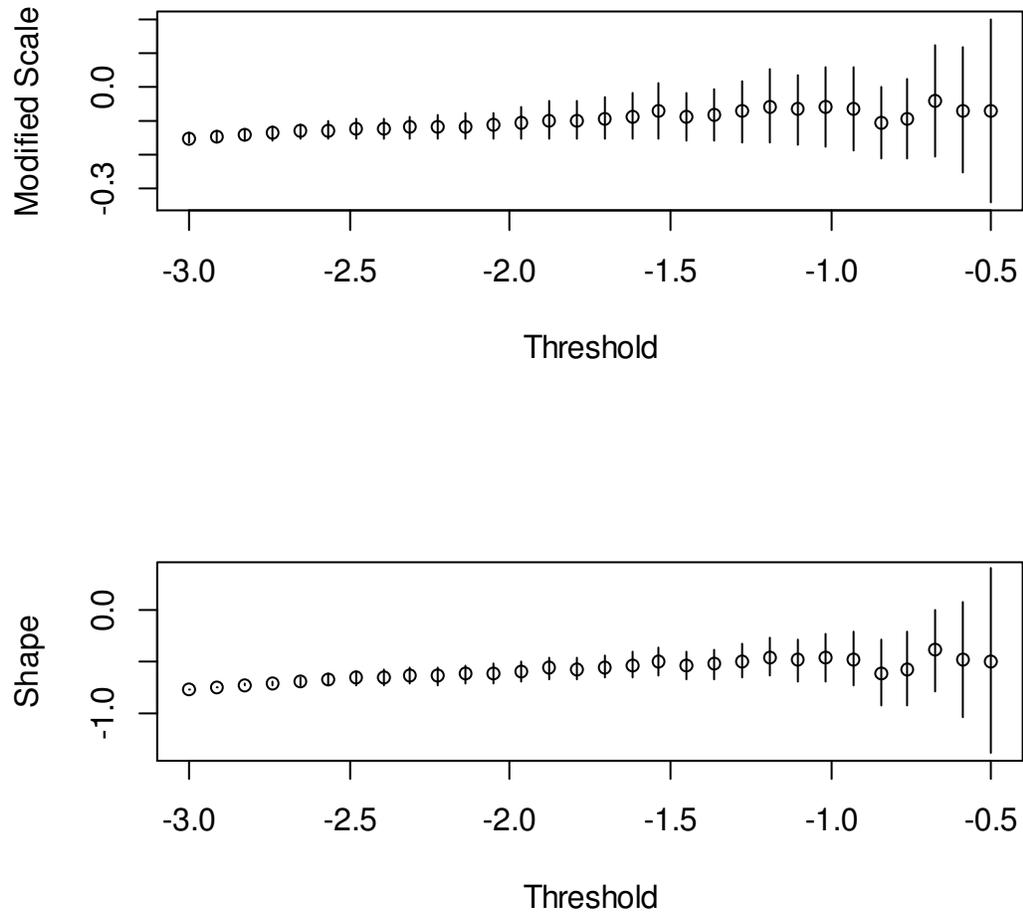


Figure 5-24: Threshold stability plot for scale and shape parameters (Metcalf intersection).

As shown in Figure 5-24, both the modified scale and shape parameters seem to be approximately constant on the range of PET (-1.1, -2.5) seconds. The range of the PET (-1.1, -2.5) seconds, was tested. Depending on step 3 in Section 5.6.3, the threshold -1.1 seconds was selected to be the best and final threshold of the Metcalfe intersection. The

number of exceedance (y) was found to be 53. The modified scale and shape parameters (σ and ξ) were (0.55 and -0.59). Figure 5-25 provides a visualization of the behaviour of the standardized residuals from the model (Metcalf intersection). The other intersections are shown in Appendix (D). The figure shows the probability and quantile plot of the exceedance (y) PET. As shown in the figure the points lie close to the unit diagonal.

Table 5-33: Summary of estimation results of the return level value at the study area.

Intersection	Threshold (u)	Scale (σ)	Shape (ξ)	N (CVE)	Exceedance (y)	n (CVE)	Return level value (z_p)
Metcalf	-1.10	0.54	-0.58	1,000,000	53	242,605	-0.21
O'Connor	-1.58	0.94	-0.66	1,000,000	79	146,628	-0.17
Bank (eastbound)	-1.27	0.68	-0.59	1,000,000	46	80,877	-0.14
Bank (westbound)	-2.12	1.56	-0.88	1,000,000	49	42,296	-0.37
Kent	-1.18	0.56	-0.55	1,000,000	183	304,426	-0.19
Lyon	-2.2	1.04	-0.45	1,000,000	66	32,935	-0.04
Bay	-1.05	0.56	-0.65	1,000,000	20	106,973	-0.22
Somerset	-1.63	0.68	-0.54	1,000,000	73	71,181	-0.41
Sussex	-1.24	0.40	-0.29	1,000,000	56	86,559	-0.07

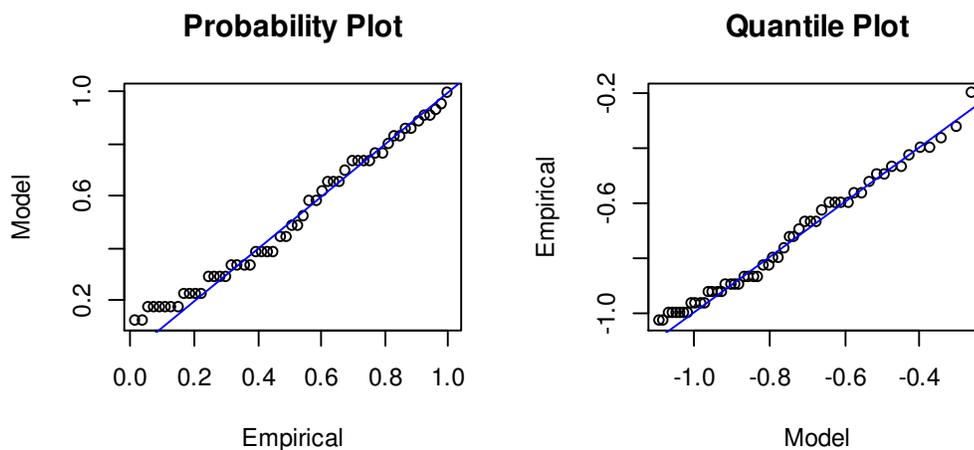


Figure 5-25: Probability and quantile plot (Metcalf intersection).

As shown in the Table 5-33, all return level values are less than zero that means there is no collision expected for all intersections for 1,000,000 CVE.

5.5.5 Correlation analysis

In order to study the strength and direction (positive or negative) of the correlation between right-turn interaction rates, return level values and cyclist violation rates among the intersections, Spearman's correlation rank was used and Spearman's correlation rank coefficient (r_s) was conducted. Mathematically, the Spearman's correlation rank coefficient can be defined as:

$$r_s = 1 - \left(\frac{6 * \sum d^2}{n^3 - n} \right) \quad (5.10)$$

where, d is the difference between two ranks, n , is the number of pairs of data.

The value of r_s should be between +1 (perfect positive correlation) and -1 (perfect negative correlation). The closer the value is to 0, the weaker the correlation. Table 5-34 shows the summary of the right-turn interaction rates, return level values and violation rates. Table 5-35 shows the results of Spearman's correlation rank coefficient.

Table 5-34: Summary of the right-turn interaction rates, intersections return level values and violation rates.

Intersection	Right-turn interaction rate *	Cyclist violation rate ***	Return level value
Metcalfe	1.16	14.29	-0.21
O'Connor	1.02	23.86	-0.17
Bank (eastbound)	2.28	13.23	-0.14
Bank (westbound)	1.68	22.72	-0.37
Kent	1.73	15.24	-0.19
Lyon	2.73	26.79	-0.04
Bay	1.84	39.09	-0.22
Somerset	2.23	15.38	-0.41
Sussex	1.94	40.10	-0.07

* The right-turn interaction rate is the ratio between the summations of all right-turn interactions for a specific intersection by the summation of all CVE for all videos for that intersection.

** Cyclist violations that are involved in interaction for example; against traffic, stopping after stop line, out-of-path stop, riding out-of-path, and cross stop line in red signal.

***The cyclist violation rate is the ratio between the summations of violation for a specific intersection by the total through cyclist volume for all videos for that intersection.

Table 5-35: Spearman's correlation rank coefficient results.

Correlation	Spearman's correlation rank coefficient
Right-turn interaction rate versus return level value	0.43
Right-turn interaction rate versus cyclist violation rate	0.08
Return value versus cyclist violation rate	0.22

As shown in Table 5-35, a positive moderate correlation between right-turn interaction rate and return level value of PET. This indicates that when the right-turn interaction rate increases the return level value of PET between motor vehicles and cyclists increase and will approach to zero or positive value. Therefore, when the return level value equals zero that means there is a collision expected among the intersections.

A positive weak correlation between return level value of PET and cyclist violation rate was shown in Table 5-35. This indicates that when the cyclist violation rate increases the return level value of PET between motor vehicles and cyclists increase and will approach to zero or positive value. Therefore, collision will be expected among the intersections.

The correlation between the right-turn interaction rate and the average cyclist violation per type were studied and the results are shown in Table 5-36.

Table 5-36: Spearman’s correlation rank coefficient results: Right-turn interaction versus the average cyclist violation per type.

Correlation	Spearman’s correlation rank coefficient	Association between variables
Right-turn interaction rate versus against traffic	0.67	large positive correlation
Right-turn interaction rate versus stopping after stop-line	0.02	minimal positive correlation
Right-turn interaction rate versus out-of-path stop	0.63	Large positive correlation
Right-turn interaction rate versus riding out-of-path	0.15	small positive correlation
Right-turn interaction rate versus cross stop-line in red signal	-0.04	minimal negative correlation

5.6 Summary

This chapter presented statistical models which can be used to evaluate the cyclist safety. These statistical models can assist in achieving proposing a ranking of intersections based on safety using non-collision, or surrogate, measures of safety. The surrogate measures of safety used in this chapter are PET measurements and non-

conformal traffic behaviour represented by traffic violations. The relationship between different candidate exposure measures and the observed number of interactions were tested. The best correlate was found between the observed interaction count and cyclist-vehicle events (CVE). The cyclist-vehicle events is the product of the right-turn vehicle volume in every observed video subsequence and the through cyclist volume in the same subsequence.

A total of 18,352 cyclists and 48,632 vehicles were observed at the selected intersections. Out of these, a total of 15,150 through cyclists, 31,604 right-turn and 17,028 left-turn vehicles were observed. The observations were made within a total period of 44 days covering the nine intersections and 707 videos were recorded.

Paired two-tail t-test was applied in order to test the statistical significance of various relationships between the data collected in the fall and summer seasons. Results showed that the average rate of right-turn or left-turn interactions (PET less or equal 3.0 seconds) for each video did not change with the different seasons at the study area. As well, the overall rate of right-turn interactions for each intersection doesn't change with the different seasons at the study area. On the other hand, the proportion of cyclists (right-turn or left-turn interactions /through cyclist movement) involved in a right-turn or left-turn interactions changes with the different seasons at each intersection.

Chi- square test was applied to test the statistical significance of the difference in the proportion of cyclists involved in a right-turn or left-turn interactions with the different seasons at each intersection. In general, the results showed that there is no statistical significance in all comparisons tested, except between the differences in average hourly cyclist volume.

In order to study the effect of the different individual variables such as treatment, seasons, and traffic volume on PET count, Poisson and Negative Binomial regression were applied under different model specifications. For all models, the classic Negative Binomial regression proved to be better than the Poisson model. The dependent variable in these models was PET count, while cyclist through volume, right-turn vehicle volume, CVE, Log (CVE), log (cyclist through volume), log (right-turn vehicle volume), treatment, and seasons were used as independent variables. In addition, the models containing the logarithm of CVE were found to provide the best model.

The correlation between the right-turn interaction rate and the average cyclist violation per type was studied. The results showed that a large positive correlation was found between the right-turn interaction rate versus against-traffic and out-of-path stop violations. A minimal positive correlation was found between the right-turn interaction rate versus stopping after stop-line and cross stop-line in red signal violations. In addition, a small positive correlation was found between the right-turn interaction rate and riding out-of-path violation.

The study presented attempts to examine the effect of pedestrian volume on cyclist-vehicle interactions after accounting for the effect of cyclist and vehicle volumes. Pedestrian volume was found to have a positive (desirable) impact on cyclist safety making the vehicle to yield to the cyclist when vehicle is making a right turn and crosses the cyclist path. This was clear on the PET proximity level (0, 2] and (0, 3] seconds. However, when motorists failed to yield to cyclists, the effect was not significant. The most important effect found is that the increase in pedestrian volume is associated with reduction in the count of positive PET values below 2 and 3 seconds (after accounting for

variability in cyclist and vehicle volumes). It may be argued that the presence of pedestrians may promote safer crossing of cyclists across right-turn vehicles when motorists decided to yield to cyclists. However, when motorists decided not to yield, the change in pedestrian volume did not have an effect on the count of PET values below 3, 2, and 1 seconds. Thus, it can be concluded that when yielding to cyclists is more prevalent, cyclist safety could be improved when the move along with higher pedestrian volume.

The extreme value theory (EVT) was applied in order to estimate the probability of extreme rare events from frequent events. The results showed that all return level values are less than zero which means that there is no collision expected for all intersections for 1,000,000 CVE. As well, a positive moderate correlation between right-turn interaction rate and return level value of PET. This indicates that when the right-turn interaction rate increases the return level value of PET between motor vehicles and cyclists increases and will approach to zero or positive value. Therefore, when the return level value equals to zero that means there is a collision expected among the intersections.

Different shortcomings occurred when the video data was collected. For example, it was difficult to find a suitable position to install the video cameras to cover all the intersections that have been studied in this chapter during the same period of time. Therefore, whenever the intersections had common data a specific statistical analysis strategy was used.

Three different types of recording cameras were used to collect the video data. After checking the videos, it was found that, videos collected using the mobile and stationary cameras operated by the City of Ottawa were not appropriate to measure PET

between cyclists and motor vehicles using the automated measurement method presented in Chapter Four. This can be attributed to the lower height of these cameras and the video resolution that was not suitable to identify the interaction between road users or to separate road users from each other and track each road user.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary and Conclusions

6.1.1 Summary

This thesis presented computer vision technique (tracker system) that can track road users (motor-vehicles, and cyclists) over time. Furthermore, areas with intersecting paths can be assessed on the basis of data gathered from videos. The primary focus in this thesis was on cyclists, with focus on speed measurement and analyzing cyclist-vehicle interactions.

Chapter Two presented a review of the literature with focus on studying the potential of automatically measuring cyclist speed. The knowledge of cyclist speed is a key parameter in the design of cycling facilities, especially in urban areas. Such cyclist facilities include: control cyclist path (a cyclist lane separated from motorized traffic lanes by marking only), segregated cyclist path (a cyclist lane with physically separated from motorized traffic lanes), shared path with other road users (same lane is shared by cyclists and non-cyclists), and control devices (markings, signs, and signals). For example, cyclist crossing speed, which is the cyclist speed when crossing an intersection, is a key parameter in the timing of traffic signal intervals serving heavy cyclist volumes and/or dedicated exclusively for cyclists. The lengths of green and clearance intervals provided by a traffic signal to cyclists to cross an intersection would clearly depend on the cyclist crossing speed. Furthermore, cyclist crossing speed is an important factor in examining cyclist safety at intersections because it is a major factor affecting the duration of exposure to the risk of collision with motor vehicles at the intersection.

Several techniques can be used to measure the speed of cyclists, *e.g.*, manual, semi-automated and automated techniques. Different devices can be used to measure the speed of cyclists, *e.g.*, radar gun, stopwatch, pneumatic tube counter, Global Positioning System (GPS), bike speedometer, and video image analysis. The majority of the studies reviewed in this thesis which focused on measuring cyclist speed were based on manual speed measurements; either in the field or from video recordings. Selecting an appropriate technique for speed measurement appears to depend on several factors, *e.g.*, sample size, required accuracy, and cost. Computer vision technique offers an appealing solution for accurate and large-scale speed measurement.

The sampling strategy followed in most reviewed studies by selecting cyclists may not be representative of the general population. The observed cyclist might have been aware of being monitored. The cyclist's who were provided with any type of tracking devices; for example, GPS and speedometer; for the purpose of collecting speed data, the technique can be argued that the data collection strategy may not yield naturalistic behaviour. In order to obtain the naturalistic cyclist behaviour, it is preferred not to interfere with road users and without any prior notice.

As previous studies have been reviewed for this study, additional judgments have been made regarding device's limitations and usage in techniques, as referred to in the following sentences.

The disadvantages in using a radar gun to measure the cyclist speed are, [i] less reliability in heavy traffic, [ii] impossible to re-check data measurements, and [iii] difficult to determine the accuracy due to the cosine error. The cosine error is defined as the angle between the line sight of the radar and the direction of the road users (Peterson,

2007). Therefore, the cosine error increases in the case of speed measurement on a curvilinear section. The disadvantages of using a pneumatic tube sensor for measuring cyclist speed are: [i] it is difficult to classify the cyclist by gender or age, [ii] limited coverage area, [iii] and it is difficult to re-check data measurements. To measure cyclist speed using an experimental bike and stopwatches may require more than two observers. One observer riding the experimental bike and two observers positioned at the beginning and at the end of a cyclist path to record the crossing time. The experiment may require more than one trail, depending on the accuracy of the observers recording. In this case the behaviour of the cyclist may not be naturalistic as the trial is repeated.

The accuracy of the GPS data depends on [i] number of satellites, [ii] changing satellite geometry, [iii] the continuity of the satellite signal capture, and [iv] cyclist position on a straight or a curvature. If the GPS is set up, so that it stores observations by real time, however, useless data will be collected when the cyclist is not moving. Therefore, to extract the collected observations from useless information will require extra efforts. It is known that the positioning accuracy can be impacted within urban canyons and in the vicinity of high buildings. Furthermore, GPS is an intrusive device for speed measurement as it is specific to cyclists participating in the survey.

The automated cyclist speed measurements using computer vision techniques might overcome above mentioned limitations. It is important to note that using the computer vision technique with a video recording device will help in providing other opportunities, *e.g.*, traffic interactions between road users, and road user violations.

In order to reduce the number of interactions between cyclists and motor-vehicles and to improve cyclist safety, different treatments have been implemented and evaluated

in several reviewed studies. For example; marking lanes, segregated cyclist lanes, advanced Stop Lines, bike-boxes, colour cycle crossings, signal phase for cyclists and different control warnings of cyclist's signs.

For example, in the reviewed studies, to provide more protection of exposure to traffic lane was found after adding segregating cyclist lane. Also, the number of interactions between cyclists and motor-vehicles was reduced after adding bike-boxes. In addition, the rate of vehicles encroachment further decreased significantly after bike-boxes were coloured.

Different cyclist-motor collisions scenarios were investigated in the reviewed studies. The results of reviewed studies showed that the majority of collisions happened when a motor vehicle driver turning right failed to notice the cyclist coming from the right side of motor vehicle driver along cyclist crossing.

Different surrogate measures of safety can be used to evaluate the safety of road users. The assessment of road safety based on collision data would introduce several possible shortcomings, such as; [i] collisions are relatively rare events, [ii] underreporting of collisions, [iii] collision data does not typically provide complete information about the collision process. Traffic conflict is considered as one of the key surrogate measures of safety, among other measures. Recently, automated data collection using video techniques were introduced in order to improve the accuracy of traffic conflict data collection and to reduce the analysis time. Traffic interaction studies depending on manual counting are time-consuming, labour-intensive and prone to observe inaccuracies. In order to eliminate the subjectivity from traffic interaction analysis, different objective proximity indicators have been suggested as surrogate measures of safety.

Chapter Three presented the technical details of comprehensive evaluation of computer vision technique to measure cyclist field-speed at signalized intersections. The average crossing speed calculated in this study is based on the average of all displacements of the constituent features (cyclist's features). The results of the video-measured speed of the cyclist are generally acceptable in terms of the accuracy (average percentage of error). This demonstrates that video-measured speed can be performed using computer vision techniques.

Two different validation approaches were performed to measure the accuracy of video-measured speeds. The differences in the average field-measured and video-measured speeds for both validation approaches were 0.18 m/s (16 cyclists), and 0.16 m/s (483 cyclists). The speed results was compared to another independent study by (Zaki *et al.*, 2013), which, also found that the difference between the average field-measured and the average video-measured speeds was 0.19 m/s (70 cyclists). However, (Zaki, *et al.*, 2013) validated the speed on a straight segment (14 meter length). In this study, turning movement with different geometry, including, straight and curve segments were included in the validation process.

The overall average speed error (absolute speed difference divided by field-measured speed) for both validation approaches were 4.83% (16 cyclists) and 4.98% (483 cyclists). A high coefficient of determination, R^2 , of 0.97 was found between field-measured and video-measured speeds. Other studies such as (Zaki, *et al.*, 2013), found lower yet still acceptable agreement between the field-measured and video-measured speeds, with a coefficient of determination, R^2 , of 0.86.

The Generalized Extreme Value was found to be the best fit distribution for both field-measured and video-measured speeds. This shows consistency in distribution between field-measured and video-measured speeds. This similarity in the distribution confirms the possibility of using the video-measured technique to measure the cyclist speed. Statistical analysis was conducted to investigate the effect of different factors on cyclist speed.

Factors such as the longitudinal grade (uphill and downhill), and acceleration significantly influence the cyclist field-measured and video-measured speeds for the validation events (483 events). This similarity in the outcome provides evidence as to the reliability of the video analysis technique. The study did not take into account characteristics such as the cyclist age, and gender classification. This limitation is according to the accuracy of the manual video observation. In addition, this study is non-intrusive related to the position of the camera that was placed on the 12th floor of the building.

Limitations of Chapter Three

This study is non-intrusive with regard to the position of the camera, which was placed on the 12th floor of the building. However, different shortcomings occurred due to the position of the video camera. The video camera was laterally far from the study location and the angle of the camera made it challenging to obtain the tracks of road users using the tracker system when there was traffic congestion. During the traffic congestion, some cyclists would disappear behind vehicles as well as, when cyclists crossed the intersection within a group of cyclists. These shortcomings can be attributed to using improper or suboptimal configuration tracking parameters. A careful consideration of this

issue was conducted in this study, but the problem of over-grouping still occurred. Over-grouping is defined as two or more different objects are grouped or merged, incorrectly, by the tracker system and considered as one object. In addition, the study did not take into account characteristics that were not easy to determine in the videos such as the cyclist age and gender.

Chapter Four presented a technique for automated measurement of a proximity measure of cyclist-vehicle interactions. The chapter presented an automated video analysis technique to measure the Post-Encroachment Time (PET) between cyclists and motor vehicles at signalized intersections. Three measurement methods were used to measure PET: [i] manual video timer measurement (MVTM) method, [ii] manual frame count measurement (MFCM) method and [iii] automated measurement (AM) method. The results of recorded cyclist-vehicle interactions with PET less than or equal to 3.0 second showed that the average PET value using the MVTM method was 2.53 seconds with a standard deviation of 0.64 second. The average PET value using MFCM method was 1.77 seconds with a standard deviation of 0.60 second. The average PET value using the AM method was 1.83 seconds with a standard deviation of 0.60 second. Depending on the lowest value of the Log-likelihood, the best-fit distribution of measured PET was identified. The Generalized Extreme Value was found to be the best fitting distribution for both MFCM and AM methods.

The AM method was evaluated based on the difference between its measurements and those obtained from the MFCM method. The results showed that the PET values obtained from both methods were very close. A high coefficient of determination, R^2 , was found between the AM and MFCM methods. The average value of the difference

(signed) in PET was -0.06 second over the range of PET difference from -0.56 to 0.43 seconds. The average absolute difference (unsigned) between AM and MFCM methods was 0.12 second. This is a very promising finding given that the precision is 0.033 second. Furthermore, there was a very good agreement between AM and MFCM methods in PET counts at three different proximity levels: levels of (0, 1], (1, 2] and (2, 3] seconds. On the other hand, agreements between the MVTM and both the AM and the MFCM were not as strong for the three severity levels. The AM method was proven to be accurate and reliable as demonstrated by comparing the results obtained using three different measurement methods.

The AM method was proven to be accurate and reliable as demonstrated by comparing the results obtained using three different measurement methods. This technique has a promising practical benefit as it can be used to measure the PET with high accuracy while taking into account the physical extremities (rear and front parts) of the analyzed road users. The proposed technique also objectively defines the potential area of collision as the point where the paths of the interacting vehicle and cyclist intersect. Thus, it removes the need to subjectively define the potential area of collision or visually determine the location where the paths of the cyclist and the vehicle intersect for the purpose of measuring PET. The proposed technique requires very little human intervention and therefore can substantially help users save time and effort and reduce human errors. Overall, the time difference required for calculating the PET for processing all events using MFCM and AM methods in this study is estimated to be (18 hours). Therefore, the process time is decreased by 67% of the time when using the AM method.

Practically, the proposed technique can support large-scale data collection for investigating the relationship between PET counts and collisions.

Limitations of Chapter Four

Different shortcomings occurred while measuring the PET between cyclists and motor vehicles using the tracker system. For example, there were some bounding boxes which partially covered road users. However, in case of partial coverage, the part of road users covered by the bounding box is used to calculate PET and this may explain why there is residual error in automated PET calculation. Although the tracker system is improved and most trajectories are detected by the tracker system, over-segmentation still occurred especially on long vehicles. Over-segmentation, which means that more than one object, can be seen representing real road users. In addition, other factors were found to impact the quality of the tracks, such as, shades and other disruptive textures.

In Chapter Five, different statistical techniques were conducted which can be used to evaluate the cyclist safety at signalized intersections. For the purpose of studying cyclist safety, video data was collected at nine selected intersections in the Downtown of Ottawa, Ontario. The intersections were selected based on feedback from the City of Ottawa. The study area consists of two groups of intersections, the first group is treatment (segregated) intersections, the second group of intersections is control (un-segregated) intersection. Focus was allocated to interactions between cyclists and right/left-turn vehicles at intersections. Different traffic movement of cyclists and motor vehicles were observed and naturalistic movements of road users within these intersections were recorded. A number of performance parameters were identified in this study. The performance parameters are related to interactions between cyclists and vehicles and

individual cyclist behaviour. The main interaction pattern is the temporal proximity of cyclist and right-turn vehicles while crossing each other's paths. A number of categories of non-conformal traffic movements were identified and observed for cyclists at the nine intersections.

A total of 18,352 cyclists and 48,632 vehicles were observed at the selected intersections. Out of these, a total of 15,150 through cyclists, 31,604 right-turn and 17,028 left-turn vehicles were observed. The observations were made within a total period of 44 days covering the nine intersections and 707 videos were recorded.

Paired two-tail t-test was applied in order to test the statistical significance of various relationships between the data collected on Fall and Summer seasons, *e.g.*, [i] the difference in average hourly cyclist volume, [ii] the difference in average of cyclist right-turn interactions (the average = $((\text{right-turn interactions} / \text{CVE}) * 1000)$ for each video. $\text{CVE} = \text{total of right-turn vehicle volume} * \text{total of cyclist through volume}$ for each video, [iii] the difference in average of cyclist right-turn interactions for each intersection (the average = $((\text{total right-turn interactions} / \text{total CVE}) * 1000)$, [iv] the difference in average of cyclist left-turn interactions (the average = $((\text{left-turn interactions} / \text{CVE}) * 1000)$ for each video. $\text{CVE} = \text{total of left-turn vehicle volume} * \text{total of cyclist through volume}$ for each video, [v] the difference in the proportion of cyclist involved in interactions ($(\text{proportion} = \text{total right-turn/or left-turn interactions} / \text{total through cyclist movement}) * 100$), and [vi] the difference in cyclist violation rate ($(\text{violation rate} = \text{total violation} / \text{total cyclist volume}) * 100$).

In addition, Chi- square test was applied to test the statistical significance of the difference in the proportion of cyclists involved in a right-turn or left-turn interactions

with the different seasons at each intersection. In general, the results showed that there is no statistical significance in all comparisons tested, except between the differences in average hourly cyclist volume.

In order to study the effect of the different individual variables such as treatment, seasons, and traffic volume on PET count, Poisson and Negative Binomial regression were applied under different model specifications. For all models, the classic Negative Binomial regression proved to be better than the Poisson model. This is consistent with crash prediction theory. The dependent variable in these models was PET count, while cyclist through volume, right-turn vehicle volume, CVE, Log (CVE), log (cyclist through volume), log (right-turn vehicle volume), treatment, and seasons were used as independent variables. Furthermore, for all proximity thresholds, the best model based on the lower AIC, was the model which uses the natural logarithm of CVE as the measure of exposure. In addition, the effect and the significance of each intersection on PET counts was tested based on p -value and the results were considered negligible.

The effect of the pedestrian volume on cyclist safety was conducted. Pedestrian volume was found to have a positive (desirable) impact on cyclist safety making the vehicle to yield to the cyclist when vehicle is making a right turn and crosses the cyclist path. This was clear on the PET proximity level (0, 2] and (0, 3] seconds. However, when motorists failed to yield to cyclists, the effect was not significant. The most important effect found is that the increase in pedestrian volume is associated with reduction in the count of positive PET values below 2 and 3 seconds (after accounting for variability in cyclist and vehicle volumes). It may be argued that the presence of pedestrians may promote safer crossing of cyclists across right-turn vehicles when motorists decided to

yield to cyclists. However, when motorists decided not to yield, the change in pedestrian volume did not have an effect on the count of PET values below 3, 2, and 1 seconds. Thus, it can be concluded that when yielding to cyclists is more prevalent, cyclist safety could be improved when the move along with higher pedestrian volume.

The extreme value theory (EVT) was applied in order to estimate the probability of extreme rare events from frequent events. The results showed that all return level values are less than zero, which means that there is no collision expected for all intersections for 1,000,000 CVE. As well, a positive moderate correlation between right-turn interaction rate and return level value of PET. This indicates that when the right-turn interaction rate increases the return level value of PET between motor vehicles and cyclists increases and will approach to zero or positive value. Therefore, when the return level value equals to zero that means there is a collision expected among the intersections

The correlation between the right-turn interaction rate and the average cyclist violation per type was studied. The results showed that a large positive correlation was found between the right-turn interaction rate versus against-traffic and out-of-path stop violations. A minimal positive correlation was found between the right-turn interaction rate versus stopping after stop-line and cross stop-line in red signal violations. In addition, a small positive correlation was found between the right-turn interaction rate and riding out-of-path violation

Limitations of Chapter Five

Different shortcomings occurred when the video data was collected. For example, it was difficult to find a suitable position to install the video cameras to cover all the intersections that have been studied in this chapter during the same period of time.

Therefore, whenever the intersections had common data a specific statistical analysis strategy was used.

Three different types of recording cameras were used to collect the video data. After checking the videos, it was found that, videos collected using the mobile and stationary cameras operated by the City of Ottawa were not appropriate to measure PET between cyclists and motor vehicles using the automated measurement method presented in Chapter Four. This can be attributed to the lower height of these cameras and the video resolution that was not suitable to identify the interaction between road users or to separate road users from each other and track each road user.

A common limitation in all chapters which utilized the tracker system is that the classification of cyclist tracks was performed manually. That is, the road user generating a track is observed in the video and classified as a vehicle or a cyclist. Automating this process can prove useful in processing larger scale of data.

6.1.2 Conclusions

Based on the previous summary and the finding of this study, the following can be concluded:

- 1- This study presents a reliable video analysis technique to measure cyclist speed at signalized intersections. The developed technique enables the automated observation of large volumes of naturalistic cyclist movements in an accurate and resource-efficient manner.
- 2- A very good fit was found between field-measured and video-measured speeds, with a high coefficient of determination, R^2 , of 0.97. This result of the coefficient

of determination can be considered as a validation of the video-measured technique to measure cyclist speed.

- 3- Different distributions were conducted and based on the lowest value of the log likelihood the Generalized Extreme Value was found to be the best fit distribution for both field-measured and video-measured speeds. This shows consistency in distribution between field-measured and video-measured speeds.
- 4- A very weak correlation (-0.1119) exists between field-measured speed and error value between cyclist field-measured and video-measured speeds. This provides evidence that the accuracy of the video analysis technique does not deteriorate at higher speeds.
- 5- Analysis of Variance (ANOVA) was applied, and based on p -value, factors such as the longitudinal grade (uphill and downhill), and crosses without stopping were found to be significantly influencing the cyclist field-measured speed. The same factors were found to be significant based on analysis of video-measured speed for the same validation events.
- 6- Side judgments indicate due to specific environmental and field conditions, such as, weather, and high buildings, it is thought difficult to measure the cyclist speed using certain techniques and equipment. As well, it is difficult to measure the extent of the behaviour of the cyclist using intrusive devices, *e.g.*, radar gun, experimental bike and GPS. In addition, it is difficult to reduce the cosine error when using the radar gun to measure the cyclists speed, this error increases on a

curve due to the large cosine effect angle and changing proportional on curve radius and cyclist speed.

- 7- Video monitoring is less susceptible to the impact of high physical objects, *e.g.*, high-rise buildings, which abound in urban environments (as compared to traditional GPS tracking techniques). As well, monitoring movements of road users are non-intrusive, thus avoiding any necessary participation of observed subjects. Observed subjects are typically unaware of video monitoring.
- 8- Cyclist speed can be measured using video camera along straight, as well as, curved paths, as compared to other devices. Furthermore, video cameras offer an inexpensive and mobile data collection platform which provides a permanent record of observations. Video analysis permits continuous tracking of cyclists to calculate a sequence of instantaneous speed observations over a finite period of time.
- 9- This study presented an automated video analysis technique (automated measurement method) to measure the Post-Encroachment Time (PET) between cyclists and motor vehicles.
- 10- The automated video analysis technique has a promising practical benefit as it can be used to measure the PET with a high degree of accuracy, while taking into account the physical extremities (rear and front parts) of the analyzed road users.
- 11- The automated video analysis technique, also, objectively defines the potential area of interaction as the point where the paths of the interacting motor vehicle

and cyclist intersect. Thus, it removes the subjectivity to define the potential area of interaction for the purpose of measuring PET.

12- Practically, the automated video analysis technique can support large-scale data collection for investigating the relationship between PET counts and collisions. In addition, this technique has a great potential for use in behavioural studies, especially when the studied events are relatively infrequent.

13- Among different levels of video camera positions that been selected in order to collect the video data for this study, the higher level of video camera positioned with respect to the level of road users provided more information that been analyzed using the developed technique. However, the developed technique is applicable for different levels of camera positions.

14- The average of cyclist right-turn or left-turn interactions (PET less or equal 3.0 seconds) for each video does not change with the different seasons at the study area. On the other hand, the proportion of cyclists involved in a right-turn or left-turn interactions changes with the different seasons at each intersection.

15- As all estimated return level values are less than zero, which this means that there is no collision expected for all intersections for 1,000,000 CVE. As well, there is a positive moderate correlation between right-turn interaction rate and return level value of PET. This indicates that when the right-turn interaction rate increases, the return level value of PET between motor vehicles and cyclists also increases and will approach a zero or positive value. Therefore when the return level value

equals to zero, this means that there is a collision expected among the intersections.

16- Pedestrian volume was found to be associated with a positive (desirable) impact on cyclist safety when vehicles yielded to the cyclists while making a right turn movement across the cyclist path. This was statistically significant for PET proximity levels (0, 2] and (0, 3] seconds. However, when motorists failed to yield to cyclists, the effect was not significant. Furthermore, the variability in pedestrian volume was not associated with vehicles blocking cyclist path while making a right-turn movement. It may be argued that the increased presence of pedestrians may promote safer crossings of cyclists across right-turn vehicles when motorists yield to cyclists.

17- A large positive correlation was found between the right-turn interaction rate versus against traffic and out-of-path stop violations. Also, a minimal positive correlation was found between the right-turn interaction rate versus stopping after stop-line and cross stop-line in red signal violations. A small positive correlation was found between right-turn interaction rate and riding out-of-path violation.

6.2 Recommendations and Future Work

Based on the results from this study, the following can be recommended:

1- The presented technique can benefit from automated road user classification (cyclist or vehicle), thus improving the level of automation. Furthermore, the presented technique offers the potential and flexibility to be expanded to consider other road users. This technique can be used to develop and validate traffic flow

models by analyzing traffic scenes to estimate different microscopic and macroscopic traffic flow measures.

- 2- The presented technique demonstrated the capability of using automated video analysis to detect and track road users in traffic scene. In addition, there is ability to identify the important events that may lead to collisions.
- 3- The presented technique is suitable to calculate several conflict indicators other than the Post-Encroachment Time (PET) that have been used in this study.
- 4- Using the computer vision techniques offers a suitable solution to demands for more efficient and accurate methods of cyclist data collection.
- 5- Statistical analysis was conducted in order to investigate the impact of several factors on cyclist speed such as weather conditions, roadway conditions, time of the day, longitudinal grade slope, and crosses without stopping. The results of this analysis provide useful information into the considerations required for the design of cyclist facilities under different operations and physical conditions.
- 6- Throughout the process of collecting video data in this study, three different levels of cameras were positioned. It is recommended to use the level that will provide the most information and that which would make it easy to identify all road users of the study area.
- 7- Cyclist crossing speed is an important factor in examining cyclist safety at intersections because it is a major factor affecting the duration of exposure to the risk of collisions with motor vehicles at the intersections. It is recommended to monitor the movements of road users using video camera because it is more

suitable than other devices used by the previous studies In addition, the video camera records the variability of cyclist speed with time that can be measured along straight as well as curved paths.

References

- Akira, Y., Hiroshi, K. & Katsumi, U., 2012. *A Study of Bicycle Travel Speed*. Traffic Engineerin Division, Road Department.
- Allen, B., Shin, B. & Cooper, P., 1978. Analysis of Traffic Conflicts and Collisions. *Transportation Research Record*, Volume 667, pp. 67-74.
- Allen, D., Bygrave, S. & Harper, H., 2005. *Behaviour at Cycle Advance Stop Lines, Prepared for Transit for London*, London, England.
- Amundsen, F. & Hydén, C., 1977. *First Workshop on Traffic Conflicts*. Oslo, Norwegian Council for Scientific and Industrial Research.
- Archer, J., 2004. *Methods for the Assessment and Prediction of Traffic Safety at Urban Intersections and their Application in Micro-simulation Modeling*. Stockholm: Royal Institute of Technology.
- Ballard, D. & Brown, C., 1982. *Computer Vision*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Bradski, G. & Kaehler, A., 2008. *Learning OpenCV: Computer Vision with the OpenCV Library*, Oreilly.
- Caries, S., 2009. A Comparitive Simulation Study of the Annual Maxima an dthe Peaks-over-threshold Methods. *Delft University of Technology*.
- Chin, H. C. & Quek, Q., 1997. Measurement of Traffic Conflicts. *Safety Science*, 26(3), pp. 169-185.

Christopher, C. & Min, H., 2009. *Alternative Methods of Measuring Operating Speed of Electric and Traditional Bikes in China-Implications for Travel Demand Models*, Proceeding of the Eastern Asia Sociatey for Transportation Studies.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1).

Coifman, B., Beymer, D., Mclauchlan, P. & Malik, J., 1998. A Real-Time Computer Vision System for Vehicle Tracking and Traffic Surveillance, *Transportation Research Part C*, Volume 6, pp. 271-288.

Coles, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London, UK.

Cooper, P. J., 1984. Experience with Traffic Conflicts in Canada with Emphasis on “Post Encroachment Time” Techniques. In: E. Asmussen, ed. *International Calibration Study of Traffic Conflict Techniques*. Springer Berlin Heidelberg, pp. 75-96.

Davies, D., Chinn, L., Buckle, G. & Reid, S., 2003. *Cycling in Vehicle Restricted Areas*, UK.

Dill, J. & Gliebe, J., 2008. *Understanding and Measuring Bicycling Behavior: a focus on travel time and route choice*, Portland: Oregon Transportation Research and Education Consortium (OTREC).

Dill, J., Monsere, C. M. & McNeil, N., 2012. Evaluation of Bike Boxes at Signalized Intersections. *Accident Analysis and Prevention*, 44(1), pp. 126-134.

- Duane, S., McIntyre, J., Anderson, M. & Moreau, S., 2002. *Handbook of Simplified Practice for Traffic Studies*, Ames: Iowa DOT Project TR-455.
- El-Geneidy, A., Krizek, K. J. & Iacono, M., 2007. Predicting Bicycle Travel Speeds Along Different Facilities Using GPS Data: a proof of concept model. Washington DC., *Transportation Research Board Annual Meeting*.
- Famoye, F. & Singh, K., 2006. Zero-Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data, *Journal Data Science*, Volume 4, pp. 117-130.
- Federal Highway Administration, 2011. *Pedestrian and Bicycle Data Collection*. H PP1-30.
- Gettman, D. & Head, L., 2003. *Surrogate Safety Measures From Traffic Simulation Models*. Federal Highway Administration, FHWA-RD-03-050.
- Gettman, D., Pu, L., Sayed, T. & Shelby, S., 2008. *Surrogate Safety Assessment Model and Validation: Final Report*, U.S. Department of Transportation.
- Giancarlo, B., Aluisio, B. J., Janaina, S. V. d. & Denise, G. P., 2010. Cycling to Work in Brazil: Users Profile, Risk Behaviors, and Traffic Accident Occurrence. *Accident Analysis and Prevention*, Volume 42, pp. 1025-1030.
- Google, 2011. Available at: <https://www.google.ca/maps/place/Mackenzie+King+Bridge,+Ottawa,+ON/@45.4237468,-75.6904941,17z>, [Online, Accessed 2011].
- Google, 2013. Available at: <https://www.google.ca/maps/preview/@45.4180604,-75.6989937,17z>, [Online, Accessed 2013].

- Google, 2011. *Feature Detection*. Available at: http://docs.opencv.org/modules/imgproc/doc/feature_detection.html, [Online, Accessed January 2011].
- Google, 2011. *opencv-feature-tracker*. Available at: <http://code.google.com/p/opencv-feature-tracker>, [Online, Accessed 2011].
- Gray, A. J., Jenkins, D., Andrews, M., Mark, H., Taaffe, D. R. & Glover, M. L., 2010. Validity and Reliability of GPS for Measuring Distance Travelled in Field-Based Team Sports. *Journal of Sports Sciences*, 28(12), pp. 1319-1325.
- Hayward, J. C., 1972. Near-miss Determination Through Use of a Scale of Danger. *Highway Research Record*, 384(6), pp. 24-34.
- Huang, F., Yu, H. & Wang, W., 2012. Identifying if VISSIM Simulation Model and SSAM Provide Reasonable Estimates for Field Measured Traffic Conflicts at Signalized Intersections. *Accident Analysis and Prevention*, Volume 50, pp. 1014-1024.
- Huang, L. & Jianping, W., 2004. A Study on Cyclist Behavior at Signalized Intersections. *IEEE Transactions on Intelligent Transportation Systems*, 5(4), pp. 293-299.
- Huang, P. & Pant, P. P., 1994. Simulation-Neural Network Model for Evaluating Dilemma Zone Problems at High-Speed Signalized Intersections. *Transportation Research Record*, Volume 1456, pp. 34-42.
- Hunter, M. & Rodgers, M., 2012. *Evaluation Of Intersection Countermeasures on High-Speed Rural Multi-Lane Facilities*. Georgia Department of Transportation.

- Hunter, W. W., 2000. Evaluation of Innovation Bike-Box Application in Eugene, Oregon. *Transportation Research Record*, 1705(00), pp. 99-106.
- Hunter, W. W., Harkey, D. L., Stewart, J. R. & Birk, M. L., 1999. Evaluation of Blue Bike-Lane Treatment in Portland, Oregon. *Transportation Research Record*, 1705, pp. 107-115.
- Hunter, W. W., Srinivasan, R. & Martell, C. A., 2009. *An Examination Of Bicycle Counts and Speeds Associated With The Installation Of Bike Lanes In St. Petersburg, Florida*, North Carolina: Highway Safety Research Center.
- Hydén, C., 1987. The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique. *Lund University of Technology*.
- Ismail, K., 2010. *Application of Computer Vision Techniques for Automated Road Safety Analysis and Traffic Data Collection*. University of British Columbia, Vancouver: PhD Thesis, Department of Civil Engineering.
- Ismail, K., Sayed, T. & Saunier, N., 2013. A Methodology for Precise Camera Calibration for Data Collection Applications In Urban Traffic Scenes. *Canadian Journal of Civil Engineering*, 40(1), pp. 57-67.
- Jensen, S. U., 2008. Bicycle Tracks and Lanes: A Before-After Study. *Transportation Research Record*, 2095(10), pp. 1-10.
- Jensen, S. U., 2008. Safety Effects of Blue Cycle Crossings: A Before-After Study. *Accident Analysis and Prevention*, 40(2), pp. 742-750.

- John, P. & Rotheram, J., 2010. Design Speeds and Acceleration Characteristics of Bicycle Traffic for Use In Planning, Design and Appraisal. *Transportation Policy*, 17(5), pp. 335-341.
- Kendrick, C. M., Moore, A., Haire, A., Bigazzi, A., Figliozzi, M., Monsere, C. M. & George, L., 2011. Impact of Bicycle Lane Characteristics on Exposure of Bicyclists to Traffic-Related Particulate Matter. *Transportation Research Record*, 2247(1), pp. 24-32.
- Khan, A. M. & Langlois, R. G., 2011. Factoring Cycling in Transportation Infrastructure: Design Considerations Based on Risk Exposure. *Institute of Transportation Engineers*, (8), pp. 49-53.
- Kiefer, R., LeBlanc, D. & Flannagan, C., 2005. Developing An Inverse Time-To-Collision Crash Alert Timing Approach Based on Drivers' Last-Second Braking and Steering Judgments. *Accident Analysis and Prevention*, 37(2), pp. 295-303.
- Koorey, G. & Mangundu, E., 2010. *Effects on Motor Vehicle Behavior of Color and Width of Bicycle Facilities at Signalized Intersections*. Washington, DC, Transportation Research Board Annual Meeting.
- Korve, M. J. & Niemeier, D. A., 2002. Benefit-Cost Analysis of Added Bicycle Phase at Existing Signalized Intersection. *Journal of Transportation Engineering*, 128(1), pp. 40-48.
- Laureshyn, A., Svensson, A. & Hydén, C., 2010. Evaluation of traffic safety, based on micro-level behavioural data: theoretical framework and first implementation. *Accident Analysis and Prevention*, 42(6), pp. 1637-1646.

- Lee, J.B., Lee, C., Shin, H. C., Park, J. & Eom, K., 2011. *Measurement of Desirable Minimum Bike Lane Width Using RTK GPS*. Washington DC, Transportation Research Board 90th Annual Meeting.
- Litman, T., Blair, R., Demopoulos, B., Eddy, N., Fritzel, A., Laidlaw, D., Maddox, H. & Froster, K., 2006. *Pedestrian and Bicycle Planning: A Guide to Best Practices*, Victoria Transport Policy Institute, Victoria, Canada.
- Lord, D., 1996. Analysis of Pedestrian Conflicts with Left-Turning Traffic. *Transportation Research Record: Journal of the Transportation Research Record*, Volume 1538, pp. 61-67.
- Lord, D. & Mannering, F., 2010. The Statistical Analysis of Crash-frequency Data: a review and Assessment of Methodological Alternatives. *Transportation Research Part A*, Volume 42, pp. 291-305.
- Loskorn, J., Mills, A. F., Brady, J. F., Duthie, J. & Machemehi, R. B., 2013. Effects of Bicycle Boxes on Bicyclists and Motorists Behavior at Intersections in Austin, Texas. *Journal of Transportation Engineering*, 39(10), pp. 1039-1046.
- Lucas, B. & Kanade, T., 1981. An Iterative Image Registration Technique with an Application to Stereo Vision. *IJCAI*, Volume 130, pp. 121-130.
- Mette, M. & Tove, H., 2008. Cyclists' Perception of Risk in Roundabouts. *Accident Analysis and Prevention*, 40(3), pp. 1055-1062.

- OpenCV, "OpenCV User Guide," 2012. [Online]. Available:
http://docs.opencv.org/doc/user_guide/user_guide.html. [Accessed January 2012].
- OpenCV, "Feature Detection," 2011. [Online]. Available:
http://docs.opencv.org/modules/imgproc/doc/feature_detection.html. [Accessed January 2011].
- Opiela, K. S., Khasnabis, S. & Datta, T., 1980. Determination of the Characteristics of Bicycle Traffic at Urban Intersections. *Transportation Research Record*, Volume 743, pp. 30-38.
- Parker, M. & Zegeer, C., 1989. *Traffic Conflict Technique for Safety and Operations: Observers Manual*, Report FHWA-IP-88-026, FHWA, U.S. Department of Transportation.
- Parkin, J. & Meyers, C., 2010. The Effect of Cycle Lanes on the Proximity Between Motor Traffic and Cycle Traffic. *Accident Analysis and Prevention*, 42(1), pp. 159-165.
- Parkin, J., Wardman, M. & Page, M., 2007. Models of Perceived Cycling Risk and Route Acceptability. *Accident Analysis and Prevention*, 39(2), pp. 364-371.
- Peesapati, L., Hunter, M. & Rodgers, M., 2013. *Evaluation of Post Encroachment Time as a Surrogate for Opposing Left-turn Crashes*. Washington DC, Transportation Research Board Annual Meeting Compendium of Papers.
- Persaud, B. N. & Mucsi, K., 1995. Microscopic Accident Potential Models for Two-Lane Rural Roads. *Transportation Research Record*, Volume 1485, pp. 134-139.
- Peterson, C., 2007. *Driver's Guide to police Radar*. 1st. ed.

- Phillips, R. O., Bjørnskau, T., Hagman, R. & Sagberg, F., 2011. Reduction in Car-Bicycle Conflict at a Road-Cycle Path Intersection: Evidence of Road User Adaptation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(2), pp. 87-95.
- Pickands, J., 1975. Statistical Inference Using Extreme Order Statistics. *Annals of Statistics*, Volume 3, pp. 119-131.
- Rao, V. T. & Rengaraju, V. R., 1998. Modeling Conflicts of Heterogeneous Traffic at Urban Uncontrolled Intersections. *Journal of Transportation Engineering*, 124(1), pp. 23-34.
- Räsänen, M., Koivisto, I. & Summala, H., 1999. Car Driver and Bicyclist Behavior at Bicycle Crossings Under Different Priority Regulations. *Journal of Safety Research*, 30(1), pp. 67-77.
- Räsänen, M. & Summala, H., 1998. Attention and Expectation Problems in Bicycle-Car Collisions: an in-depth study. *Accident Analysis and Prevention*, 30(5), pp. 657-666.
- R_Core_Team, 2012. R: A Language and Environment for Statistical Computing., Vienna: R Foundation for Statistical Computing.
- Rezendes, J., 2006. Bicycle Helmet: Overcoming Barriers to Use and Increasing Effectiveness. *Journal of Pediatric Nursing*, Volume 21, pp. 35-44.
- Sakshaug, L., Laureshyn, A., Svensson, A. & Hydén, C., 2010. Cyclists in Roundabouts-- Different Design Solutions. *Accident Analysis and Prevention*, 42(4), pp.1338-1351.
- Sarosh, K. & Winai, R., 2001. Characteristics of Passing and meeting Maneuvers on Exclusive Bicycle Paths. *Transportation Research Record*, 1776(1), pp. 220-228.

Saunier, N. & Sayed, T., 2006. *A Feature-based Tracking Algorithm for Vehicles in Intersections*. The 3rd Canadian Conference on Computer and Robot Vision.

Sayed, T., Brown, G. & Navin, F., 1994. Simulation of Traffic Conflicts at Unsignalized Intersections with TSC-Sim. *Accident Analysis and Prevention*, 26(5), pp. 593-607.

Sayed, T., Zaki, M. H. & Autey, J., 2013. Automated Safety Diagnosis of Vehicle-Bicycle Interactions Using Computer Vision Analysis. *Safety Science*, Volume 59, pp. 163-172.

Singh, A., 2012. *A Study of Cyclist Behaviour and Perceptions on Shared Space*. London, UK.

Sobhani, A., Young, W., Bahrololoom, S. & Sarvi, M., 2012. *Calculating Time-To-Collision for Analysing Right Turning Behaviour at Signalised Intersections*. Perth, UK.

Songchitruksa, P. & Tarko, A. P., 2006. The Extreme Value Theory Approach to Safety Estimation. *Accident Analysis and Prevention*, 38(4), pp. 811-822.

- Songchitruksa, P. & Tarko, A., 2007. Practical Method for Estimating Frequency of Right-Angle Collisions at Traffic Signals. *Transportation Research Record*, Volume 1953, pp. 89-97.
- Spek, A.C., Wieringa, P. & Janssen, W.H., 2006. Intersection Approach Speed and Accident Probability. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9(2), pp. 155-171.
- Summala, H., Pasanen, E., Räsänen, M. & Sievänen, J., 1996. Bicycle Accidents and Drivers' Visual Search at Left and Right Turns. *Accident Analysis and Prevention*, 28(2), pp. 147-153.
- Tarko, A., Davis, G., Saunier, N., Sayed, T. & Washington, S., 2009. Subcommittee on Surrogate Measures of Safety and Committee on Safety Data Evaluation and Analysis. *Transportation Research Record*, Washington, DC.
- Thompson, D. C., Rebolledo, V., Thompson, R. S., Kaufman, A. & Rivara, F. P., 1997. Bike Speed Measurements in a Recreational Population: Validity of Self Reported Speed. *Injury Prevention*, Volume 3, pp. 43-45.
- Thompson, H. T. & Perkins, D. D., 1983. Surrogate Measures for Accident Experience at Rural Isolated Horizontal Curves. *Transportation Research Record*, Volume 905, pp. 142-147.
- Townshend, A., Worringham, C. & Stewart, I., 2008. *Assessment of Speed and Position During Human Locomotion Using Nondifferential GPS*, Med Sci Sports Exerc.
- Transport Canada, 2010. *A Quick Look at Fatally Injured Vulnerable Road Users*. Available at: <http://www.tc.gc.ca>. [Online, Accessed 10 November 2013].

- Van der Horst, R., De Goede, M., De Hair-Buijssen, S. & Methorst, R., 2014. Traffic Conflicts on Bicycle Paths: A Systematic Observation of Behaviour From Video. *Accident Analysis and Prevention*, Volume 62, pp. 358-368.
- Viera, A. & Garrett, J., 2005. Understanding Interobserver Agreement: The kappa Statistic. *Family Medicine*, 37(5), pp. 360-363.
- Virkler, M. R. & Balasubramanian, R., 1998. Flow Characteristics on Shared Hiking/Biking/ Joggin Trails. *Transportation Research Record*, Volume 1636, pp. 43-46.
- Vogel, K., 2003. A Comparison of Headway and Time To Collision as Safety Indicators. *Accident Analysis and Prevention*, 35(3), pp. 427-433.
- Wachtel, A., Forester, J. & Pelz, D., 1995. Signal Clearance Timing for Bicyclists. *ITE Journal*, 65(3), pp. 38-45.
- Wang, C. & Stamatiadis, N., 2010. Bicyclist Injury Severity in Bicycle-Motor Vehicle Crashes at Unsignalized Intersections in Kentucky. *Transportation Research Record*, pp. 1-13.
- Wang, D., Feng, T. & Liang, C., 2008. Research on Bicycle Conversion Factors. *Transportation Research Part A: Policy and Practice*, 24(8), pp. 1129-1139.
- Wang, Y. & Nihan, N., 2004. Estimating the Risk of Collision Between Bicycles and Motor Vehicles at Signalized Intersections. *Accident Analysis and Prevention*, 36(3), pp. 313-321.

- Washington, S. P., Karlaftis, M. G. & Mannering, F. L., 2011. *Statistical and Econometric Methods for Transportation Data Analysis*. Second ed. U.S.: CRC Press, Taylor & Francis Group.
- Wayne, P., 1997. Bicyclist Performance on a Multiuse Trail. *Transportation Research Record*, Volume 1578, pp. 127-131.
- Weigand, L., 2008. *A Review of Literature: Intersection Treatments to Improve Bicycle Access and Safety*, Portland, Oregon.
- Wheeler, N., Conrad, R. & Figliozzi, M., 2010. *A Statistical Analysis of Bicycle Rider Performance: The Impact of Gender on Riders' Performance at Signalized Intersections*. Washington D C, TRB 89th Annual Meeting Compendium of Papers DVD.
- Witte, T. & Wilson, A., 2004. Accuracy of Non-Differential GPS for the Determination of Speed Over Ground. *Journal of Biomechanics*, 37(12), pp. 1891-1898.
- Wood, J. M., Lacherez, P. F., Marszalek, R. P. & King, M. J., 2009. Drivers' and Cyclists' Experiences of Sharing the Road: Incidents, Attitudes and Perceptions of Visibility. *Accident Analysis and Prevention*, 41(4), pp. 772-776.
- Zaki, M. H., Sayed, T. & Cheung, A., 2013. *Automated Collection of Cyclist Data Using Computer Vision Techniques*. Washington DC., Transportation Research Board Annual Meeting.
- Zheng, L., Ismail, K. & Meng, X., 2014. Freeway Safety Estimation Using Extreme Value Theory Approaches: A Comparative Study. *Accident Analysis and Prevention*, Volume 62, pp. 32-41.

Appendix A : Sample of Tracker Inputs and Outputs for Speed Measurement.

Video_path	Video_name	Format	Event_starts	Duration
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:00:08	0:00:10
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:01:18	0:00:10
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:04:26	0:00:13
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:04:55	0:00:13
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:10:37	0:00:12
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:11:23	0:00:13
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:11:57	0:00:08
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:12:41	0:00:05
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:13:01	0:00:12
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:13:20	0:00:15
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:14:56	0:00:08
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:16:06	0:00:14
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:16:43	0:00:12
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:17:46	0:00:09
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:21:43	0:00:16
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:24:59	0:00:10
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:25:24	0:00:08
F:\AVI-Videos\Meckenzi-King-Bridge-Intersection\JUN.2011\JUN.8.2011\	JUN.0807	avi	0:26:51	0:00:10

Figure A-1: Sample of video recording information methodology.

```

#config 2.0
#####
##### FEATURE GROUPER CONFIGURATIONS#####
#####
#tracks are features that have been detected in the video
#compress_visualization will shrink the output video by half
homography_point_correspondence=0
debug_gui=0
#visualize_with_coordinates=0
log_tracks_info=0
compress_visualization=1
#number of frames an object must exist before it is activated
min_frames_tracked=5
#minimum distance a feature has to move in min_frames_tracked frames before it is considered
to be activated
min_distance_moved_required=.02
#distance to look around an activated feature before it is added to a feature group
maximum_distance_activated=1.3
#how far two tracks can be from eachother before they are segmented
segmentation_threshold=0.7
#the minimum distance between two feature tracks
min_distance_between_tracks=.3
#chooses whether to use the bottom of an object rather than the middle to represent position
use_bottom_position=1
##### Visualizer Options#####
#draw the trailing object line
draw_track_line=1
#draw the object features and connecting features
draw_graph=1
#draw the ids of the objects
draw_component_id=1
#draw the current frame in the top left corner
draw_frame=1
#draw the bounding object ellipse
draw_ellipse=0
##### KLT TRACKER OPTIONS#####
tracker_max_corners=900
tracker_quality_level=0.07
tracker_min_distance=1
tracker_block_size=3
tracker_use_harris_detector=0
tracker_k=15

```

Figure A-2: Sample of the tracker system configuration file.



Figure A-3: Video frame samples: before and after tracker system application.

Appendix B : Tracker System Improvement.

Many improvements have been made to the existing open-source feature-based vehicle tracking system (<http://code.google.com/p/opencv-feature-tracker>), in terms of adding new functionalities and mechanisms, and improving already existing ones. Next is a summary of those improvements and the motivation behind them. They are presented along with snapshots and diagrams demonstrating the changes made.

1- Object position

Each object, *e.g.*, vehicle, cyclist, and pedestrian, detected by the tracker system is represented by a set of features, *i.e.*, vertices. Each feature, *t*, is represented in terms of its *x* and *y* coordinates along the *X* and *Y* axis of the Cartesian plane.

Since the existing open-source version of the tracker system had the only the center *C* node identified, new object nodes have been added. These nodes are: Top-Right, Right, Bottom-Right, Bottom, Bottom-Left, Left, Top-Left, and Top (as shorthand will be denoted *TR*, *R*, *BR*, *B*, *BL*, *L*, *TL*, and *T* respectively). However, the center *C* node has been used to find and compute the other nodes as shown in Table 4-1. These nodes can help to obtain a detailed profile of the object (vehicle, cyclist, or others) in order to pick the desired or required location of the object, for example; front, rear, and side. The developed tracker system was extended to identify all these nodes.

In this tracking system, each object of interest, a motor vehicle or cyclist, was represented by a set of features. A bounding box, for each object, is formed as the smallest upright rectangle containing all object features as shown in Figure B-1. This box is obtained using functions from a function in the Open CV project (OpenCV, 2012).

(a)



(b)

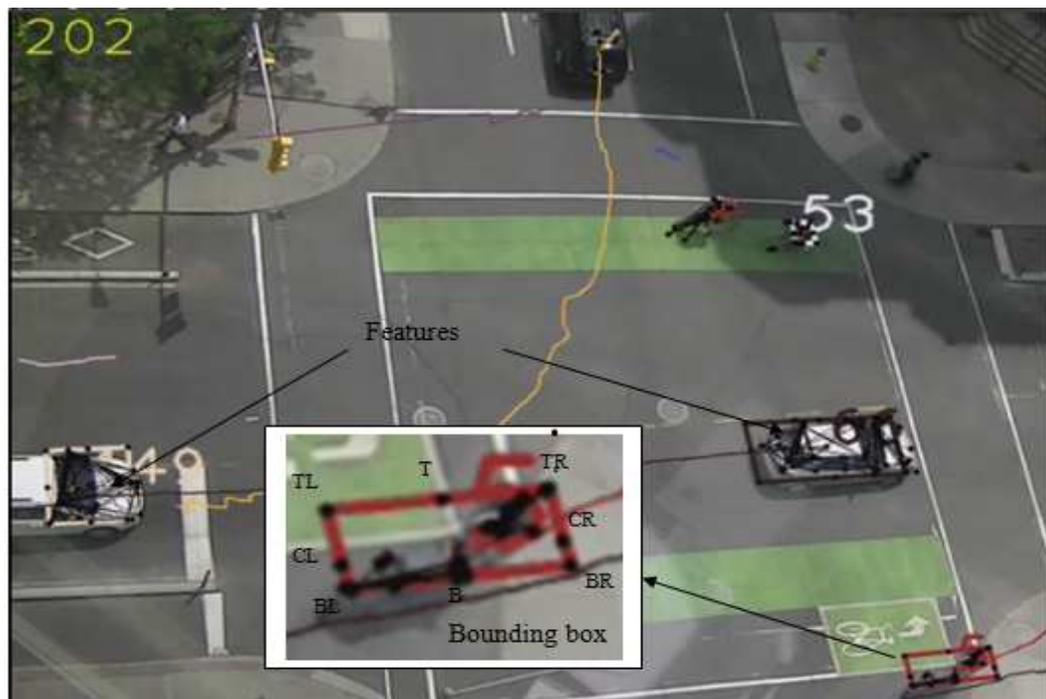


Figure B-1: Object bounding box. (a) before adding the bounding box with center tracks only. (b) after adding the bounding box.

2- Speed parameter

Each object detected by the tracker system is represented by a graph. Graph is formed by the features of the object and the edges, *i.e.*, connections, between those features as shown in Figure B-2. When a set of features is detected, edges are formed between those features based on the distance parameter, `maximum_distance_activated`. If the distance between two features, t_1 and t_2 , is less than or equal to the distance parameter then a connection, or edge, is formed between t_1 and t_2 . The distance between t_1 and t_2 is computed and maintained by the tracker system as $d(t_1, t_2)$.

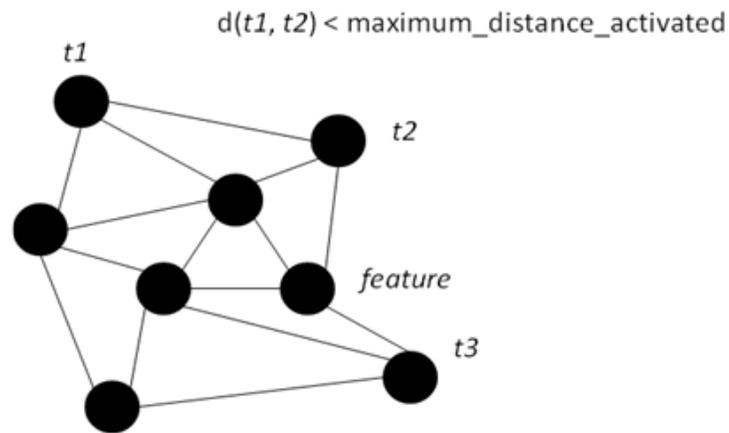


Figure B-2: Features of an object and the connections, edges, between them.

As shown in Figure B-2 there is a connection, edge, between feature t_1 and feature t_2 because the distance between them is less than or equal to the `maximum_distance_activated` threshold, while no connection is formed between features t_2 and t_3 because the distance between them is greater than the `maximum_distance_activated` threshold. The threshold is the limit allowed for those two features to be away from each other before the edge that connects them is severed.

After the initial graph for an object is formed, the tracker system keeps updating the distance of an edge as the two features move comparing it against the distance threshold to determine whether an edge should be kept or severed. The `maximum_distance_activated` parameter is the only parameter taken under consideration when forming the initial graph for an object. In some cases, the distance threshold was not enough or sufficient in dealing with the issue of object over-grouping. Over-grouping is defined as the two or more different objects are grouped or merged, incorrectly, by the tracker system and considered as one object as shown in Figure B-3. In particular, Over-grouping can occur between two features. For example, features t_1 and t_2 , as shown in Figure B-2, belong two different objects because the distance between them is within the distance threshold, and therefore the two objects are merged, incorrectly, into one object.



Figure B-3: The rear of a vehicle and a pedestrian being over-grouped as one object.

A configuration speed parameter, or speed threshold, was added to overcome the issue of over-grouping. The speed threshold is define as, the difference in speed allowed between the two features of an edge before the edge is severed. The speed is calculated

and stored for the feature at each frame, i , for which the feature exists. For example, the difference in speed between features t_1 and t_2 can be calculated as:

$$\text{Difference-in-Speed } (t_1, t_2) = |V_{t1} - V_{t2}|$$

where, V_{t1} is the speed for the feature t_1 (m/s),

V_{t2} is the speed for the feature t_2 (m/s).

However, in some cases, the speed threshold was found after many trails not enough or sufficient in dealing with the issue of object over-grouping. A new configuration parameter, `direction_angle` threshold, was added to the tracker system in order to improve its ability to deal with the issue of over-grouping. The `direction_angle` threshold is defined as the difference in direction allowed between the two features of an edge before the edges is severed. The direction is expressed by an angle (Θ) measured in degree.

The difference in direction between two features, t_1 and t_2 can be calculated as

$$\Theta = \cos^{-1} \frac{t_1 \cdot t_2}{|t_1| \cdot |t_2|}$$

$t_1 \cdot t_2$ is defined as the dot product between t_1 and t_2 , where,

$$t_1 \cdot t_2 = (Dx_{t1} * Dx_{t2}) + (Dy_{t1} * Dy_{t2})$$

where Dx and Dy , are the rates of change for the x and y coordinates, respectively.

The testing and analysis conducted proved that by using those two new thresholds, speed threshold and `direction_angle` threshold, in most cases the developed tracker system was successfully able to deal with over-grouping as shown in Figure B-4.



Figure B-4: No over-grouping between objects.

After describing the speed threshold and direction_angle threshold, the algorithm followed by the tracker system to form or keep edges is given below.

1. For each pair of edges, t_1 and t_2 as shown in Figure B-2, detected by the tracker, check if the distance between them, $d(t_1, t_2)$, is less than or equal to the maximum_distance_activated threshold. If yes, then an edge between (t_1, t_2) is formed, otherwise no edge is formed.
2. After all edges are formed, the object is created, and an id number is assigned for this object.
3. At each frame in the object's lifetime, for each edge between t_1 and t_2 , belonging to the object:
 - If the distance between t_1 and t_2 is greater than the maximum_distance_activated threshold, then the edge between t_1 and t_2 is severed.

- If the difference in speed between the two features t_1 and t_2 , is greater than the `speed_threshold`, then the edge is severed.
- If the direction difference between the two features t_1 and t_2 , is greater than the `direction_angle` threshold, then the edge is severed.

3- Mask operation

In many cases the video being processed may contain objects that are considered irrelevant or unimportant to the conflict analysis such as pedestrians, or other irrelevant.

The mask is an operation applied by the tracker to mask out particular regions or areas in the video the user wishes to ignore. Ignoring those objects can substantially speed up the tracking process and obtain tracks in shorter time. In general, the more the mask covers, the less objects the tracker has to deal with and so the process of detecting and tracking features becomes significantly faster. Figure B-5 shows the two conditions, before and after adding the mask to the video. The mask can easily be created by the user manually and passed to the tracker as an image file (*e.g.*, bmp, png, and jpeg).

The testing and analysis conducted indicate that using the mask can speed up the tracking process of a video in terms of running time. For an average video with duration of 14 seconds, without using a mask, the tracker takes 413 seconds to process the video. While when the mask is applied, the processing time decreases to 168 seconds, which is almost three times less than the time taken when no mask is applied.

(a)



(b)



Figure B-5: Video mask. (a) before adding the mask. (b) after adding the mask.

When dealing with low-quality videos, the mask could also be used to mask out distracting objects, for example shades and other disruptive textures that are affecting the tracks quality of the objects. In such cases, one might attempt to apply a restrictive version of the mask, *i.e.*, a restrictive mask, to concentrate the tracker resources and

efforts to a particular area or region in the video by masking out other objects. Figure B-6 shows before and after using the restrictive version of the mask.

(a)



(b)

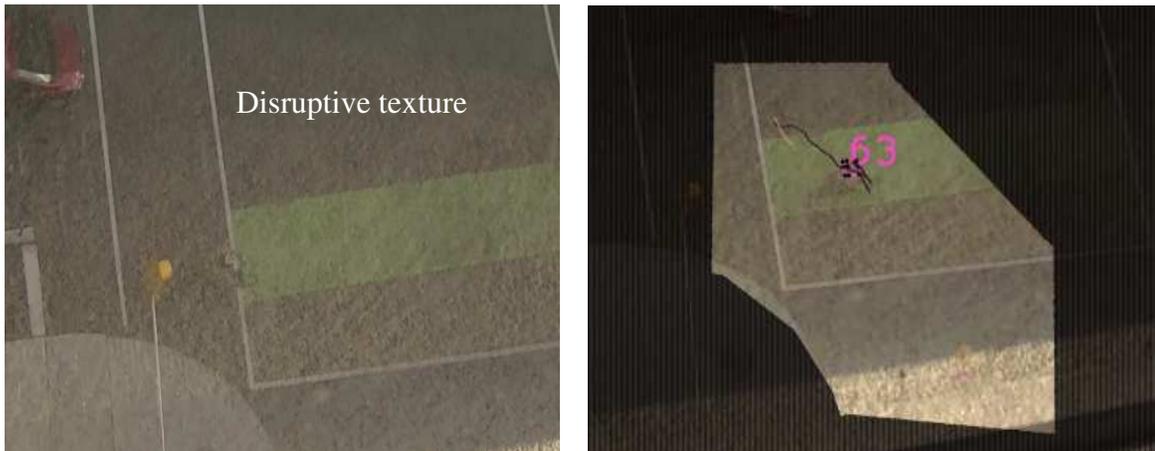


Figure B-6: Video restrictive mask. (a) before adding the restrictive mask. (b) after adding the restrictive mask.

4- Minimum features

The `minimum_features` parameter is already implemented in the open-source version of tracker system. The value of this parameter was set constant at the source code level. The `minimum_features` parameter represents the number of features that are required to be gathered from the object, in order for the object to be tracked by the tracker system. This parameter is closely related to the issue of partial-tracks and it is affected by the quality of the video. Partial-tracks occur when the number of features goes below the threshold assigned in the configuration file for the object to be tracked and detected, and this causes a discontinuity in the tracks of the object where the tracks at those lost frames are unknown or uncalculated. Having the `minimum_feature` parameter editable, the parameter can be assigned the appropriate value, based on the quality of the video, in order to overcome partial-tracks. A standard value of five was used to the `minimum_feature` parameter, however in some cases a lower value was used to deal with the low quality videos that contain shades and disruptive textures. The issue of partial-tracks is discussed in detail in the subsequent section.

5- Configuration file

The configuration file is a text file that the user uses to configure or assign values for specific parameters or conditions that the tracker system will depend on. These parameters were found to depend on: field condition, illumination, weather conditions, video quality, traffic density, and whether the tracked cyclist stands out from the background colour. When the tracker system begins executing, it consults the values and conditions in the configuration file in order to see what parameters should be in effect.

Figure B-7 shows the configuration with existing parameters from the open-source version of the tracker system.

```
##### FEATURE GROUPER CONFIGURATIONS#####  
#tracks are features that have been detected in the video  
#compress_visualization will shrink the output video by half  
homography_point_correspondence=0  
debug_gui=0  
#visualize_with_coordinates=0  
log_tracks_info=0  
compress_visualization=1  
#number of frames an object must exist before it is activated  
min_frames_tracked=5  
#minimum distance a feature has to move in min_frames_tracked frames before it is considered to be  
activated  
min_distance_moved_required=.01  
#distance to look around an activated feature before it is added to a feature group  
maximum_distance_activated=2  
#how far two tracks can be from each other before they are segmented  
segmentation_threshold=1.0  
#the minimum distance between two feature tracks  
min_distance_between_tracks=.5  
##### KLT TRACKER OPTIONS#####  
tracker_max_corners=1300  
tracker_quality_level=0.07  
tracker_min_distance=1  
tracker_block_size=3  
tracker_use_harris_detector=0  
tracker_k=15
```

Figure B-7: Configuration file before the improvement.

New parameters were added to the configuration file in order to give the user as much control over certain values and conditions and as possible. By having those parameters set from the configuration file, the user can choose or pick the values that suit or fit the video being processed the best. Figure B-8 shows the new configuration file after new parameters were added. The highlighted lines indicate the new parameters that have been added.

```

##### FEATURE GROUPER CONFIGURATIONS#####
#tracks are features that have been detected in the video
#compress_visualization will shrink the output video by half
homography_point_correspondence=0
debug_gui=0
#visualize_with_coordinates=0
log_tracks_info=0
compress_visualization=1
#number of frames an object must exist before it is activated
min_frames_tracked=5
#minimum distance a feature has to move in min_frames_tracked frames before it is considered to be activated
min_distance_moved_required=.01
#distance to look around an activated feature before it is added to a feature group
maximum_distance_activated=1.50
#how far two tracks can be from each other before they are segmented
segmentation_threshold=1.0
#how much speed difference do we allow between vertices before their edge is severed
speed_threshold= 0.50
#the maximum angle (theta) in degrees allowed between two vertices before their edge is severed
direction_angle_threshold= 90.0
#the minimum distance between two feature tracks
min_distance_between_tracks=.2
#chooses whether to use some type of mask (-1: no mask, 0: static mask)
use_mask=0
##### Visualizer Options#####
#draw the trailing object line
draw_track_line=1
#chooses the position to draw
use_bottom_position=0
use_center_position=1
use_top_position=0
use_topRight_position=0
use_topLeft_position=0
use_bottomLeft_position=0
use_bottomRight_position=0
#draw the object features and connecting features
draw_graph=0
#draw the ids of the objects
draw_component_id=1
#draw the current frame in the top left corner
draw_frame=1
#draw the bounding object ellipse
draw_ellipse=0
#draw the bounding box that contains each object(plus the positions' locations)
draw_objects_bounding_box=1
#set the font size
label_font_size=5.0
##### Low-Level Configuration Options #####
#the minimum number of features (i.e. vertices) required for an object to be tracked
minimum_features=5
#choose whether to merge two components as one component if two features were close enough to one another
merge_components=1
##### KLT TRACKER OPTIONS#####
tracker_max_corners=2750
tracker_quality_level=0.03
tracker_min_distance=1
tracker_block_size=3
tracker_use_harris_detector=0
tracker_k=15

```

Figure B-8: Configuration file after the improvement.

6- Partial tracks

As mentioned in the previous section the partial-tracks occur when the number of features goes below the threshold (*i.e.*, `minimum_feature` parameter). Figure B-9 shows the sample of tracked and untracked object in one scene.

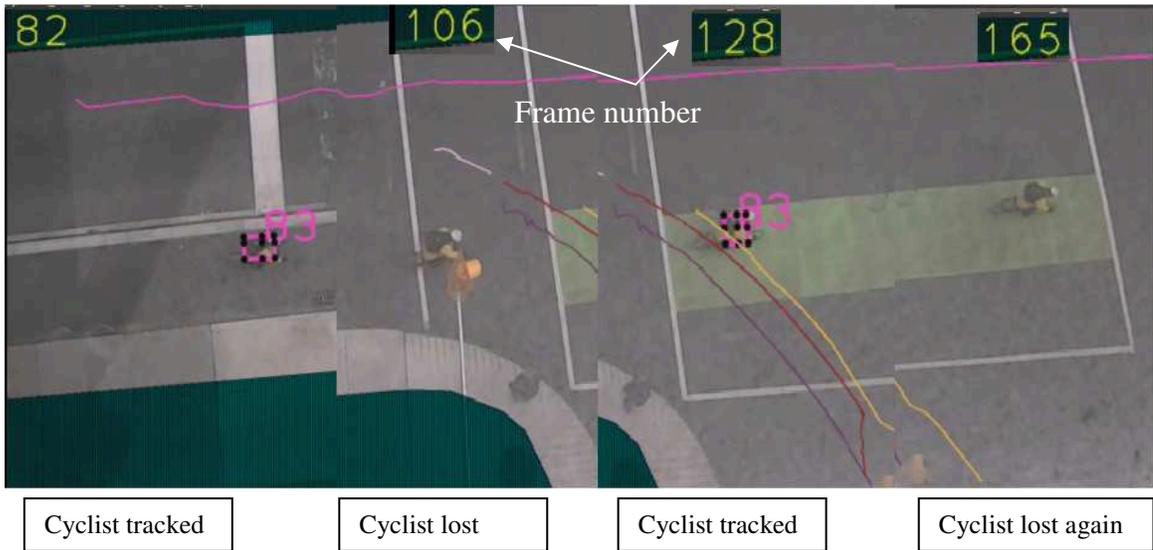


Figure B-9: Explanation of tracked and untracked object.

Improvement:

In order to deal with the issue of partial-tracks, a simple interpolation algorithm has been implemented to construct the lost data of the object based on its other available data (other existing frames). Object's data, at certain frames, are considered lost when the object is not tracked over those frames. This helps to have better tracks especially when dealing with low quality videos. The required mechanism was added to the tracker in order to detect partial-tracks. When partial-tracks occur, the user can now be aware of this issue when analyzing the text output associated with the video. When writing the profile of each object, the tracker indicates the number of frames that were lost. Also, the question mark symbol, `?`, can now be seen in the text output to indicate the loss of data

$$Z = \frac{X_a + X_b}{2}$$

where;

Z : the average data value between the last and first data tracked of the object

X_a : the last data value tracked of the object

X_b : the first data value tracked again of the object

When data are lost over two frames or more, the interpolation algorithm always places the calculated lost data, Z , in the middle of the set, as the average value, and continues to follow this approach for each next calculated lost data in order to obtain a gradual and averaged range among the values of the set. For example, when the first lost value in the set is computed as the average of the values in frame 8 and frame 17, the value is inserted into frame 12, since 12 is the average value between frames 8 and 17. Table B-1 shows a sample calculation demonstrating the process of the interpolation algorithm.

7- Stitching Object

In some cases, the same object might be tracked inconsistently with different id numbers. For example, as shown in Figure B-11, object 120, appearing first at frame 69 and continued to frame 140, is actually the same as object 194, which appears first at frame 180 and continued to frame 245. But due to partial tracks, object 120 was lost and then was tracked again under a new id number 194. The user can identify this case by manual observation of the video output.

Table B-1: Sample calculation of the interpolation algorithm.

Frame	Frame	Frame	Frame	Frame	Frame	Frame	Frame	Frame	Frame
8	9	10	11	12	13	14	15	16	17
41.1187	?	?	?	41.0021	?	?	?	?	40.8860
41.1187	?	41.0604	?	41.0021	?	40.9440	?	?	40.8860
41.1187	41.0895	41.0604	41.0312	41.0021	40.9730	40.9440	40.9150	?	40.8860
41.1187	41.0895	41.0604	41.0312	41.0021	40.9730	40.9440	40.9150	40.9005	40.8860
41.1187	41.0895	41.0604	41.0312	41.0021	40.9730	40.9440	40.9150	40.9005	40.8860
	Adjusted value data								

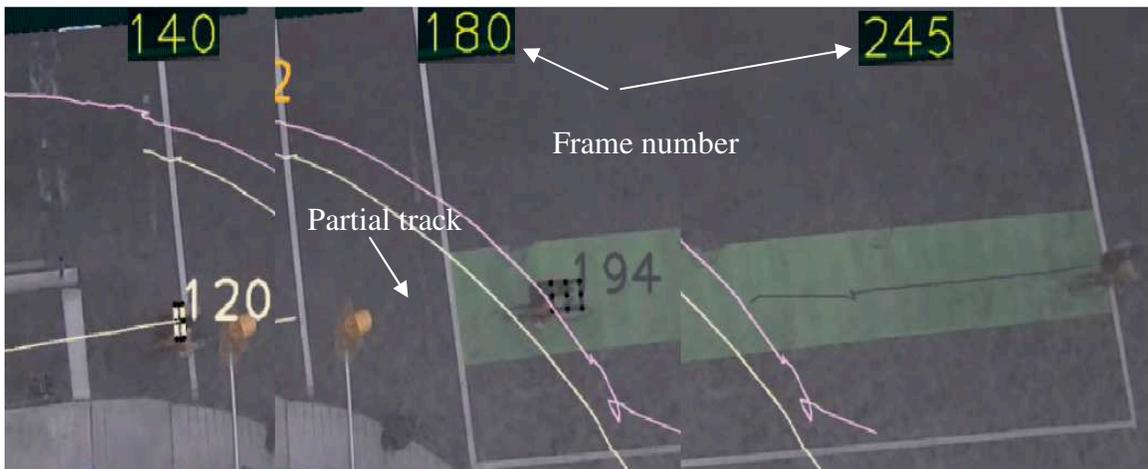


Figure B-11: Visualization video sample with two different ids (120, 194) for one object.

Improvement:

A simple algorithm to stitch those different objects, representing the same entity, has been developed. The different id numbers for those objects are given, by the user, to the algorithm as input. The algorithm merges the tracks of those objects and then applies the simple average linear interpolation algorithm as described in partial-tracks section.

To construct a new profile of this object, both object 120 and 194, must be stitched or combined together, and Linear-Interpolation must be implemented to compute the lost tracks between the two objects. The algorithm to stitch and combine the two objects can be summarized in the following outline:

1. Given object A and object B, where A's first and last frames are A-first and A-last, respectively; and B's first and last frames are B-first and B-last, respectively.
2. If A has partial tracks, then fix the partial tracks by applying Linear-Interpolation on the lost tracks.
3. If B has partial tracks, then fix the partial tracks by applying Linear-Interpolation on the lost tracks.
4. Construct a new object with id C, let C-first frame be A-first and C-last frame be B-last.

Appendix C : Layout of the intersections.

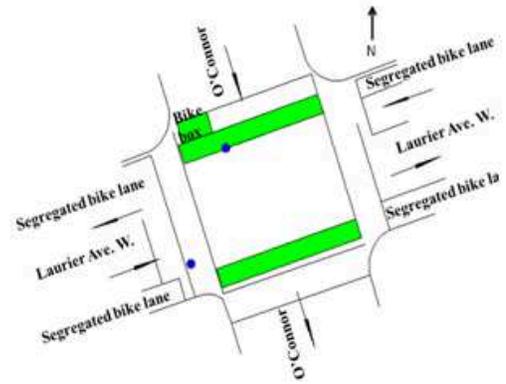
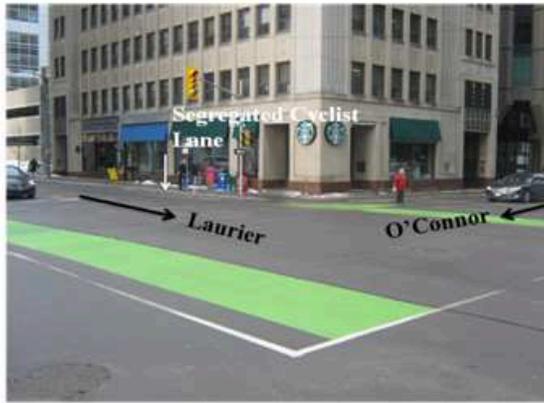


Figure C-1: Layout of Laurier- O'Connor signalized intersection



Figure C-2: Layout of Laurier- Bank signalized intersection.

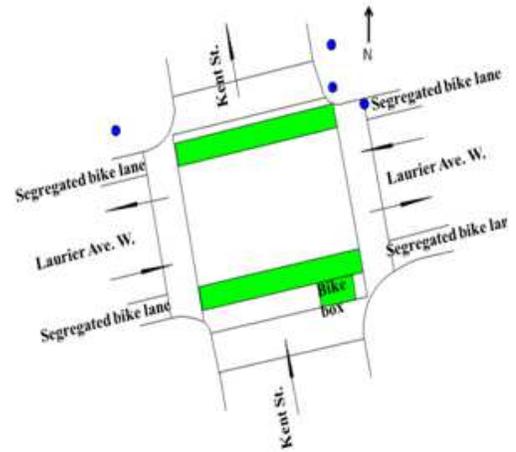
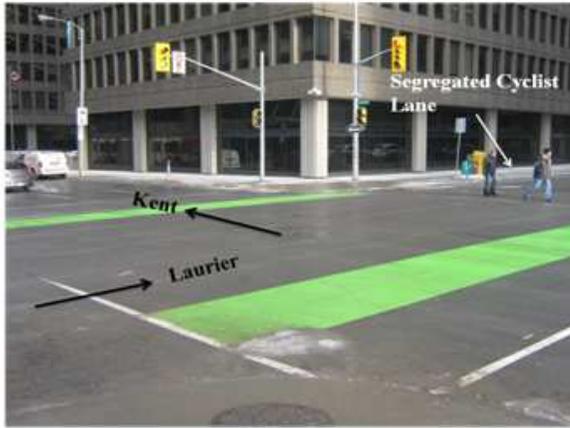


Figure C-3: Layout of Laurier- Kent signalized intersection.

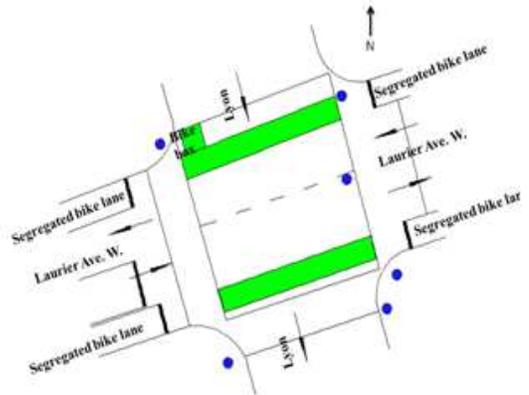


Figure C-4: Layout of Laurier- Lyon signalized intersection.



Figure C-5: Layout of Laurier- Bay signalized intersection.

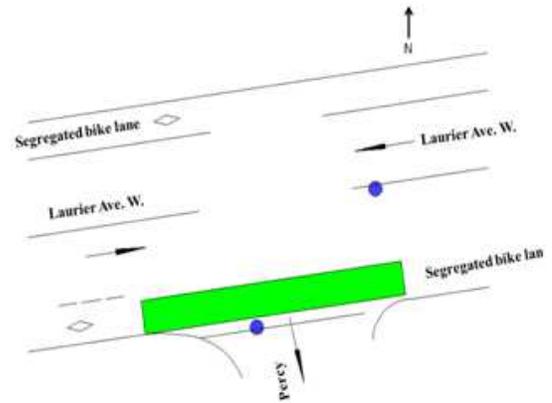


Figure C-6: Layout of Laurier- Percy intersection.

Appendix D : Graphical Diagnostic of Probability and Quantile Plots.

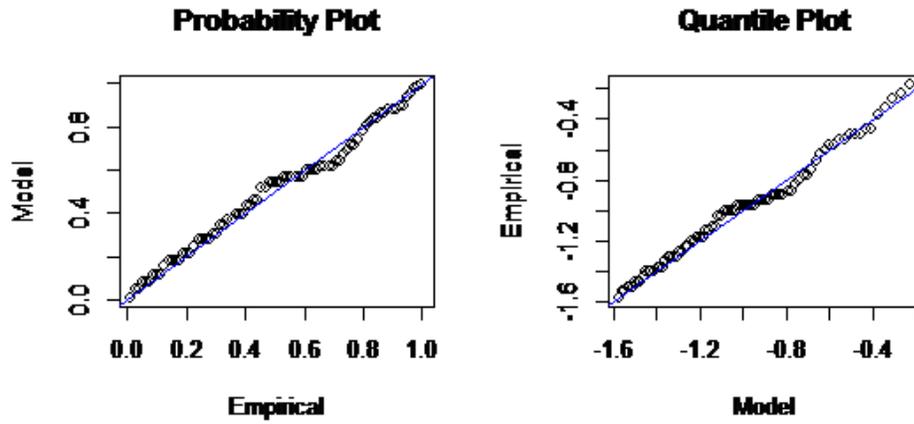


Figure D-1: Probability and quantile plots of Laurier-O'Connor signalized intersection.

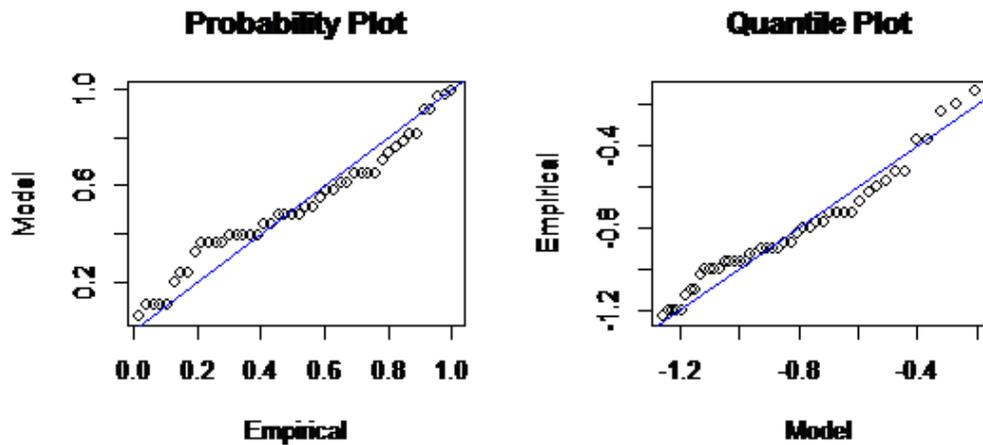


Figure D-2: Probability and Quantile plots of Laurier-Bank (eastbound) signalized intersection.

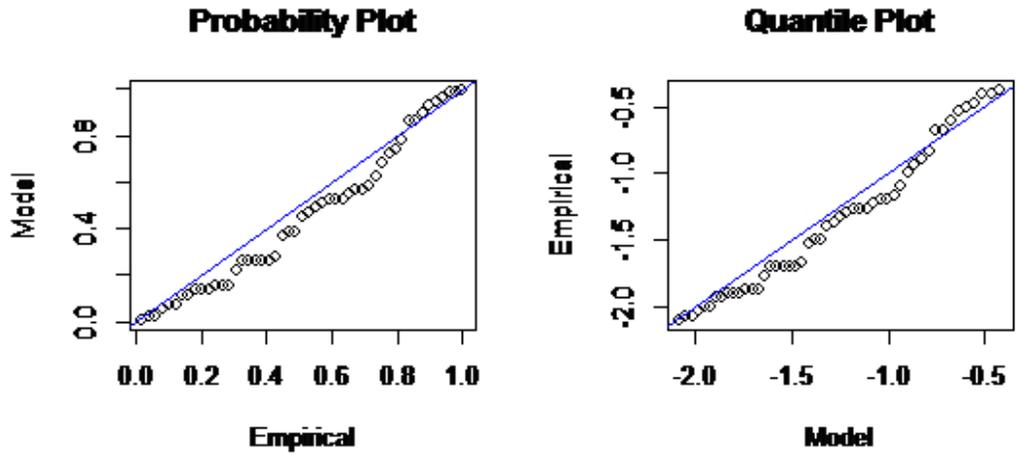


Figure D-3: Probability and Quantile plots of Laurier-Bank (westbound) signalized intersection.

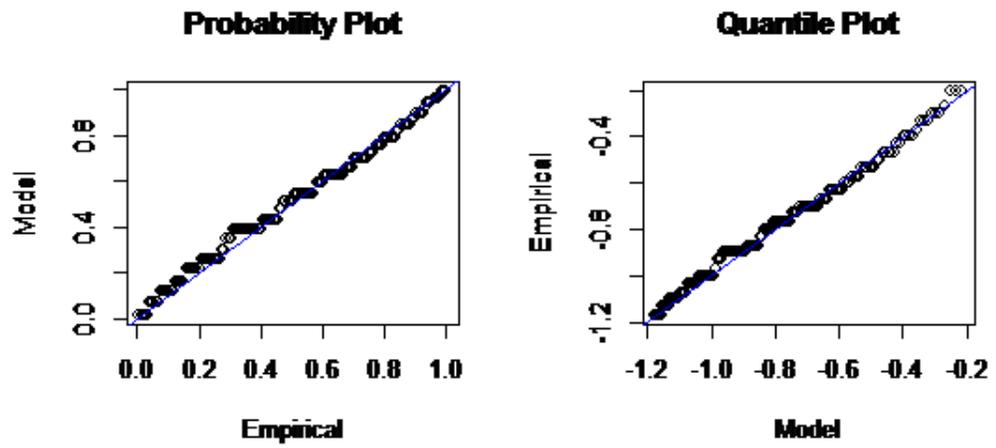


Figure D-4: Probability and Quantile plots of Laurier-Kent signalized intersection.

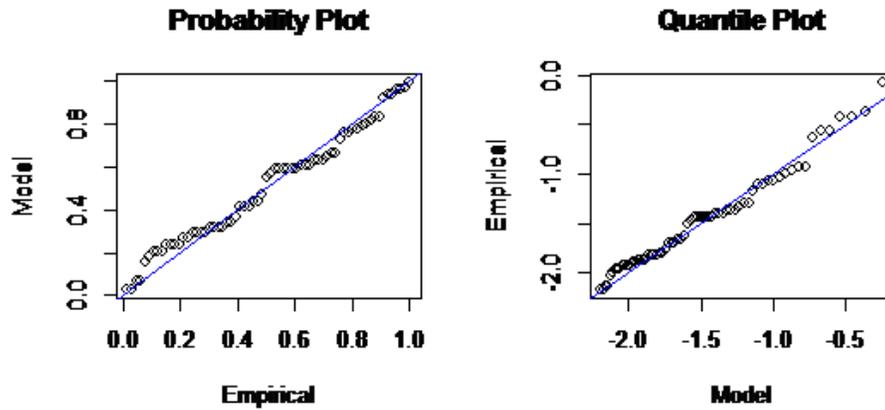


Figure D-5: Probability and Quantile plots of Laurier-Lyon signalized intersection.

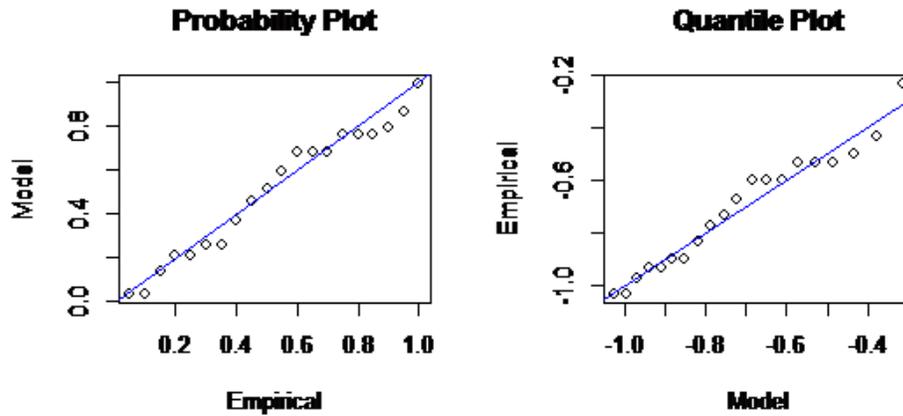


Figure D-6: Probability and Quantile plots of Laurier-Bay signalized intersection.

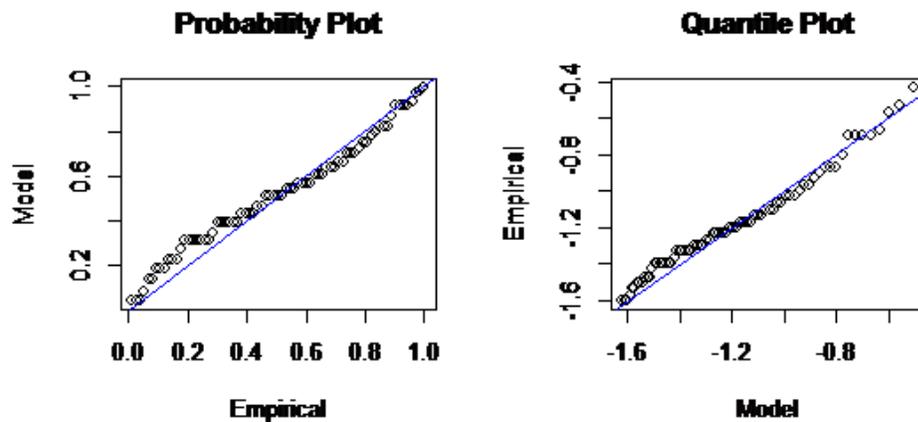


Figure D-7: Probability and Quantile plots of Laurier-Somerset signalized intersection.

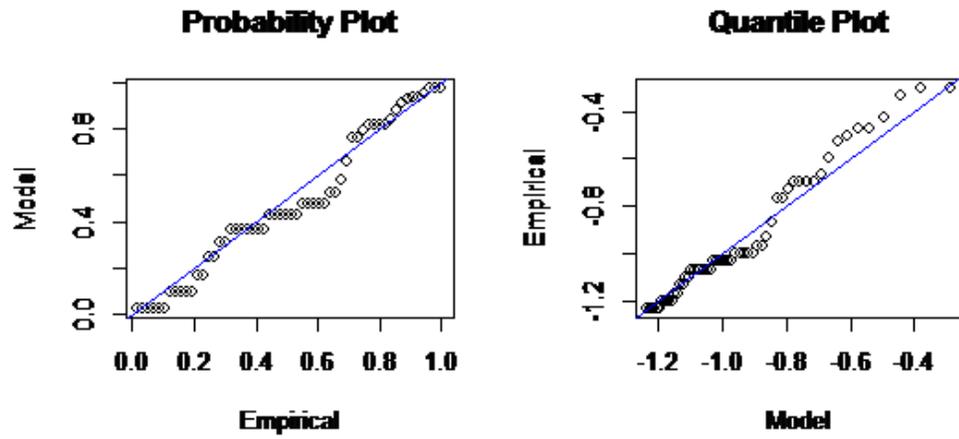


Figure D-8: Probability and Quantile plots of Laurier-Sussex signalized intersection.

Appendix E : Published Papers and Conference Presentation.

Information that was written in this thesis for Chapters 1, 2, 3, 4 and 5 have appeared in the following publications.

- 1- **Kassim, A.**, K. Ismail, and Y. Hassan, "Automated Measuring of Cyclist-Motor Vehicle Post Encroachment Time at Signalized Intersections", *Canadian Journal of Civil Engineering*, 41(7), pp. 605-614. 2014.
- 2- **Kassim, A.**, K. Ismail, and Y. Hassan, "Automated Measuring of Cyclist/Motor Vehicle Post Encroachment Time at signalized Intersections", in TRB 93 rd Annual Meeting, Washington DC 2014.
- 3- **Kassim, A.**, K. Ismail, and Y. Hassan, "A Methodology for Measuring the Severity of Cyclist/Motor Vehicle Conflicts at Signalized Intersections", CSCE Annual Conference, 2013.
- 4- **Kassim, A.**, L. Pascoe, K. Ismail, and A.O., Abd El Halim, "A Vision-Based Analysis of Cyclist Crossing Speed", in TRB 91st Annual Meeting Compendium of papers, Washington DC, 2011.