AUTOMATED ANALYSIS OF PEDESTRIAN-VEHICLE CONFLICTS USING VIDEO DATA

Karim Ismail, M.A.Sc.
Research Assistant, Department of Civil Engineering
University of British Columbia
karim@civil.ubc.ca

Tarek Sayed, PhD, P.Eng.
Distinguished University Scholar
Professor of Civil Engineering
University of British Columbia
tsayed@civil.ubc.ca

Nicolas Saunier, PhD
Research Associate, Department of Civil Engineering
University of British Columbia
saunier@civil.ubc.ca

Clark Lim, M.A.Sc., P.Eng.
Acuere Consulting
clim@acuere.ca

Word Count = 7000 words
[5000 words + 1 table, 6 figures + 1 appendix; *references 1200*]

AUTOMATED ANALYSIS OF PEDESTRIAN-VEHICLE CONFLICTS USING VIDEO DATA

ABSTRACT

Pedestrians are vulnerable road users, and despite their limited representation in traffic events, pedestrian-involved injuries and fatalities are overrepresented in traffic collisions. However, little is known about pedestrian exposure to the risk of collision, especially when compared to the amount of knowledge available for motorized traffic. More data and analysis is therefore required to understand the processes that involve pedestrians in collisions. Collision statistics alone are inadequate for the study of pedestrian-vehicle collisions because of data quantity and quality issues. Surrogate safety measures, as provided by the collection and study of traffic conflicts, were developed as a proactive complementary approach to offer more in-depth safety analysis. However, high costs and reliability issues have inhibited the extensive application of traffic conflict analysis. This paper presents an automated video analysis system that can: 1) detect and track road users in a traffic scene, and classify them as pedestrian and motorized road users; 2) identify important events that may lead to collisions; 3) calculate several severity conflict indicators. The system seeks to classify important events and conflicts automatically, but can also be used to summarize large amounts of data that can be further reviewed by safety experts. The functionality of the system is demonstrated on a video dataset collected over two days at an intersection in Downtown Vancouver, British Columbia. Four conflict indicators are automatically computed for all pedestrian-vehicle events and provide detailed insight in the conflict process. Simple detection rules on the indicators are tested to classify traffic events. This study is unique in its attempt to extract conflict indicators from video sequences in a fully automated way.

INTRODUCTION

There is a growing movement toward emphasizing sustainability into the transportation system by promoting public transit and improving the traffic conditions for non-motorized modes of transport. Walking is a key non-motorized mode of transport that connects different components of a multimodal transport network and interfaces with external activity areas (land use). Building safe and walking-friendly pedestrian facilities is fundamental to encouraging and accommodating walking activities. For example most modern municipalities are required to have in place official community plans (OCP) to manage growth and many, if not most, of them contain policies that promote pedestrian activities. Furthermore, on February 4, 2008, U.S. Secretary of Transportation announced a $68 billion budget for the U.S. Department of Transportation's 2009 fiscal year which highlighted funds for safety programs that focus on problem areas such as pedestrian injuries.

The study of pedestrian safety focuses on the interaction between pedestrians and other motorized and non-motorized traffic, as well as the conformity to traffic control regulations. Traffic safety analysis has traditionally relied on historical collision data. However, there are some shortcomings to this approach:

1. Traffic collisions are rare and highly random events that usually require extended observation times, usually in the order of years, and sophisticated statistical techniques. As well, many extraneous factors can change during the observation period, further complicating the analysis.
2. Collision-based safety analysis is a reactive approach, which means that a significant number of collisions has to be recorded before action is taken.
3. There are well-known concerns with the quantity and quality of collision data (1). Collision data reporting is often incomplete and biased toward highly damaging collisions. Collision auditing is conducted after collision occurrence, at which time the causes, specific location, and behavioral aspects of the event are subject to judgment – if ever reported.

These shortcomings of using collision data for pedestrian safety analysis are even more acute. For example, collisions involving pedestrians are less frequent than other collision types. Pedestrian-involved collisions accounted from 1992 to 2001 for 3.6% of the total number of collisions in British Columbia (2) and 2.4% in Canada (3). In addition, pedestrian traffic volumes are less readily available than motorized traffic volumes due to the difficulties of collecting pedestrian data. The identification of pedestrian exposure to the risk of collision is therefore difficult. Pedestrians, being vulnerable road users, when involved in collisions, have considerably higher chances of being severely injured, with little chance of the collision being classified as property-damage-only. From 1992 to 2001, pedestrians accounted for 14.8% of traffic collision victims (i.e. injured or killed) in British Columbia and 15.2% in Canada.

The use of surrogate safety measures has been advocated as a complementary approach to address these issues and to offer more in depth analysis than relying on accidents statistics alone. One of the most developed methods relies on traffic conflict analysis (4) (5) (6). Traffic Conflict Techniques (TCTs) involve observing and evaluating the frequency and severity of traffic conflicts at an intersection by a team of trained observers. The concept was first proposed by Perkins and Harris in 1967 (7). A traffic conflict takes place when "two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged" (8). Traffic conflicts are more frequent than traffic collisions. The "safety-relatedness" of traffic conflicts (9), i.e. their relationship to collisions, must be established to use traffic conflicts as surrogates to collisions for safety analysis. A common theoretical framework ranks all traffic interactions by their severity in a hierarchy, with collisions at the top and undisturbed passages at the bottom (1).

TCTs were shown to produce estimates of average accident frequency that are comparable to accident-based analysis (10). Traffic conflicts are manually collected by a team of trained observers, either on site or offline through recorded videos. Despite the considerable effort that is put into the development of training methods and the validation of the observers' judgment, such data collection is subject to intra- and inter-observer variability. This can compromise the reliability and repeatability of traffic conflict data collection. In addition, the training and employment of human observers makes traffic conflict studies costly. In a recent study (11), the effort for extracting pedestrian and motorist data from videos was deemed "immense". This type of data is not only difficult to collect, but also its usefulness is subject to the level of accuracy and precision of the collection process.

Due to the issues and limitations of manual data collection, a growing trend of the use of automated data collection systems has caught on in the field of transportation engineering. In particular, automated video analysis has attracted considerable interest, as video sensors are now widely available (traffic cameras are already installed on many roadways) and inexpensive (1).

Previous work on the automated analysis of video data in transportation has mainly focused on vehicular traffic, e.g. (12) (13). This reflects the fact that the automated detection and tracking of pedestrians in video data is still a distinctively difficult problem. Specific problems for pedestrians arise from their complex movement dynamics and groupings, varied appearance, non-rigid nature, and the generally less organized nature of pedestrian traffic as compared to vehicular traffic that are subject to standard "rules of the road" and lane discipline.

This work strives to address some of the previous shortcomings and research recommendations. This paper discusses the development and testing of an automated video-analysis system that seeks to satisfy the following objectives:

1. Detect and track road users in a traffic scene, and classify them into pedestrian and motorized traffic.
2. Identify important events in a video sequence. The definition of an important event in this study is "any event that involves a crossing pedestrian and a conflicting vehicle in which there exists a conceivable chain of events that could lead to a collision between these road users". To be conceivable, a reasonable chain of events leading to a collision should be considered. The actual quantitative interpretation of this general definition is given in the experimental study.
3. Report objective measures of severity indicators for all events.

The system can either work completely automatically, or be used to assist human experts by sifting through large amounts of video data and identifying the important events that deserve further investigation. The system was tested on video data recorded for two days at a location in the Downtown area of Vancouver, British Columbia. The task of calculating traffic conflict indicators for each event that involved a pedestrian-vehicle interaction was performed in a fully automated way. To the authors' knowledge, little similar work (if any) exists in the automated collection and analysis of pedestrian-vehicle conflicts.

PREVIOUS WORK

Pedestrian-vehicle conflicts

Cynecki (9) described a conflict analysis technique for pedestrian crossings, citing fundamental differences between vehicle-vehicle and pedestrian-vehicle conflicts, and indicating desirable characteristics to conduct a conflict study. Two of these characteristics, repeatability and practicability of traffic conflict studies, can greatly benefit from automated video analysis, which offers a cost-efficient and objective means for traffic conflict analysis. In subsequent bodies of work, several studies adopted traffic conflict analysis to study the level of safety of pedestrian crossings, e.g. (14-23). While the majority of past work was based on observer-based traffic conflict analysis, few studies, e.g. (20), developed a relationship between conflict indicators and automatically measured parameters, such as motorist deceleration rate. In a recent study (24), an automated analysis of video data was performed to investigate the interactions between pedestrians and vehicles at roundabout approaches.

Severity conflict indicators

Various conflict indicators have been developed to measure the severity of an interaction by quantifying the spatial and temporal proximity of two or more road users. The main advantage of conflict indicators is their ability to capture the severity of an interaction in an objective and quantitative way. Concerns however remain regarding the lack of a consistent and accurate definition of conflict indicators (25). Conflict indicators developed in the literature are capable of capturing and connoting different proximal, situational, and behavioral aspects of traffic conflicts. Each indicator however possesses drawbacks that limit their ability to measure the severity of recognized traffic events. For a review of conflict indicators and their relative advantages and limitations, the readers are referred to (26).

Pedestrian detection and tracking

To study pedestrian-vehicle conflicts, all road users must be detected, tracked from one video frame to the next, and classified by type, at least as pedestrians and motorized road users. This is a challenging task in busy open outdoor urban environments. In addition to specific problems when tracking pedestrians, common problems are global illumination variations, multiple object tracking, and shadow handling. For a good illustration of the challenges and techniques, the readers are referred to (27), although it is geared towards the study of human motion at a finer scale than this study requires. In (27), the different approaches are classified into:

- Tracking by detection: detection of objects is done using background modeling and subtraction with the current image (24) (28) (29) (30), or deformable templates, i.e. a model of image appearance using color distribution, edge characteristics, and texture. Image classifiers can be trained on labeled data to detect pedestrians (31). In many cases, especially if the objects are well separated, this approach works well.
- Tracking using flow: selecting good “interest points” and features, and matching them between successive images provide feature tracks that can be clustered into object trajectories. This approach is also called feature-based tracking and has been applied to traffic monitoring in (32) (33), and pedestrian counting in (34).
- Tracking with probability: it is convenient to see tracking as a probabilistic inference problem in a Bayesian tracking framework. In simple cases, independent Kalman filters can be run successfully for each target (Extended Kalman Filters are used for individuals and groups of pedestrians in (35)), but will fail in scenes where the objects interact and

occlude each other. This is called the “data association problem” and can be addressed using particle filters and Markov chain Monte Carlo methods for sampling.

Although great progress has been made in recent years, the tracking performance of the various systems is difficult to report and compare, especially when many of these systems are not publicly available or their details disclosed, and when benchmarks are rare and not systematically used. Tracking pedestrians and mixed traffic in crowded scenes is still an open problem. To the authors’ knowledge, no attempt has yet been made to develop a fully functional video-based pedestrian conflict analysis system. The collected datasets are typically small, and in some cases, e.g. (24), require significant manual input to correct the automated results and supplement with additional data.

VIDEO-BASED SYSTEM FOR AUTOMATED PEDESTRIAN CONFLICT ANALYSIS

This Section describes the development of a video-based system for the automated analysis of pedestrian conflicts. The system has 5 basic components (Figure 1): 1) video pre-processing; 2) feature processing; 3) grouping; 4) high-level object processing; and 5) information extraction. Furthermore, the steps required in preparing the system for use, are discussed in brief.

Camera Calibration

The main purpose of camera calibration is to establish a set of camera parameters in order to find a relationship between world coordinates and image plane coordinates. The inverse transformation that recovers world coordinates of objects in the video images can be obtained from the camera parameters. Camera parameters are classified into extrinsic and intrinsic parameters. Extrinsic camera parameters specify the translation and rotation of the camera’s coordinates relative to world coordinates. Intrinsic parameters are required to establish a perspective projection of objects defined in the camera’s coordinates onto the image plane. Both sets of parameters can be obtained by minimizing the difference between the projection of geometric entities, e.g. points and lines, onto world or image plane spaces, and the actual measurements of these entities in projection space. The mapping from homogeneous world coordinates $\mathbf{P}$ to homogeneous image plane coordinates $\mathbf{p}$ can be described as follows:

$$\mathbf{P} = \mathbf{A}\,[\mathbf{R}|\mathbf{t}]\,\mathbf{P} \tag{1}$$

where $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{t}$ are the intrinsic projection, rotation and translation matrices respectively.

The calibration data used in this study was composed of a set of 22 points selected from salient features in the monitored traffic scene that appear in the video image, as shown in Figure 2(a) and (b). The world coordinates of the calibration points were collected from an orthographic image of the location obtained from Google Maps (36). The intrinsic parameter considered in this study is the camera focal length. The mapping in Equation (1) imposes a reduction in dimensionality due to the projection onto a plane. The inverse projection is defined only if one of the world coordinates, or a relationship thereof, is known. In our application, image plane coordinates are re-projected onto the road surface, i.e. the plane Z=0.

The optimization algorithm used in finding the optimal set of parameters is the Nelder-Mead simplex method available in the Matlab Optimization Toolbox (37). An initial estimate for the camera position was obtained using an approximate position for the camera set-up location and the rotation angles using an orthographic satellite image that contains the camera set-up location and the monitored traffic scene.

The calibration accuracy obtained by applying the previous procedure to a Vancouver intersection (as will be described later in the subsequent section) was satisfactory. The average percentage error in coordinate estimates was less than 1%. The camera calibration problem faced in this study was relatively simple due to the abundance of lane marking features that appear in the orthographic image of the traffic scene.

Figure 3 shows the projection of a sample of pedestrian tracks on an orthographic satellite image of the traffic scene. Similar studies in the literature used artificial construction of an orthographic image using video image rectification e.g. (38). The approach followed in this study by projecting the video data on an independent site map proved helpful in visually verifying the accuracy of the resulting projection - especially with the difficulties faced in obtaining calibration data. In addition, it was possible to collate pedestrian tracks obtained from different camera settings into a single site map, whereas video image rectification produces a setting-dependent site map.

Video Formatting

Depending on the video source, it may be necessary to encode the video in a suitable format for later processing, as well as correct recording artifacts such as interlacing. For this study, a digital video recorder was used that encoded video to a suitable AVI format.

Feature Tracking and Grouping

A feature-based tracking system was initially developed for vehicle detection and tracking as part of a larger system for automated road safety analysis (33)(39). Feature-based tracking is preferred because it can handle partial occlusion. The tracking of features is done through the well known Kanade-Lucas-Tomasi feature tracker (40). Stationary features and features with unrealistic motion are filtered out, and new features are generated to track objects entering the field of view. Since a moving object can have multiple features, the next step is to group the features, i.e. deciding what set of features belongs to the same object, using cues like spatial proximity and common motion. The grouping method described in (41) was extended to handle intersections (33). A graph connecting features is constructed over time. Two parameters are crucial for the success of the method: the connection distance $D_{connection}$, i.e. the maximum distance between two features for their connection, and the segmentation distance $D_{segmentation}$, i.e. the maximum difference between the minimum and maximum distance between two features. The tracking accuracy for motor vehicles was measured to be between 84.7% and 94.4% on three different sets of sequences (33). This means that most trajectories are detected by the system, although over-grouping and over-segmentation can still occur.

High-level Object Processing

Difficulties occur in scenes where the traffic is mixed and the road users have very different sizes, e.g. vehicles and pedestrians, and the connection and segmentation distances can only be adjusted for one type of road user. To address this issue, the original system has been extended by identifying the types of the road users. The parameters are adjusted for pedestrians, and consequently the motorized vehicles are over-segmented. Once the groups of features belonging to motorized vehicles are identified, the feature are processed a second time by the grouping algorithm using larger connection and segmentation distances.

In the current system, a simple test using a threshold on the maximum speed of each road user is sufficient to discriminate between pedestrians and motorized road users in most cases. This test will typically classify bicyclists as motorized road users, which may lead to consider pedestrian-vehicle conflicts that are in fact pedestrian-bicyclist conflicts. Road user classification will be improved in the future by using object classifiers based on background subtraction and image appearance (31).

System Operator and User

The point of an automated system is to minimize user input, especially to eliminate the need for continuous supervising. Global optimization methods to adjust parameters are still lacking, as performance is difficult to evaluate completely automatically. The role of the system operator is therefore to find good parameter values by trial and error, and by visual inspection of the results. Since the world coordinates are recovered, the parameters can be used unchanged in various scenes. The system was developed in an open manner in order to provide data for analysis and visualization purposes. The results are currently stored in plain text files, but could also be stored in a database, and can be mined for the needs of the end user.

DATA COLLECTION AND APPLICATION

The system was tested on traffic video recorded for two days during daytime at a crosswalk in Downtown Vancouver. The objective of the case study is to assess the capability to identify instances of important events, and to calculate severity conflict indicators for each of these events.

Site Description and Data Collection

The study area is the intersection between Pender St. and W. Georgia St. in the Downtown area of Vancouver, British Columbia, Canada. The main interacting movements are pedestrian crossing and left-turn vehicles. Left-turn traffic at signalized intersections poses a particularly increased risk of collision for pedestrians (see the relevant references in (14)). Furthermore, this intersection is unique in that it is a skewed intersection within a corridor grid of streets all containing right-angle intersections. Hence, there is a high possibility of observing an adequate number of important interactions between pedestrians and motorists that involve a risk of collision. In this study, important events occurred when a pedestrian and a vehicle co-existed inside the monitored crosswalk.

A video camera was set on the 6^{th} floor of a building that overlooks the intersection and aimed towards the west. Video recording was conducted for a total of 20 hours over two business days. Approximately, a total of 7000 left-turning vehicles and 2100 pedestrians were observed. These volume estimates are derived from the automated video analysis.

Calculation of Conflict Indicators

The system detects all events constituted by the pairs of pedestrians and vehicles that are in the traffic scene simultaneously. Among these events, this study is interested in important events as defined in the introduction, and traffic conflicts, which are a subset of important events. The complement of important events over the space of all traffic events are defined as undisturbed passages.

In order to compensate for the limitations of individual conflict indicators, four conflict indicators were calculated in this study. One of the most widely used conflict indicators is Time-to-Collision (TTC). TTC is defined as “…the time that remains until a collision between two vehicles would have occurred if the collision course and speed difference are maintained.” (42). An accurate estimation of TTC however requires considerable field measurement of road user positions, speed and direction of movement. This work relies on the traditional operational definition of a collision course, extrapolating the road users' movements with constant velocity (used in (1) for example). This hypothesis is however simple and may lead to unrealistic collision-course estimates as will be discussed later.

Other conflict indicators are used to capture different proximity aspects. Post-Encroachment Time (PET) suggested by Cooper (43) is the time difference between the moment an offending road user leaves an area of potential collision and the moment of arrival of a conflicted road user possessing the right of way. Gap time (GT) is a variation on PET that is calculated at each instant by projecting the movement of the interacting road users in space and time (26). Deceleration-to-Safety Time (DST) is defined as the necessary deceleration to reach a non-negative PET value if the movements of the conflicting road users remain unchanged (44). Allen et al. (45) ranked GT, PET and Deceleration Rate as the primary measures for left-turn conflicts. DST was selected since it captures greater details of the traffic event. TTC was selected since it is the primary traffic conflict indicator in the literature. The values of conflict indicators used in event detection are the minimum TTC, the minimum GT, the maximum DST and PET. Figure 4 shows sequences of severity conflict indicators calculated for a traffic conflict event. *Appendix* 1 shows the description of the method used in this study to calculate these severity indicators in algorithmic form.

Validation

Various manually designed detection conditions defined over the composite values of the severity conflict indicators are used to identify automatically important events. These results are compared on a sample of events manually classified by a human observer, using the definition of important events given in this paper and the US FHWA observer's guide (45). The pre-condition for an important event to occur in this study is that a left turning vehicle enters the monitored crosswalk in the presence of a pedestrian or a group of pedestrians already in the crosswalk. Excluded were the events that involved the following unlikely chain of events: a vehicle reverting its travel direction, a pedestrian changing movement from walking to running (> 3.5 m/s), and a collision involving pedestrians standing beyond the curb line.

Sources of mismatch that can lead to inaccurate indicator values and misclassifications of traffic events are:

1. Errors in pedestrian and vehicle detections. These errors include: noise in tracked object position that could lead to unrealistic extrapolation of a road user's position, multiple detection of the same road user, lost detections of a road user, appearing or disappearing during a traffic event.
2. Incapability of the used conflict indicators to measure the level of severity of a traffic event.

While in some cases, it was evident why the erroneous classification of the traffic event took place, it was difficult in other cases to explicate the error source. In order to follow an objective evaluation, the overall performance of the system was considered with respect to detecting and tracking road users, as well as making judicious use of the severity information measured by the conflict indicators.

In this study, the detection conditions used for identifying conflicts and important events are defined by scaling serious conflict threshold values that delimit serious conflicts from other traffic events by a severity factor. Table 1 shows the details of the detection conditions and the summary of detection results for various severity factor values. The total number of conflict events in the analyzed video sequence is 17. The number of traffic conflicts considers the actual number of pedestrians involved, e.g. a conflict involving a vehicle and two pedestrians is counted as two conflicts.

Only PET may allow detecting important events as well as conflicts separately from the other indicators. This is consistent with a study in the literature that used PET for conflict detection (20). Other conflict indicators however could not solely detect an adequate percentage of important events and traffic conflicts. A combination of the four conflict indicators could enable the system to automatically capture 89.5% of the conflicts and 71.7% of important events while however detecting 54.5% of undisturbed passage events as important events.

DISCUSSION

One of the functional purposes of the developed system described in this paper is to automatically identify important pedestrian-vehicle events, including conflicts, and relay their record to a human observer for further examination. Combining information from four conflict indicators proved successful in reporting the majority of conflicts identified by a human observer. Figure 5 shows sample frames of important events automatically detected by the system.

The capability of each conflict indicator to characterize important events was compared to manually annotated events in the dataset. As shown in Table 1, none of the conflict indicators was solely capable of capturing important events. The following limitations of the selected conflict indicators were identified in this study:

1. A prerequisite for TTC is also the existence of road users on a collision-course, that is vehicles will collide if their "movements remain unchanged" (1). The existence of a collision-course is not however a necessary condition for capturing "dangerous proximity." Some dangerous interactions could not be captured by TTC because the involved road users were not on a collision-course. A typical case occurs when a motorist passes behind a pedestrian at a perilously close distance. A perturbation however of the speed or direction of movement of the motorist, or slight delay on the part of the pedestrian, could potentially create a collision-course.
2. The extrapolation of road users' movements with constant speed and direction could lead to erroneously small values of TTC and DST. Figure 6 shows the distribution of calculable values of *min* TTC in conflicts and regular events. It is observable that while TTC can function as a severity measure, it overestimates the actual conflict severity in many events. A typical situation occurs when a pedestrian is considered on a collision course with turning vehicles of which the velocity vector happens to point at the pedestrian. However, this method of road users' movement extrapolation is widely used in the literature.
3. PET was the most reliable parameter for detecting important events. Despite its simple definition, PET has inherent drawbacks in its ability to accurately capture conflict severity. Events in the video sequence in which motorists decelerated to near-stop to avoid collision usually have PET values that do not reflect the true severity of the event.

A potential improvement to current conflict indicators is to consider the continuum of all possible actions by road users in a probabilistic framework. Recent work discussed the representation of conflicts and collision in a single theoretical framework that considers the different possibilities of evasive actions (47). The establishment of the distribution of possible movements requires a data-intensive study of behavioral and situational aspects of road users during normal driving conditions as well as traffic conflicts. The video analysis system presented in this paper has been used to demonstrate this approach, extracting the typical motion patterns of road users to compute the collision probability of any pair of interacting road users (48).

CONCLUSIONS AND FUTURE WORK

This paper presents an automated system and methodology that furthers the development of previous work on video analysis to capture the movements of pedestrians at crossing locations. The movement paths of pedestrians and transversal trajectories of vehicles were analyzed and a group of conflict indicators were calculated for each pedestrian-vehicle interaction. The system provides the ability to automatically calculate conflict indicators and report important interactions to a human observer for further examination of traffic interactions. The quality of four conflict indicators, Time-to-Collision, Post-Encroachment Time, Gap Time, and Deceleration-to-Safety Time, were assessed in regard to their ability to comprehend the severity of traffic conflicts. None of the conflict indicators were capable of capturing all dangerous interactions between road users alone. However, a combination of the four indicators proved useful in the identification of important events and traffic conflicts. A planned continuation of this work involves the collection of additional video data at traffic intersections with high pedestrian-involved collision hazard potential. Future work also includes testing, as well as improving, the system’s accuracy to detect and track road users in more crowded traffic scenes. As evidenced in this study, there is a need to develop safety measures that address the limitations of current conflict indicators, and draw on the extensive movement data made available by automated methods, such as the automated video analysis system described herein.

ACKNOWLEDGMENT

The authors would like to thank Sam Young and staff of the MMM Group (Vancouver office) for supporting the video survey, and Hsu Hua Lu for assisting in the manual annotation of the video sequences.

REFERNCES

1. *Estimating the severity of safety related behaviour.* **Svensson, Å. and Hydén, C.** 2006, Accident Analysis and Prevention, Vol. 38, pp. 379-385.
2. *British Columbia Traffic Collision Statistics.* s.l. : Insurance Corporation of British Columbia, 2005. 1203-8008.
3. http://www.tc.gc.ca/roadsafety/tp/tp13951/2001/page2.htm. [Online] Transport Canada, 2008.
4. *Traffic conflict standards for intersections.* **Sayed, T. and Zein, S.** s.l. : Transportation Planning and Technology, Vol. 22, pp. 309–323.
5. *Simulation of Traffic Conflicts at Unsignalised Intersections with TSC-Sim.* **Sayed, T., Brown, G. R. and Navin, F.** 5, s.l. : Accident Analysis & Prevention, 1994, Vol. 26, pp. 593-607.
6. *Traffic Conflict Techniques and Other Forms of Behavioural Analysis: Application to Safety Diagnoses.* **Muhlrad, Nicole.** 1993.
7. **Perkins, S. and Harris, J.** *Criteria for Traffic Conflict Characteristics.* s.l. : General Motors Corporation, 1967. Report GMR 632.

8. **Amundson, F. and and Hydén, C.** Oslo : Proceedings of First Workshop on traffic Conflicts, Institute of Economics, 1977.
9. *Development of Conflicts Analysis Technique for Pedestrian Crossings.* **Cynecki, M. J.** 1980, Transportation Research Record, Vol. 743.
10. **Migletz, D.J., Glauz, W.D. and Bauer, K.M.** *Relationships between traffic conflicts and accidents.* s.l. : US Department of Transportation, Federal Highway Administration, 1985. FHWA/RD-84/042.
11. **Fitzpatrick, Kay, et al.** *TCRP Report 112/NCHRP Report 562. Improving Pedestrian Safety at Unsignalized Crossings.* Transportation Resarch Board. 2006.
12. *Pedestrian Walking Speed, Step Size, and Step Frequency from the Perspective of Gender and Age: Case Study in Beijing, China.* **Hui, Xiong, et al.** 2007. Source Data: Transportation Research Board Annual Meeting 2007 Paper No. 07-1486.
13. **Messelodi, S. and Modena, C. M.** *A computer vision system for traffic accident risk measurement: A case study.* s.l. : Technical Report, ITC-irst , ITC, 2005. T05-06-07.
14. *Analysis of pedestrian conflicts with left-turning traffic.* **Lord, D.** 1996, Transportation Research Record, Journal of the Transportation Research Board, Vol. 1538.
15. *Effect on Vehicle-Pedestrian Conflicts of "Turning Traffic Must Yield to Pedestrians" Sign.* **Abdulsattar, H N, et al.** s.l. : Transportation Research Record: Journal of the Transportation Research Board, 1996, Vol. 1553, pp. 38-45.
16. *Using auditory pedestrian signals to reduce pedestrian and vehicle conflicts.* **Van Houten, R., et al.** 1997, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1578, pp. 20-22.
17. *Conflict analysis for prediction of fatal crash locations in mixed traffic.* **Tiwari, G., Mohan, D. and Fazio, J.** 1998, Accident Analysis and Prevention, Vol. 30, pp. 207-215.
18. *Parameters for evaluating pedestrian safety problems in signalised intersections using the traffic conflict analysis technique—a study in Sao Paolo, Brazil.* **Tourinho, L.F.B. and Pietrantonio, H.** 3, s.l. : Transportation planning and technology, 2003, Vol. 29, pp. 183-216.
19. *Reducing conflicts between motor vehicles and pedestrians: The separate and combined effects of pavement markings and a sign prompt.* **Huybers, S., Houten, R. Van and Malenfant, J. E. Louis.** 2004, Journal of applied behavior analysis, Vol. 37, pp. 445-456.
20. *The development of an automatic method of safety monitoring at Pelican crossings.* **Malkhamaha, S., Miles, T. and Montgomery, F.** 2005, Accident Analysis and Prevention, Vol. 37, pp. 938-946.
21. *In-Street Pedestrian Crossing Signs and Effects on Pedestrian-Vehicle Conflicts at University Campus Crosswalks.* **Medina, Juan, Benekohal, Rahim and Wang, Ming-Heng.** 2008.
22. *Pedestrian, Bicycle, and Vehicle Traffic Conflict Management in Big Ten University Campuses.* **Rodriguez-Seda, Jarice D., Benekohal, Rahim F. and Morocoima-Black, Rita.** 2008. Transportation Research Board.
23. *Pedestrian/Vehicle Conflicts: An Accident Prediction Model.* **Davies, S., Robertson, D. H. and King, E. K.** 1989, Transportation Research Record: Journal of the Transportation Research Board, TRB, National Research Council, Washington, D.C., Vol. 1210, pp. 1–11.
24. *Empirical Study of Pedestrian-Vehicle Interactions in the Vicinity of Single-Lane Roundabouts.* **Chae, K. and Rouphail, N. M.** 2008. Transportation Research Board Annual Meeting Compendium of Papers. 08-2898.

25. *Measurement of Traffic Conflicts.* **Chin, C. H. and Quek, S. T.** 1997, Safety Science, Vol. 26, pp. 169-187.
26. **Archer, J.** *Methods for the Assessment and Prediction of Traffic Safety at Urban Intersections and their Application in Micro-simulation Modelling.* Royal Institute of Technology. 2004. Academic Thesis.
27. *Computational Studies of Human Motion: Part 1, Tracking and Motion Synthesis.* **Forsyth, D.A., et al.** 2005, Foundations and Trends in Computer Graphics and Vision, Vol. 1, pp. 77-254.
28. *Behavioral Priors for Detection and Tracking of Pedestrians in Video Sequences.* **Antonini, Gianluca, et al.** 2006, International Journal of Computer Vision, Vol. 69, pp. 159-180.
29. *Using Low-Cost Infrared Detectors to Monitor Movement of Pedestrians: Initial Findings.* **Kerridge, J., et al.** 2004, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1878, pp. 11-18.
30. *Video-Based Monitoring of Pedestrian Movements at Signalized Intersections.* **Malinovskiy, Yegor, Wu, Yao-Jan and Wang, Yinhai.** 2008.
31. *Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors.* **Wu, B. and Nevatia, R.** s.l. : Kluwer Academic Publishers, 2007, International Journal of Computer Vision, Vol. 75, pp. 247-266.
32. *Vehicle Segmentation and Tracking in the Presence of Occlusions.* **Kanhere, N. K., Birchfield, S. T. and Sarasua, W. A.** 2006. Transportation Research Board Annual Meeting Compendium of Papers.
33. *A feature-based tracking algorithm for vehicles in intersections.* **Saunier, N. and Sayed, T.** s.l. : IEEE, 2006.
34. *Counting Crowded Moving Objects.* **Rabaud, V. and Belongie, S.** 2006. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Vol. 1, pp. 705-711.
35. *Estimating pedestrian counts in groups.* **Kilambi, P., et al.** s.l. : Elsevier Science Inc., 2008, Computer Vision and Image Understanding, Vol. 110, pp. 43-59.
36. *Google Maps.* s.l. : Google.
37. **Mathworks.** Optimization. *Mathworks Products.* [Online] 2008. http://www.mathworks.com/products/optimization.
38. *Automated video analysis as a tool for analysing road user behaviour.* **Laureshyn, Aliaksei and Ardö, Håkan.** 2006. ITS World Congress.
39. *Automated Road Safety Analysis Using Video Data.* **Saunier, N. and Sayed, T.** s.l. : Transportation Research Record: Journal of the Transportation Research Board, 2007, Vol. 2019, pp. 57-64.
40. **Birchfield, S. T.** KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker. http://www.ces.clemson.edu/~stb/klt/.
41. *A Real-time Computer Vision System for Measuring Traffic Parameters.* **Beymer, D., et al.** s.l. : IEEE Computer Society, 1997. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. pp. 495-501.
42. *Near-miss determination through use of a scale of danger.* **Hayward, J. Ch.** 1968, Highway Research Record, Vol. 384, pp. 24–34.
43. *Experience with traffic conflicts in Canada with emphasis on post encroachment time techniques.* **Cooper, J.** 1984. Vol. vol. F5, pp. 75–96.
44. *Deceleration to safety time (DST)—a useful figure to evaluate traffic safety.* **Hupfer, C.** 1997.

45. *Analysis of traffic conflicts and collision.* **Allen, B.L., Shin, B.T. and Cooper, D.J.** s.l. : Transportation Research Record: Journal of the Transportation Research Board, 1978, Vol. 667, pp. 67-74.
46. **Parker, M. and Zegeer, C. V.** *Traffic Conflict Techniques for Safety and Operations Observer's Manual.* US Department of Transportation, Federal Highway Administration. McLean, Virginia : Federal Highway Administration, 1989. FHWA-IP-88-027.
47. *Outline of Causal Theory of Traffic Conflicts and Collisions.* **Davis, G., Hourdos, J. and Xiong, H.** Washington DC : Transportation Research Board, 2008. TRB 87th Annual Meeting Compendium of Papers DVD.
48. *A Probabilistic Framework for the Automated Analysis of the Exposure to Road Collision.* **Saunier, N. and Sayed, T.** 2008. Transportation Research Board Annual Meeting Compendium of Papers.

LIST OF TABLES

LIST OF FIGURES

TABLE 1 Summary of Validation Results

Identification Conditions[1]	Percentage of each event types correctly identified by the system: Traffic Conflict[2]	Important Events[3]	Uninterrupted Passages	Percentage of undisturbed passage falsely identified by the system as important events
$a_{TTC} = 1$	5.3	4.3	93.2	6.8
$a_{TTC} = 2$	31.6	23.9	87.2	12.8
$a_{TTC} = 5$	36.8	39.1	66.0	34.0
$a_{DST} = 1$	0.0	0.0	100.0	0.0
$a_{DST} = 2$	5.3	3.3	96.6	3.4
$a_{DST} = 5$	47.4	51.1	63.0	37.0
$a_{GT} = 1$	21.1	27.2	80.4	19.6
$a_{GT} = 2$	26.3	32.6	75.7	24.3
$a_{GT} = 5$	42.1	41.3	66.0	34.0
$a_{PET} = 1$	5.3	0.0	99.6	0.4
$a_{PET} = 2$	10.5	2.2	98.3	1.7
$a_{PET} = 5$	89.5	42.4	88.5	11.5
$a_{PET} = 1$ OR $a_{GT} = 1$ OR $a_{DST} = 1$ OR $a_{TTC} = 1$	21.1	28.3	74.5	25.5
$a_{PET} = 2$ OR $a_{GT} = 2$ OR $a_{DST} = 2$ OR $a_{TTC} = 2$	36.8	43.5	67.2	32.8
$a_{PET} = 5$ OR $a_{GT} = 5$ OR $a_{DST} = 5$ OR $a_{TTC} = 5$	89.5	71.7	45.5	54.5

1The thresholds of the identification definitions are determined by scaling the serious conflict threshold on each severity indicators by a severity factor a_X, where the subscript X refers to the concerned conflict severity indicator. The following typical severity thresholds are taken from the literature: 1.5 s, 3 m/s^2, 1s, and 1s, for TTC, DST, PET, and GT respectively. For TTC (and similarly for PET and GT), all events that involved TTC < 1.5 x a_{TTC} are detected as important events. For DST, all events that involved DST < 1.5 / a_{DST} are detected as important events. Thus defined, higher severity factors would cover events with lower conflict severity. Increasing the factors lead to a higher chance of detecting conflicts at the expense of misclassifying undisturbed passages as important events. If a severity factor is not mentioned for a indicator, it means that it is not used in the condition.

2 Observer-based conflict identification was performed according to the US FHWA Observer Manual (44).

3 The definition of an important interaction is an event that involves a crossing pedestrian and a conflicting vehicle in which there exists a conceivable chain of events that could lead to a collision between these road users. The pre-condition for an important event to occur in this study is that a left turning vehicle enters the monitored crosswalk in the presence of a pedestrian or a group of pedestrians already in the crosswalk. Excluded were the events that involved the following unlikely chain of events: a vehicle reverting its travel direction, a pedestrian changing movement from walking to running (> 3.5 m/s), and a collision involving pedestrians standing beyond the curb line.

FIGURE 1 Layout of the pedestrian detection and tracking system. The figure shows the five main layers the make up the system. Depicted also is the data flow among system modules from low-level video data to a position database of detected, tracked, and classified road users.

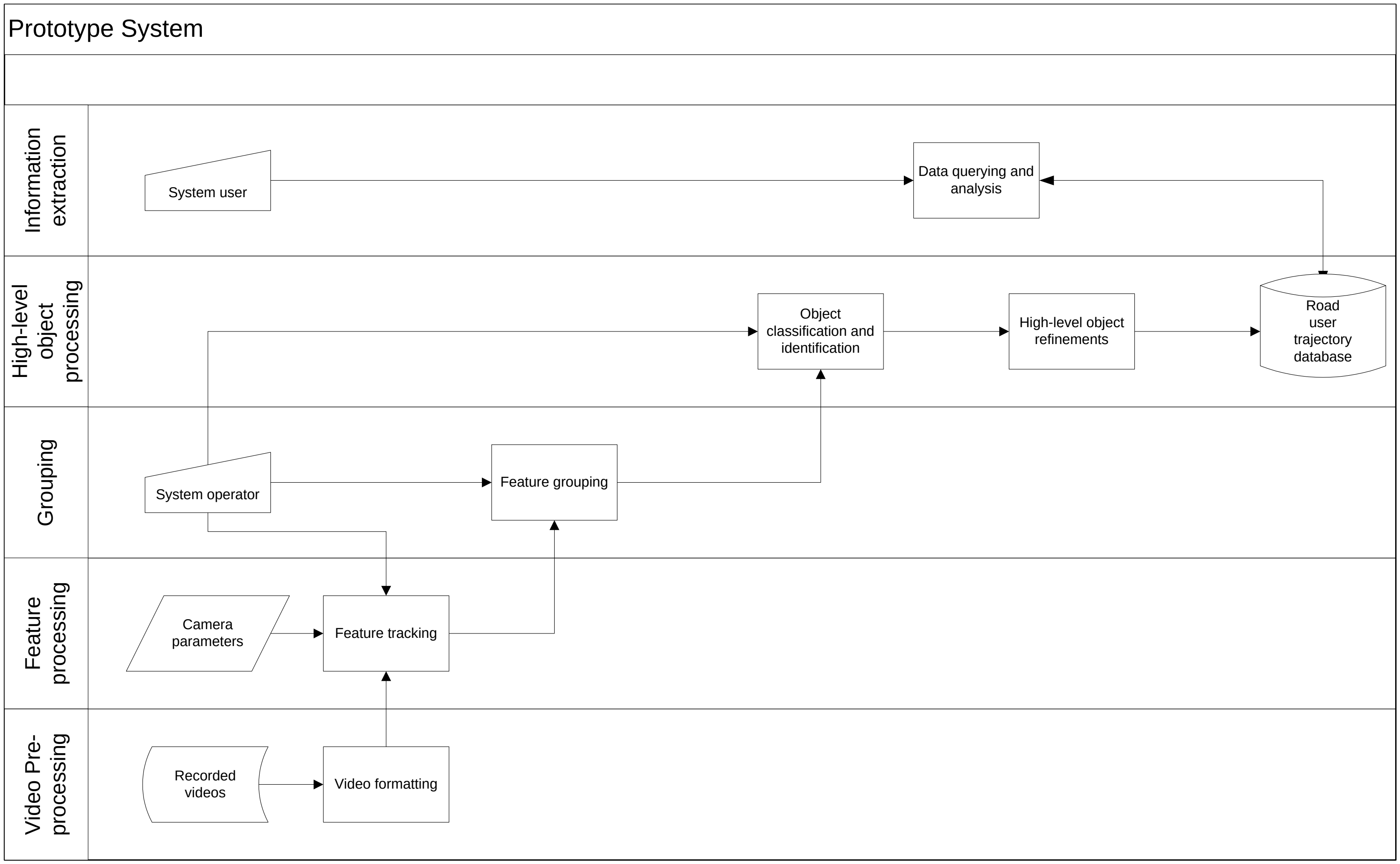


FIGURE 2 The 22 points used to estimate the camera calibration are displayed on a video frame (figure a) and on an orthographic satellite image of the traffic scene (figure b). Bulleted points (●) are manually annotated and x-shaped points (x) are projections of annotated points using the calculated camera parameters.

FIGURE 3 A sample of pedestrian tracks is projected on an orthographic satellite image of the traffic scene. Vehicle tracks are depicted in red and pedestrian tracks are in black.

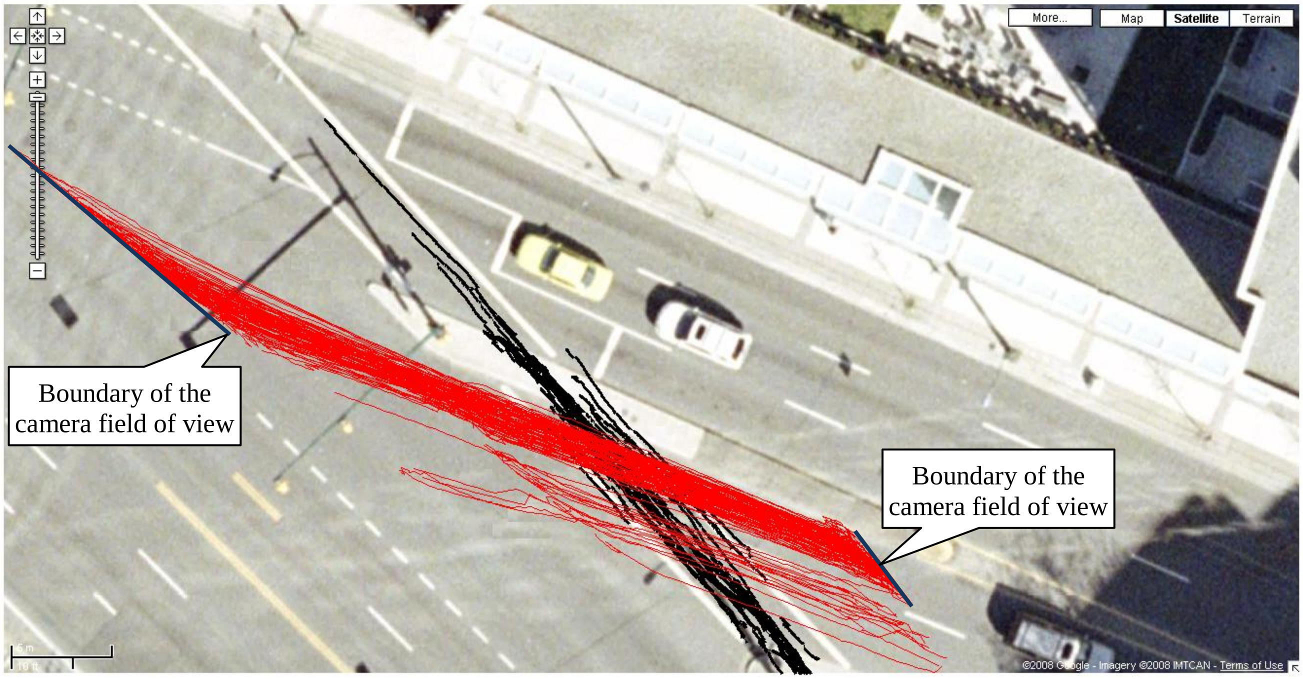


FIGURE 4 Conflict indicators for a sample traffic event. The left figure describes the traffic event shown in figure 5 (a). The right figure describes the traffic event shown in figure 5 (b).

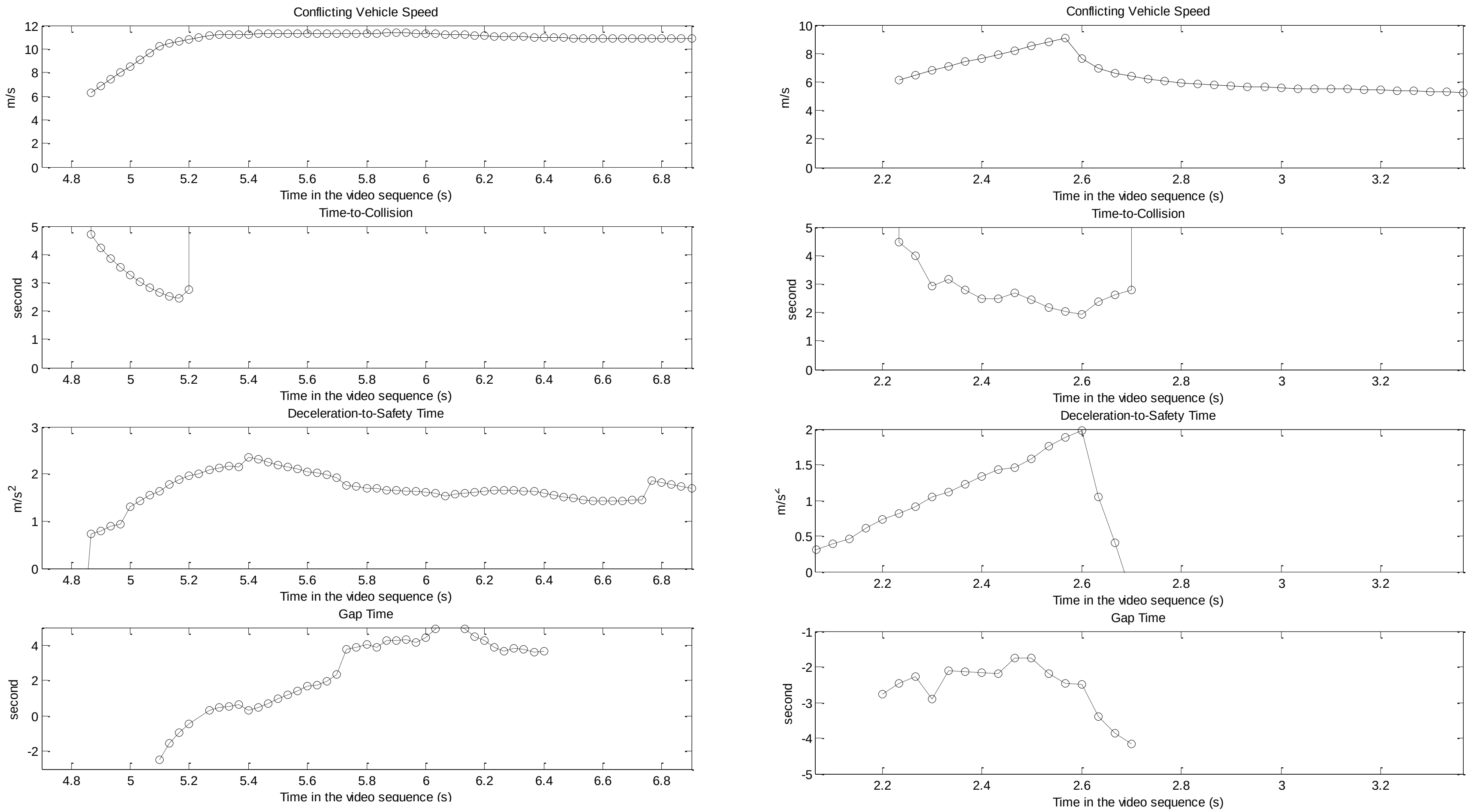


FIGURE 5 Sample of automatically detected important events with the road users' trajectories. The numbers under each image are respectively the min TTC (seconds), PET (seconds), maximum DST (m/s[2]), and min GT (s). In the images, the road user speed is indicated in m/s.

a)

2.43 | 3.63 | 2.34 | -2.47

b)

1.93 | 2.13 | 1.98 | -4.17

c)

1.27 | 3.17 | 2.83 | 0.30

d)

2.03 | 2.80 | 3.34 | 0.03

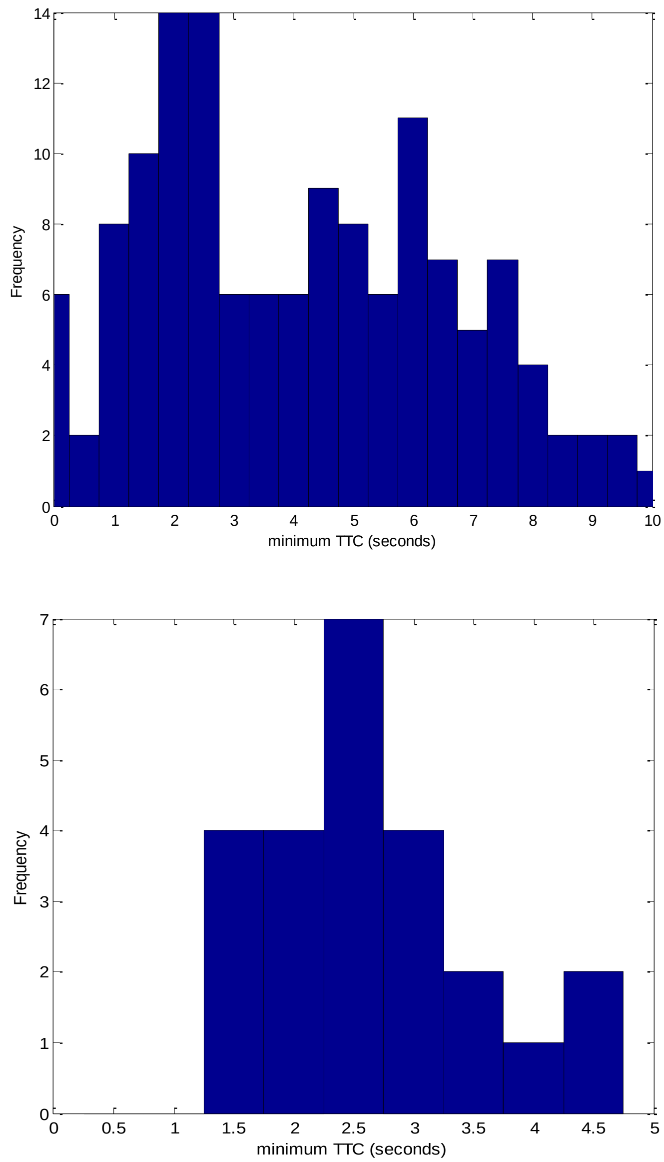
e)

1.70 | 4.00 | 1.78 | 0.57

f)

5.73 | 3.87 | 2.38 | 0.77

FIGURE 6 Distribution of the minimum Time-to-Collision (seconds) respectively for all events for which it could be computed (top) and for all manually annotated important events (bottom).



APPENDIX 1

<u>**Algorithm 1:** Algorithm for calculating conflict indicators for a pedestrian-vehicle event</u>

Definitions: 1) A generic position function $F: N^+ \rightarrow \Re^2$ returns the world-space position of a road user (x,y) at time instant t such that $F(t) = (x,y)$.

2) A generic velocity function $\dot{F}: N^+ \rightarrow \Re^2$ returns the velocity components of a road user $(\dot{x},\dot{y})$ at time instant t such that $\dot{F}(t) = (\dot{x},\dot{y})$.

3) A generic position extrapolation function $E: N^+ \times \Re^2 \times \Re^2 \rightarrow \Re^2$ returns the position at time $t + \Delta t$ of a road user with current position (x,y) and velocity $(\dot{x},\dot{y})$ at time t, $E(\Delta t, x, y) = (x,y) + (\dot{x},\dot{y}) * \Delta t$

Input: Let $P: N^+ \rightarrow \Re^2$ be the pedestrian position function, defined for $t \in [t_{p1}, t_{p2}]$

Let H_{f1}, H_{f2}, H_{r1} and H_{r2} be the position functions of the vehicle front and rear corners respectively, that are all defined for $t \in [t_{v1}, t_{v2}]$

Let $\dot{P}$ and $\dot{H}$ be the pedestrian and vehicle velocity functions, respectively

1- Let W be the segment demarcating the crosswalk that is furthest from the vehicle

Let $c_1 = 0.25\, m/s$ be a speed threshold and $\Delta t_e = 10\, s$ be the maximum extrapolation time.

Output: Time series of **TTC**, **DST**, and **GT**, and the PET

begin

for each pair consisting of a pedestrian and a vehicle whose observed trajectories intersect at a point P_0

Let T_0 be the times at which each road users occupies P_0

Find the times T_1 at which the observed vehicle rear corner positions H_{r1}, H_{r2} intersect W

PET=$max\, T_1$ -$min\, T_0$

for each $t \in [\max(t_{p_1}, t_{v_1}), \min(t_{p_2}, t_{v_2})]$ such that $\|\dot{P}(t)\| \geq c_1$ AND $\|\dot{H}(t)\| \geq c_1$

Find the intersection points P_2 between the extrapolated positions of the pedestrian $\{E(\Delta t, P(t)) \mid 0 \leq \Delta t \leq \Delta t_e\}$ and of the vehicle front corners $\{E(\Delta t_e, H_f(t)) \mid 0 \leq \Delta t \leq \Delta t_e\}$ for $f \in \{f_1, f_2\}$

Find the intersection points P_3 between the extrapolated positions of the vehicle rear corners $\{E(\Delta t_e, H_r(t)) \mid 0 \leq \Delta t \leq \Delta t_e\}$ for $r \in \{r_1, r_2\}$ and W

2- Find the times T_2 and T_3 at which each road user occupies the intersection points in P_2 and P_3

Calculate Δt =**TTC**(t) such that $E(\Delta t, P(t))$ lies inside the extrapolated positions of the vehicle outline

Calculate **GT**$(t) = min\ T_2 - max\ T_3$

and **DST**$(t) = 2\dfrac{(\max T_2 - t\ ) * \left\|\dot{H}(t)\right\| - \max_{p \in P_2}\left\|P(t) - p\right\|}{(\max T_2 - t\ )^2}$

if the pedestrian leaves the conflict area before the vehicle **then**

Recalculate **GT**(t) and PET such that it is the time between the instant a pedestrian clears the conflict area and the instant of arrival of the front of the conflicting vehicle arrival.

Notes:

1- This definition of a "conflict area" is adopted from Lord (14)
2- Several algorithmic details were implemented to deal with tracking errors, e.g. tracked objects that are detected or lost during the traffic event. Details are omitted for brevity.