Automated Pedestrian Safety Analysis Using Video Data in the Context of Scramble Phase Intersections

Karim Ismail, M.A.Sc., PhD. Student

Tarek Sayed, P.Eng., Ph.D.

Nicolas Saunier, Ph.D.

The University of British Columbia

Paper prepared for presentation

at the ________________Session

of the 2009 Annual Conference of the
Transportation Association of Canada

Vancouver, British Columbia

Acknowledgment

The video data used in this research was obtained in a previous study conducted at the Institute of Transportation Studies, UC Berkeley Traffic Safety Center. The key contact persons were Jenna Hua and Prof. David Ragland. The authors would like to sincerely thank them. Observation of the traffic conflicts was conducted by Varun Ramakrishna at UBC and Indian Institute of Technology, Madras, Chennai.

Automated Pedestrian Safety Analysis Using Video Data in the Context of Scramble Phase Intersections

Abstract:

Pedestrian data is critical for supporting safety studies and proper design of pedestrian facilities. There has been a traditional shortcoming in the transportation literature with regard to the availability of and quality of pedestrian data. A potential reason is that current data collection methods mainly rely on human observers which make them costly and prone to observer-based errors. In addition, automated data collection methods currently in practice are mainly developed to serve motorized traffic with few proven applications at similar level of functionality for pedestrians. This paper describes a novel technique for conducting road safety analysis using an automated video analysis system. The system was tested using video data collected from a scramble phase intersection in the downtown area of Oakland California. The context of scramble phase is unique in that conflict indicators have to be developed further in order to represent the severity of interaction outside crosswalks. This study demonstrated the feasibility of using computer vision techniques to automatically analyze pedestrian-vehicle conflicts. The ability to analyze pedestrian-vehicle conflicts will potentially reduce the cost of conflict surveys and improve the quality and quantity of safety-related data. In addition, pedestrian walking speed data was collected using other video sequences at high accuracy. An accurate measurement of pedestrian walking speed is critical for understanding pedestrian behaviour at crosswalks and for properly designing traffic signals. To the best of the authors' knowledge, there is little, if any, work similar to this study that exists in the technical literature in regard to the method of analysis and context of scramble phase intersection.

Introduction

Background and Study Aims

> *"[Exposure to the risk of collision is] very difficult to measure directly, since this would involve tracking the movements of all people at all times"* [1].

In addition to exposure it is also infeasible to obtain accurate estimates of other quantities fundamental to road safety analysis without analyzing positions of road users in space and time. From a broad perspective, this paper demonstrates a novel application of advanced video analysis techniques to address traditional limitations in collecting traffic data and to tackle challenging road safety applications. Video sensors are selected as the primary source of data in this research. Video data is rich in details, recording devices are becoming less expensive, video cameras are often already installed for traffic monitoring purpose, and automated analysis is possible using techniques developed in the field of computer vision. This research effort is positioned at the forefront of advanced traffic monitoring technologies and draws on recent developments in the realm of computer vision. This research is also at another frontier in the realm of road safety by attempting to automate traffic conflict techniques to address their cost and their error-prone implementation [2]. While the set goals are ambitious, current state-of-the-art techniques in road safety analysis and computer vision cannot readily support proved applications. Also, being simultaneously at two research frontiers is as ambitious as it is challenging. Despite the technical challenges, promising results have been obtained. This paper reports a part of ongoing research work with main focus on the study of pedestrian safety. Before delving into the research details, the motivation, significance, and focus points of this study are discussed.

Technological Impact

Extracting road user positions in space and time, tracks, from video sequence enables road safety analysis at a much higher spatial and temporal resolution than current techniques available in practice. Using computer vision techniques, precise positional measurement is also possible in an efficient way. Video data can be analyzed at a degree of automation that enables dealing with large data volumes while consuming little staff and time resources. While a well established research stream tries to confer some aspects of "intelligence" to transportation systems, this research work can be viewed as an attempt to equip it with "vision".

Special Focus on the Pedestrian

Extracting road user tracks from videos is of special importance for non-motorized modes of travel. Traditionally, non-motorized modes of transport, and walking in particular, have received research and practice focus secondary to motorized modes. As issues of urban sustainability and volatility of the oil markets have come to the fore of the public and policy makers, researchers, planners and practitioners in the field of transportation engineering are placing more emphasis on mass and non-motorized modes of travel. Another motivation for traditional modes of travel such as public transit comes from the current economic challenges. Often the case, the use of public transit is linked to other non-motorized modes of travel, especially walking.

Walking is the most important non-motorized mode of travel and is almost certainly part of any multi-modal trip. Pedestrians are aptly described as the lifeblood of urban areas [2]. Walking modal share in Vancouver, 17% of total trips, is the highest in Canada and one of the highest in North America. It equalizes transit mode share and exceeds carpooling. An emerging emphasis on the study of pedestrians in transportation engineering is manifested in recent research directives and policies, such as Transport Canada's ecoMobility. The program aims to reduce traffic emissions by promoting less polluting forms of transportation such as walking and cycling. The second Strategic Highway Research Program of the US is a $205 million research program established in 2006 with primary focus on exploring complex issues surrounding the human side of the highway system.

While non-motorized modes of travel are groomed as key alternatives for improving the sustainability of the transport system, they suffer an increased risk of collision involvement. Collisions involving non-motorized traffic, often vulnerable road user, are highly injurious and physically damaging. In the developing world, the vulnerability of non-motorized modes and the little attention paid by policymakers to these modes of travel make the situation especially dangerous. A review of 38 studies from developing countries reported that pedestrian fatalities were highest among all modes in 75% of the studies accounting for 41% to 75% of all fatalities [3]. The problem of pedestrian vulnerability is also present in developed countries. Approximately 22% of fatal road collisions in Canada and 30% in British Columbia involve vulnerable road users, 13% and 15% of which are pedestrians, respectively [4]. As society is becoming more aware of the importance of non-motorized modes of transport, existing safety issues with these modes of travel will receive increased attention from researchers and practitioners.

Shortcomings in the Current State of Knowledge

Data Availability

Despite the significant role walking plays in any multimodal transportation network, pedestrian traffic is often understudied in favour of motorized modes of transportation [5]. Availability of road safety data in what relates the study of pedestrians has been identified as a major and common challenge for practical studies in numerous sources, e.g. [6]. Examples of data needs include pedestrian volume and measures of exposure to collision risk which are often expensive and time-consuming to collect [7]. Surrogates and/or statistical predictors of these types of data are often used in practice, e.g. [8] [1]. While in practice there are developed technologies and proven applications for motorized traffic count, this is not the case for pedestrian traffic [9]. Pedestrians move in a less organized fashion, at higher densities, and in more complex and constrained spaces than vehicular traffic. Thus existing issues with data availability are compounded due to the lack of reliable automated methods.

Issues with Collision-based Safety

Road collisions are by far the most costly and most dangerous events that occur on roads. Road collisions are predicted to climb from the 10th to the 8th most common cause of death by 2030 [10]. The global number of road collision fatalities was approximately 1.3 million in 2004 [11] costing the global economy approximately $US 500 billion per year [12]. The economic cost of road collisions to Canada is immense. A recent study by Transport Canada estimates the annual cost of road collisions to the Canadian economy, including health care, environmental damage, lost productivity, and induced traffic congestion, is $ 62.7 billion [13], an enormous 5% of the Canadian GDP. This is far more than the approximately $3.7 billion annual cost of traffic congestion [14].

Despite the prominence of collisions-based safety analysis, there are several limitations to this approach due to the following problems with collision data [15]:

1. **Attribution**. The information obtained by police reports and interviews often does not allow the attribution of road collisions to a single cause. It is sometimes difficult to pinpoint the failure mechanism that lead to a road collision. In that, the safety analyst is often required to remedy or prevent events which causes are not precisely known.
2. **Data Volume**. Despite the enormous social burden of road collisions, the frequency of road collision, especially in disaggregate data form, is low. Drawing statistically stable and significant inferences from such data is typically challenging and in some cases controversial.
3. **Data Quality**. Road collision reporting is based on post-hoc description, witness accounts, and site observations. The process is fundamentally deductive and subjective. Collision records are often incomplete and lack details. The quality of road collision reporting has been deteriorating in many jurisdictions. Reporting is also biased toward highly damaging, while non-injurious collisions can go unreported.
4. **Ethical Concern**. While the object of road safety analysis is the reduction of the risk of road collisions, it is typically based on the road collision as the main data unit. That is, collisions have to occur and be recorded over an adequately long period in order to conduct safety diagnosis. Typical before-and-after studies require the observation of road collisions for a period of several years in order to draw informative information from road collision data. For the identification of hazardous locations to be proper, several years of road collision observations have to be available. The previous limitations of collision data give rise to a paradoxical situation in which the safety analyst, for the sake of methodological correctness, strives to observe events that are ought to be prevented!

Shortcomings in road collision data are even more pronounced in the study of pedestrian safety. Pedestrian-involved collisions are more injurious and less frequent. Exposure data, such as pedestrian volume, is often difficult to obtain and expensive to collect.

A Case for Traffic Conflict Techniques

The study of pedestrians draws its importance from the physical vulnerability of this type of road users and the key role that walking activities play in a sustainable transportation system. Often times, the road safety analyst is humbled by the challenges of studying pedestrian safety. One part of this challenge comes from the dearth and cost of pedestrian exposure data. Arguments that support the adoption of traffic conflicts techniques are of particular relevance and find more ground in the study of pedestrian safety. Traffic Conflict Techniques (TCTs) involve observing and evaluating the frequency and severity of traffic conflicts at an intersection by a team of trained observers [16]. Traffic conflicts are more frequent and much less costlier, if any, than road collisions. Traffic conflicts are observable events that can give insight into the failure mechanism that leads to road collisions. Before-and-after studies based on traffic conflicts can be conducted over shorter periods. This is a key advantage over collision-based analysis since this timely conduct of TCTs enables quick evaluation of engineering treatments. TCTs also have construct validity. A common theoretical framework ranks all traffic interactions by their severity[1] in a hierarchy, with collisions at the top, undisturbed passages at the bottom, and traffic conflicts in between [17]. Therefore, with knowledge of the underlying severity distribution, it is possible to draw inference on road collisions by observing traffic conflicts. The relationship between conflicts and accidents has also been demonstrated [2].

Issues of the current level of development of traffic conflict analysis arise as well in the study of pedestrian-vehicle conflicts. Inter- and intra-observer variability is a common challenge for the repeatability and consistency of results from traffic conflict surveys [18]. Field observations are costly to conduct and typically require trained staff observers. Despite decades of conceptual developments, there is no universal *operational* definition[2] of a traffic conflict [15]. Finally, the estimation of proximal conflict indicators (as shorthand will be called conflict indicator), such as Time to Collision [19] can be difficult using field observations. The proposed research overcomes most of the previous challenging by automating traffic conflict analysis and relying on objective conflict indicators.

Study Purpose, Objectives, and Significance

Based on the previous introduction, the following key arguments can be stated in support of this research work:

1. Focus in transportation engineering is shifting toward sustainable modes of travel. For many reasons, research on non-motorized modes, especially walking, of travel is gaining momentum.
2. There is a need for an automated video-based data collection method for pedestrians. Data availability has been a common challenge to pedestrian studies.
3. Qualitative and quantitative issues with road collision data are more pronounced in pedestrian safety studies. Arguments against reactive and collision-based road safety analysis are more acute in what relates to the study of pedestrians.
4. Traffic conflict techniques, if conducted in an objective and a resource-efficient fashion, can be a useful complement or substitute to collision-based road safety analysis. The proof of validity of this technique will become more obtainable if video-analysis methods are developed and adopted.

Previous work has been performed to develop a video analysis system that can automatically detect, classify, and track road users and interpret their movement [20]. This paper presents further development steps of this system in order to analyze pedestrian safety in the context of scramble phase intersections. Following are the study objectives:

1. Extract real-world positions of road users that appear in video sequences,
2. Detect, track, and identify the type of road users,

---

[1] The severity of a traffic interaction can be described by how imminent was a collision possible between the conflicting road users if their movement remained unchanged.

[2] A well-recognized conceptual definition 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 remained unchanged” [38]. The interpretation of words such as “approach” and “unchanged” into a set of rules is context dependent.

3. Calculate objective conflict indicators,
4. Evaluate the ability of the system to detect and estimate the severity of dangerous pedestrian-vehicle interactions. The evaluation is compared to observer-based detection and rating of traffic conflicts.

This research is another step in a direction that is, to the best of the authors' knowledge, unique in the field of road safety and pedestrian studies.

Literature Review

Evaluation of Scramble Phase Intersection

Pedestrian scramble is a type of traffic signal phasing that gives pedestrians an exclusive phase in which all approaches are stopped. In a scramble phase, pedestrians are permitted to walk between any two points on the intersection perimeter. Often, the traditional crosswalk marking is indented to indicate that walking during the scramble phase is permitted at any point within the intersection. The purpose of this traffic control scheme is to limit the interaction with and exposure to conflicting motorized traffic. Due to the novelty of this traffic phasing scheme, to transportation engineers as well as road users, proper evaluation of this engineering treatment is often required.

Existing evaluation work contains mixed, and sometimes paradoxical, results. The objective of the scramble phase is to reduce the potential for conflict between pedestrian traffic and turning movement. Therefore, the risk of collision will decrease. The provision of an exclusive pedestrian phase comes with the price of prolonged waiting time. Pedestrians are not permitted to cross with parallel motorized traffic and are required to wait. Since pedestrian violation rate may increase with waiting time, the scramble phase can exacerbate existing problems of unpermitted crossing, thus increasing the potential for conflicts [21]. Vaziri [22] reported a 66% reduction in crashes after the introduction of a scramble phase. The author explained this reduction by the high pedestrian as well as turning volumes at the concerned intersection. Zeeger et al. [23] indicated that scramble phase is most advantageous at intersections with high pedestrian and turning vehicle volumes. Garder [24] reported a reduction in pedestrian-vehicle conflicts after a scramble phase was introduced in three intersections in Sweden. The reduction was conditional on high pedestrian compliance rate. Kim and Teng [25] developed simulation models of scramble phase and found that they reduced pedestrian-vehicle conflicts and increased vehicular delays. Bechtel et al. [26] found a significant reduction in the rate of pedestrian-vehicle conflicts and an increase in the number of pedestrian violations. Similar findings regarding the effect of compliance rate were reported in a recent study in the City of Calgary [27].

The literature on scramble phase evaluation is relatively limited. Two open questions remain: what is the change in pedestrian-vehicle severity, temporal, and spatial distribution after implementing the scramble phase? and what is the long-run safety implication of increased violations? Finding a conclusive answer can be aided by automated safety analysis.

Video-based Safety Analysis

The objectives of a video analysis system are to detect, classify, and track road users observed in video sequences in an automated fashion. The components of the video analysis system are 1) video preprocessing; 2) feature processing; 3) grouping; 4) high-level object processing; and 5) information extraction, as presented in Figure 1. Figure 2 illustrates the tracking steps and Figure 3 illustrates the process of camera calibration. A subsequent stage is the analysis of road user tracks to measure their spatial and temporal proximity. In this study, the safety analysis was conducted by calculating four conflict indicators: Time to Collision (TTC), Post-Encroachment Time (PET) and Gap Time (GT).

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.” [28]. An accurate estimation of TTC using field observations is prohibitive, requiring considerable field measurement of road user positions, speed and direction of movement. PET 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 [29]. GT is a variant of PET that is calculated at each instant by extrapolating the movements of the interacting road users in space and time [30]. 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 and PET.

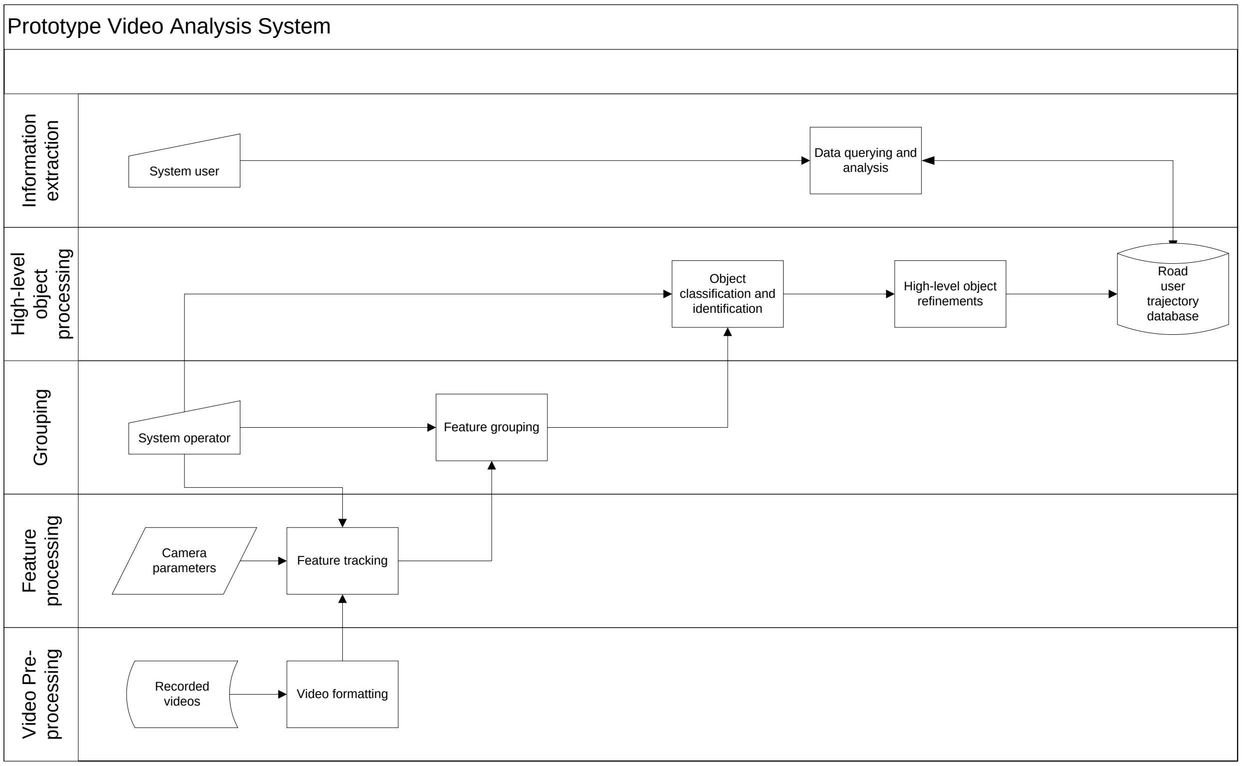


Figure 1: Layout of the video analysis system. The figure shows the five main layers the make up the system. Depicted also is the data flow among system modules from low-level video data to a position database of detected, tracked, and classified road users **[32]**.

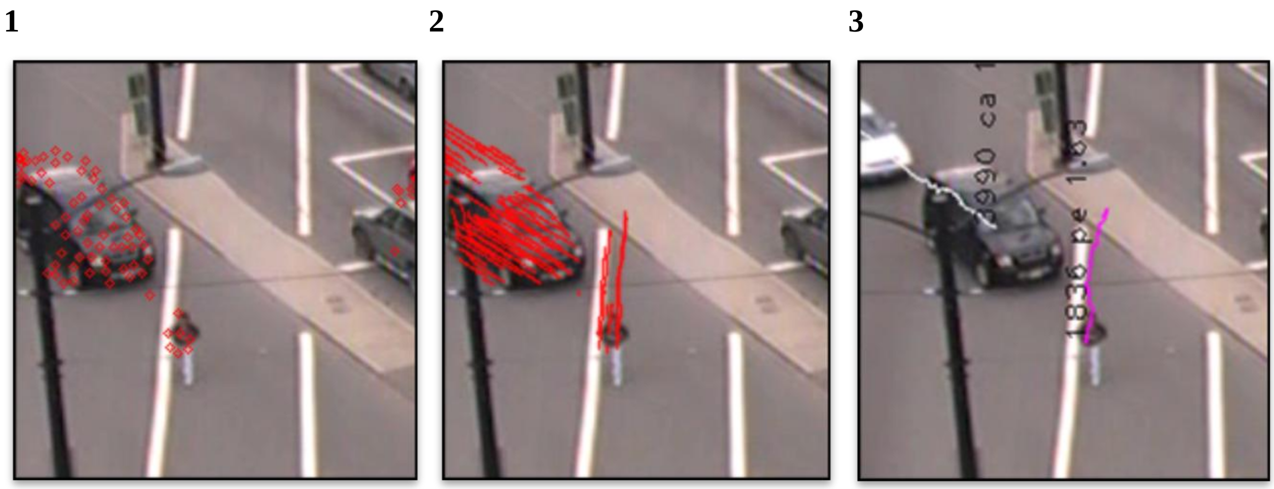


Figure 2: Features are first detected on road users as shown in frame 1. Features are tracked in consecutive frames as shown in frame 2. Features that exhibit similarity in movement are grouped and a representative trajectory is adopted as shown in frame 3. The displayed numbers are road user classification ("*ca*" denotes a car and "*pe*" denotes a pedestrian) along with instantaneous speed (m/s).

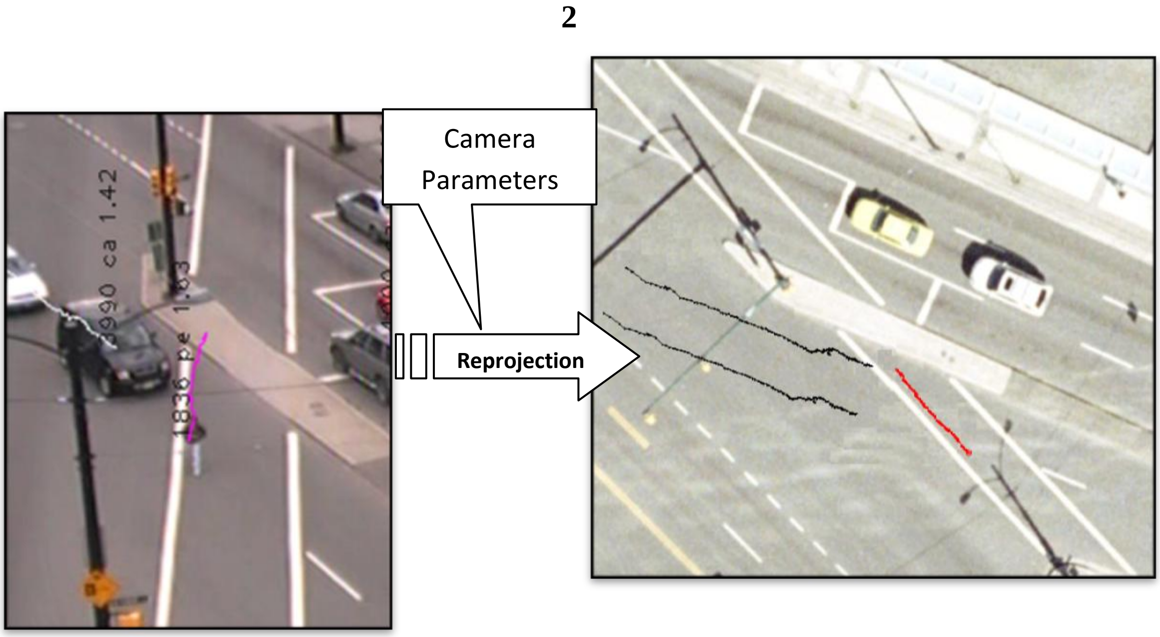


Figure 3: detected road users in subfigure 1 are reprojected (mapped or transformed) into world coordinates and overlaid on the orthophoto of the scene as depicted in subfigure 2.

Methodology

The aim of this study is to demonstrate a study of pedestrian-vehicle conflicts in the context of scramble phase intersection. The research methodology is composed of four main steps: 1) Data collection, 2) camera calibration, 3) Extraction of road user tracks, 4) calculation of conflict indicators, and 5) Validation. The methodology is applied to a video sequence collected after the introduction of the scramble phase in the Webster and 8th intersection in Oakland, California.

Site Description and Data Collection

Video sequence analyzed in this study was collected in 2002 at the 8th and Webster intersection in a Downtown intersection in Chinatown in Oakland, California. The recordings were obtained from a vantage point at the North East side of the intersection on 09/13/02. The sequence length is half an hour after the implementation of the scramble phase. The intersection approaches carry one-way traffic movement along Webster St. (North-South and right-turn movement from North to West) and 8th St. (East-West and left-turn movement from East to South). The intersection serves relatively large volume of vehicular traffic of approximately 4,000 vehicles per hour [33]. The pedestrian volume is also relatively high with approximately 3,000 pedestrians crossing per hour at peak time. Turning movements represent approximately one third of the traffic volume. A combined time of “Walk” and flashing “Don’t Walk” time before the scramble phase was 15 s for 8th and 19 s for Webster. After the scramble phase, the combined “Walk” and flashing “Don’t Walk” is 28 s. The pedestrian green ratio, the ratio between “Walk” time and cycle length decreased after the scramble phase. This reflects a more general implication of scramble phase which increased the waiting time at the expense of reduced pedestrian-vehicle interaction and reduced walking distance for pedestrians (who can cross diagonally).

Camera Calibration

A main shortcoming of using video observation is the reduction in dimensionality by projection on surface plane. The following quote explain this problem succinctly:

“*The use of video-analysis also limits the quality and scope of a safety study, where safety critical events can be difficult to detect in two-dimensional imagery, and subject to problems related to the relative positioning of the camera and the coverage this provides*” [34]. By conducting camera calibration, we attempt to address these shortcomings.

The canonical pinhole camera model is adopted to represent the perspective projection of real-world objects on the image plane. The objective of camera calibration is to find a mapping $\mathbf{T}$ from image coordinates to real-world coordinates. The matrix $\mathbf{T}$ can be decomposed into two matrices such that: $\mathbf{T} = \mathbf{M} \times \mathbf{N}$, where matrix $\mathbf{N}_{4x4}$ maps from world coordinates to camera coordinates, and matrix $\mathbf{M}_{3x4}$ maps from camera coordinates to pixel coordinates. Knowledge of extrinsic camera parameters, comprising 3 rotation angles and a translation vector, is sufficient for generating $\mathbf{N}$. Knowledge of the intrinsic camera parameters (focal length and radial lens distortion) are sufficient to generates matrix $\mathbf{M}$.

where f_x and f_y are respectively referred to as the horizontal and vertical focal length in pixels (assumed equal in this study), θ is the angle between the horizontal and vertical axes of the image plane, and (u_o, v_o) are the coordinates of the principal point.

The estimation of camera parameters can is cast as an optimization problem in which we search for the set of parameters that minimize the difference between observed features and their reprojection from the image space. A proper cost function need to satisfy the following conditions:

1. Be able to uniformly represent error terms from different geometric primitives, i.e. consistent weights and units.
2. Be perspective invariant, i.e. not sensitive to image resolution or camera-object distance.

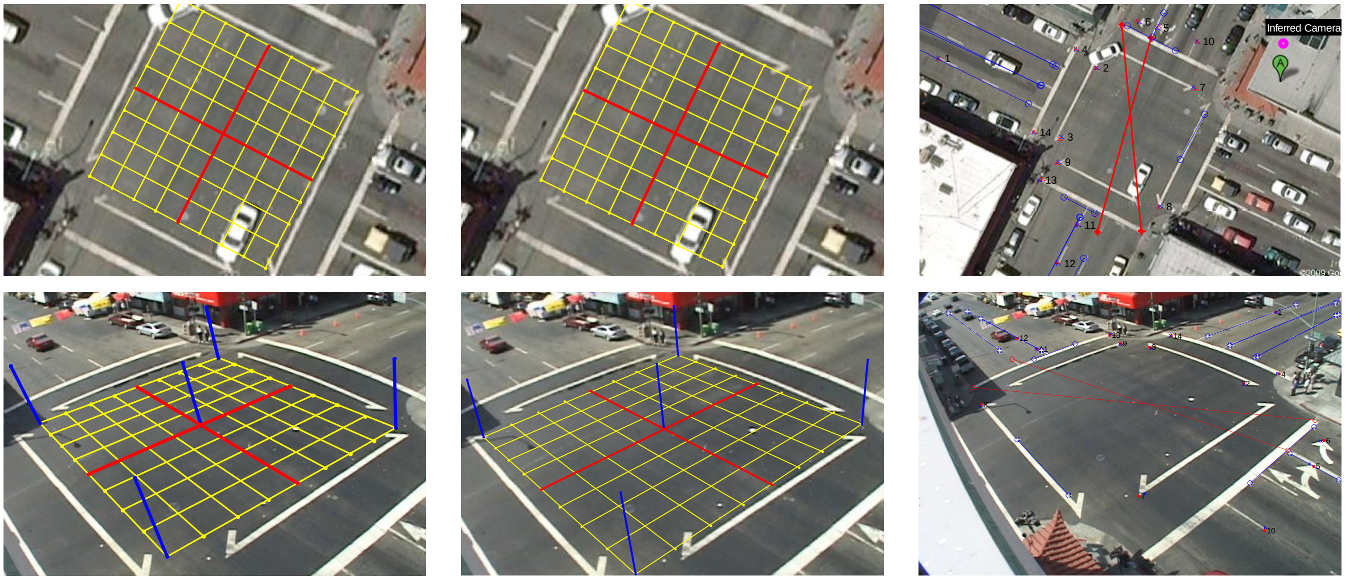
Figure 4 (Subfigure 3) depicts the features used in the proposed cost function. The following cost function is used in this study:

$$f(\mathbf{X}) = \sum_{i\in C, j\in D, k\in A} \|\Delta\mathbf{P}_i\|_2^{\ 2} + \left(\Delta s_j\right)^2 + \left(\bar{l}\tan\Delta a_k\right)^2 \qquad (1)$$

where $\mathbf{X}$ is a vector of camera parameters, $C, D, \& \ A$ are respectively the sets of calibration point-difference, distances, and angular constraints, $\|\Delta\mathbf{P}_i\|_2$ is the real-world distance between observed and back-projected calibration point *i*, and Δs_j is the difference between observed and projected distances *j*. Back-projection to the world space is performed efficiently using the homography matrix that corresponds to $\mathbf{X}$. A least square estimation of the homography matrix defined by the set of equations that characterize the projection of four points, preferably selected from C, using $\mathbf{X}$. Angular constraints are defined in terms of the angle between the back-projection of some pairs of user-defined line segments. $\bar{\mathrm{l}}$ is the average length of the back-projected line segments. $\Delta \mathrm{a}_\mathrm{k}$ is the difference between observed and calculated acute angle *k* between the back-projected line segments. The previous cost function describes linear discrepancies between observed and back-projected geometric primitives, all expressed in real-world unit distance.

The minimization of the cost function in Equation 3 over the camera parameters is performed using various optimization algorithms. Preliminary results show that the Nelder-Mead (NM) simplex algorithm is preferable to strong convergence characteristics. The initial guess of the camera parameters was obtained using an estimate of the camera position in an orthographic map of the monitored traffic intersections, of the camera height, and of the location of the back-projection of the principal point on the road surface.

Figure 4 shows the results of the camera calibration and demonstrates the success of this approach for camera calibration proved very successful. The root mean square error in estimating positions and distances was approximately 20 cm. This is estimation accuracy is adequate and exceeds the accuracy of the extracted road user tracks. It is noteworthy that in the video sequence analyzed in this study research, the only information available was the video sequence recorded for the traffic intersection. Using Google Maps, an orthographic image of the intersection was obtained at a reasonable resolution. Using matched features and additional linear and angular constraints, the camera parameters were estimated.



Subfig. 1 calibration results before including radial camera distortion.

Subfig. 2 calibration results after including radial camera distortion.

Subfig. 3 calibration features as shown in world (upper) and image (lower) spaces.

Figure 4: These figures demonstrate the ability of the proposed approach to obtain reliable camera parameters. The sample frames are collected from the video at instants in which there was little presence of road users in the intersection.

Extraction of Road User Tracks

The developed video analysis system is capable of automatically recovering road user tracks from the collected video sequence. The tracking of features is done through the well known Kanade-Lucas-Tomasi feature tracker. Since a moving object can have multiple features, the features have to be grouped, using cues like spatial proximity and common motion. In this process, a graph is constructed over time such that each node represents a detected feature and an edge represents an assignment to a common moving road user. The tracking process depends on a set of manually set parameters. A detailed description of these parameters can be found in [32]. The two most important tracking parameters for the feature grouping process are: Segmentation Distance ($D_{segmentation}$), the maximum difference between the minimum and maximum distance between two features in the same group, and the Connection Distance ($D_{connection}$), the maximum initial distance between two features in the same group. The relationship between the two parameters is complex. In general, increasing $D_{segmentation}$ will increase the chance of over-grouping, i.e. tracking multiple road users as one. Reducing $D_{segmentation}$ may lead to the opposite, over-segmentation, i.e. the tracking one road user as many. Both issues may compromise the reliability of the system to detection pedestrian-vehicle conflicts. $D_{connection}$ controls the spacing between two connected features. A large value may result in over-grouping as well. This is a typical case when a group of pedestrians are grouped as a one moving object.

In previous work, these parameters were tuned by trial and error [31] [20]. Although considerable experience has been gained in the optimal selection of the set of tracking parameters, the manual approach is time consuming, non-systematic, and holds therefore no guarantee of optimality. To address these shortcomings, the set of tracking parameters was obtained in a systematic way by casting it as a minimization problem. For the purpose of conflict detection and severity rating, the main issue is the failure to detect and track road users. Parameters are therefore selected to minimize the number of missed road users. Following is a brief description of the parameter selection process:

1. Create a ground-truth data set. By observing a test sequence, an observer is asked to label road users in each frame, i.e. draw a bounding box around vehicles and bounding ellipses around pedestrians. Figure 5 shows a sample of labelled frames.
2. Evaluate the performance of the system on the same part of the video sequence for which ground truth was obtained. This is done by analyzing the correspondence between the ground truth and the road users tracked by the system. A road user detected by the system and in the ground truth are assigned if they are similar enough, i.e. satisfying some temporal and spatial similarity conditions. A specific method was developed for this purpose and is described in more details in [37].
3. The number of missed detections is the number of unassigned ground truth labelled objects, i.e. road users that are not tracked. This is one of the most undesirable outcomes since if the tracking system errs on the side of missing road users, it will not be possible to analyze the interactions in which these road users may have been involved. While the safety analysis system is intended ultimately to function in an automated way, it could be utilized interactively to assist users. For that purpose, an observer reviews the events reported by the system and identifies traffic conflicts within this subsequence of the entire video sequence. This can reduce the time required for in-office traffic conflict analysis. Erring on the side of missing traffic conflicts defeats can greatly erode the advantage of reducing video observation time. For this purpose, the cost term of the evaluation process is the number of missed detections.

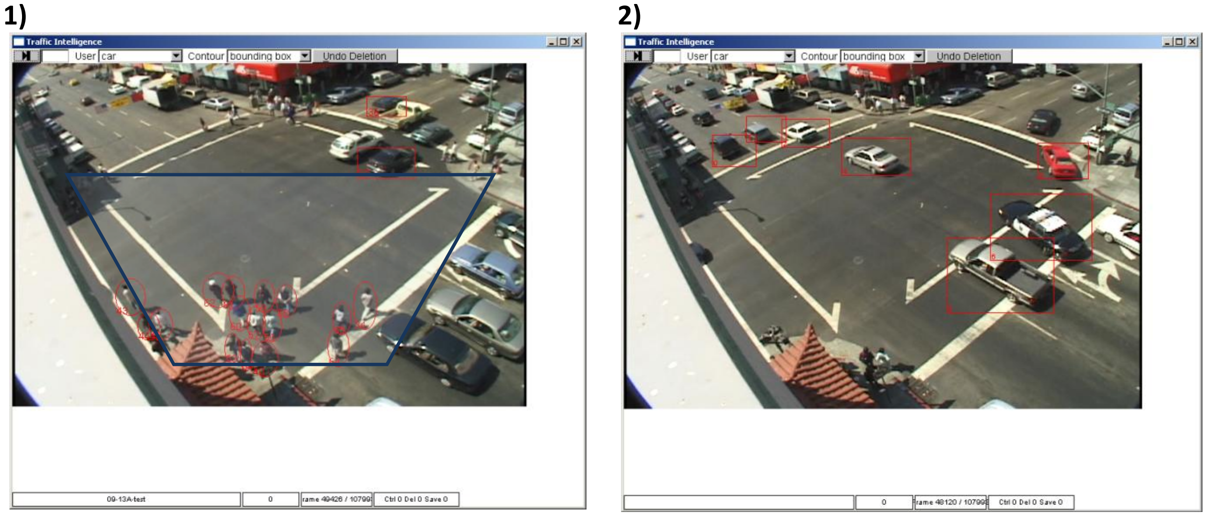


Figure 5: These figures demonstrate the ability of the proposed approach to obtain reliable camera parameters. The sample frames are collected from the video at instants in which there was little presence of road users in the intersection. Pedestrians were labelled only within the region in subfigure 1. This is called in later parts of the research *“region of detection”* for which reliable tracks are obtained.

Figure 6 shows the results of the performance evaluation process. The total number of labelled road users in this subsequence was 87 (pedestrians and vehicles). The pair of connection and segmentation parameters used in this analysis is (0.85m, 0.4m). The rate of false detection at the this parameter selection is 92% and rate of missed detection is 70%. The problem can equivalently be cast as an optimization problem in which the search for the optimal parameters is performed automatically. The simple approach employed in this study yielded similar results to those found with manual tuning (0.9m, 0.35m). However the advantage is that the automated evaluation approach is more systematic and much more time-efficient. Figure 7 shows the superimposition of a 30% sample of the tracked road users.

The calculation of the conflict indicators is conducted in a fashion similar to the algorithm presented in previous work [20]. There is however difference in calculating PET and GT since during a scramble phase, the movement of pedestrians is not restricted within a crosswalk. In that, the instant at which a vehicle "encroaches" into the right-of-way of a pedestrian is not strictly its entrance to the boundaries of a crosswalk. To follow a similar approach, a virtual crosswalk is defined around each pedestrian. This is a heuristic that quantitatively defines a pedestrian's right of way. The virtual crosswalk is defined with a width of 3m and extends before the pedestrian object start point and beyond the end point by 2 meters. Figure 8 illustrates the virtual crosswalk. In order to demonstrate the safety interpretation based on conflict indicators, a sample interaction is presented in Figure 9. The upper figures show snapshots of the two road users approaching each others at the end of and after the scramble phase. The pedestrian failed to clear the intersection before the end of the scramble phase. The motorist appears do drive unaware of the pedestrian and at some point, and therefore could collide with her if the vehicle's movement remained unchanged. Figure 8 shows also the vehicle speed measured at every frame (1/30 s), TTC, and GT. PET does not evolve with time and was evaluated based on observed road user trajectories.

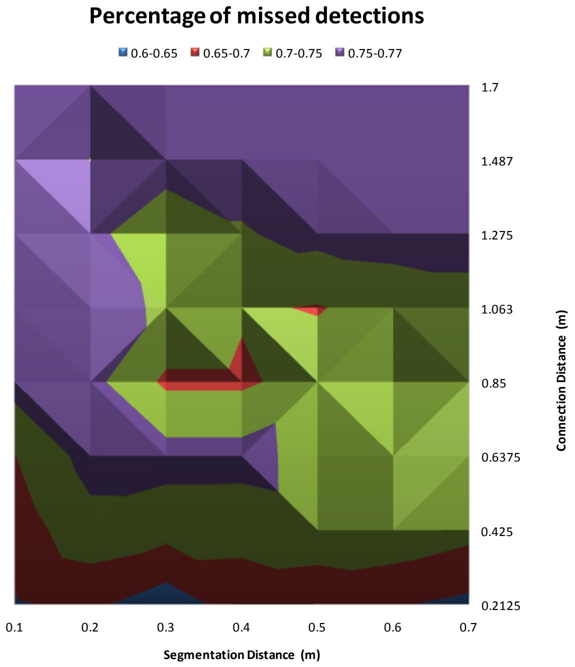


Figure 6: The areas in red represent desirable sets of segmentation and connection distances. The area around (0.35, 0.85) is selected in this study since it represents a balanced tracking accuracy between pedestrians and vehicles.

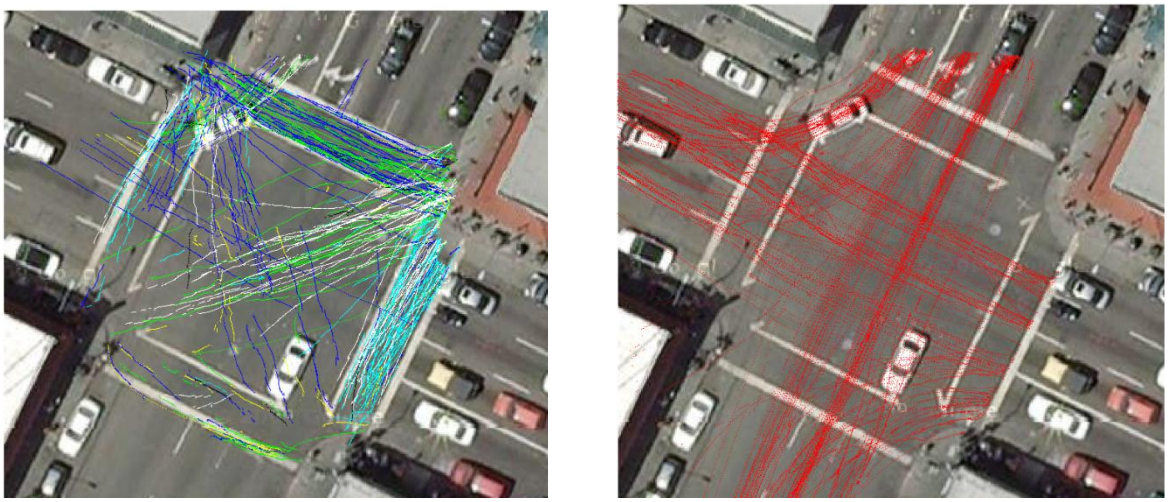


Figure 7: Superimposition of road user tracks. Tracks in red are those of motorists. Pedestrian tracks are displayed in other colours depending on a 12-category K-means clustering. The clustering variable are the set of average pedestrian directions at the first, middle, and last 20% of the road user's lifetime.

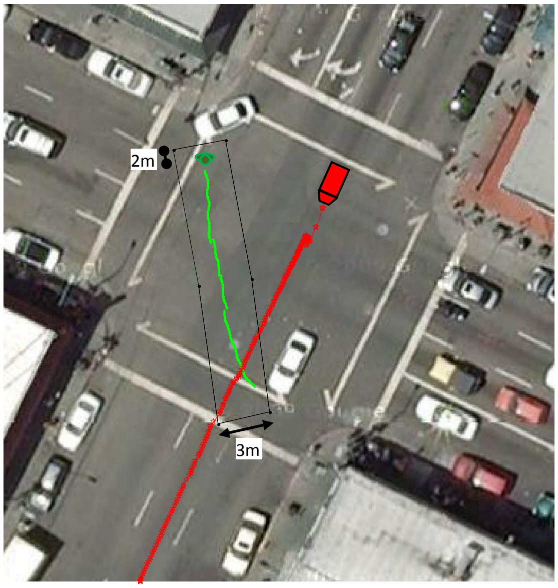


Figure 8: Pedestrian track (in green) and conflicting vehicle track (in red). Marked in black are the six points defining the boundary of a virtual crosswalk.

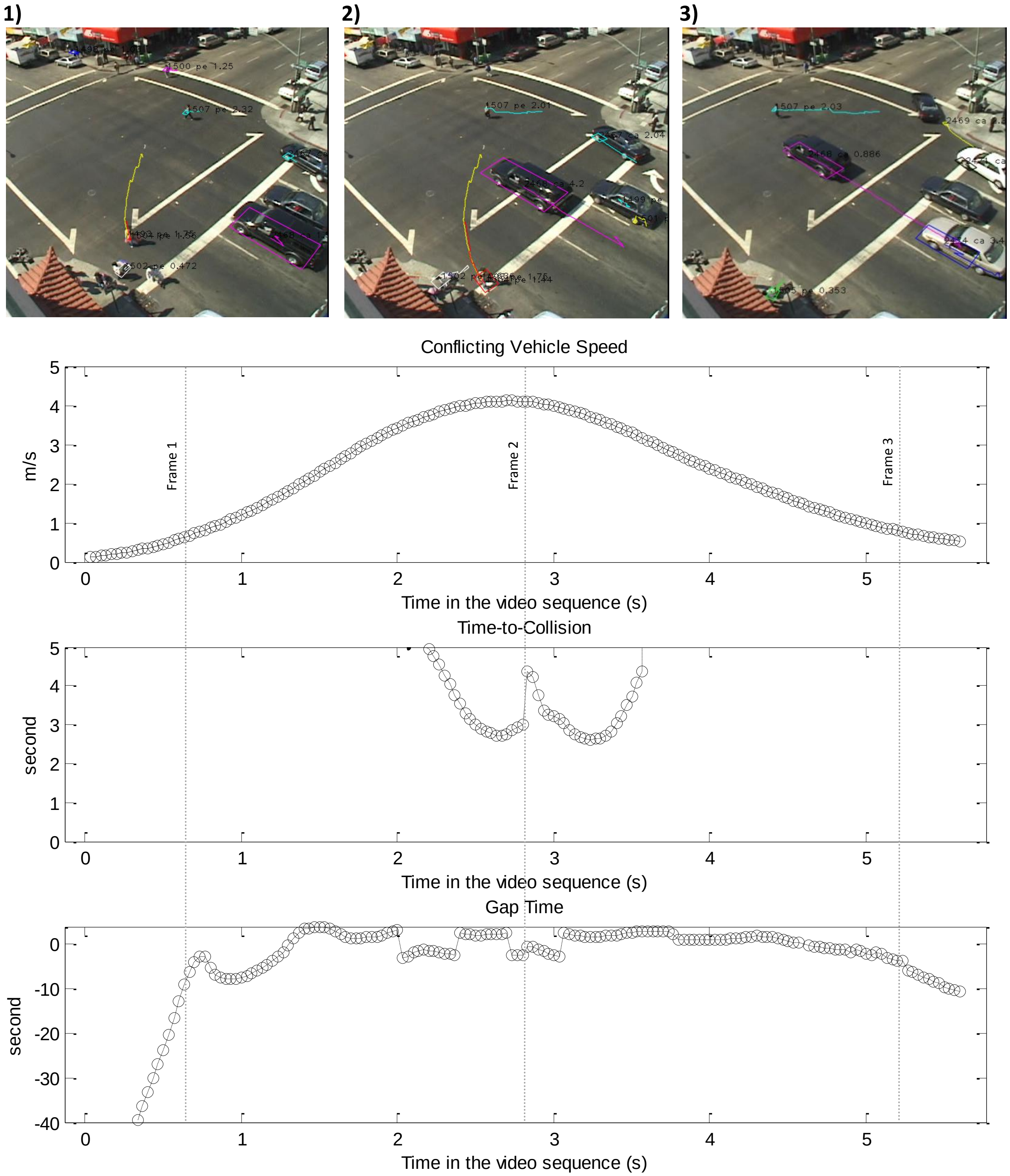


Figure 9: A sample pedestrian-vehicle interaction. The pedestrian entered the intersection during the "Don't Walk" part of the scramble phase and failed to clear the intersection before the end of the phase. A clear conflict with a southbound vehicle is created. For this event, the minimum recorded TTC is 2.6s. The minimum GT is 0.015s and PET is 3.1s.

Results

The camera is set the designated location in Figure 4 (Subfigure 3 upper right corner). This enables good quality capture of the road user movement in the closer quarter of the intersection. In total, 2622 pedestrian objects and 2620 vehicle objects were detected and tracked. The heuristic used for conflict detection is as follows:

$$Detection = \begin{cases} 1 & if\ (TTC_{min} < 3) OR\ (|PET| < 3) OR (|GT_{min}| < 3) \\ 0 & otherwise \end{cases}$$

where $Detection$ takes a value of 1 (true) if the right-hand condition is satisfied. Pedestrians using southern and western crosswalks were not reliably tracked. All conflicts that took place within this region (conflicts types 1 and 2) were detected. The rate of false alarm (the ratio of interactions detected as conflicts but not labelled as such) is 12.5%. The majority of these false alarms are due to misclassification of pedestrians moving during the scramble phase as vehicles. Also, some slow-moving vehicles were misclassified as pedestrians, and therefore triggered a lot of conflicts. Table 1 shows a summary of events labelled as conflicts by a human observer. The criteria in the FHWA observer guide were used to detect conflicts [37].

Table 1: Summary of automated safety analysis

Conflict	Frame	TTC_{min}	PET	GT_{min}	Comments
1	20634	-	1.26	-	Significant part of the pedestrian track is outside the region of detection.
2	34108	2.6	2.6	0.06	Complete capture of the interacting road user tracks.
3	32490	-	-	-	Outside region of detection
4	34950	-	-	-	Outside region of detection

Conclusions

This study presented a novel video analysis system for the detection and severity rating of traffic conflicts. The data analyzed in this study is a subsequence of a larger set of video observations of the intersection at 8th and Webster in Oakland, California. Further analysis is being conducted in the future. Based on the results presented in his paper, the detection of traffic conflicts using video analysis is clearly feasible. This approach overcomes traditional shortcomings of collision-based road safety analysis and also limitations to field observations of traffic conflicts. Three main continuations of this research are being pursued:

1. The improvement to the traditional definition of objective conflict indicators.
2. Improved road user classification. A major source of false alarms is the relatively high rate of false classification (pedestrians misclassified as vehicles and the contrary case). The heuristic used in this study was based only on the maximum observed speed. When this value exceeds a threshold of 3 m/s, the object was classified as a vehicle.
3. Conducting the analysis on different video sets. This video data was not collected for the purpose of automated analysis. The camera setting could support better analysis if set at a higher location. The quality of the image at the further parts of the intersection was not adequate to enable reliable tracking.

References

[1] Greene-Roesel, R., Diogenes, M.C. and Ragland, D., *Estimating Pedestrian Accident Exposure: Protocol Report.* Institute of Transportation Studies. UC Berkeley Traffic Safety Center. 2007. p. 14.

[2] Zein, T. Sayed and S., "Traffic Conflict Standards for Intersections." s.l. : Transportation Planning and Technology, 1999, Vol. 22, pp. 309-323.

[3] , *A Policy on Geometric Design of Highways and Streets.* s.l. : American Association of State Highway and Transportation Officials (AASHTO), 2001.

[4] Odero, W., Garner, P. and A., A. Zwi., "Road traffic injuries in developing countries: A comprehensive review of epidemiological studies." Trop Med int Health, 1997, Vol. 2, pp. 445-60.

[5] An., *British Columbia Traffic Collision Statistics.* Insurance Corporation of British Columbia. 2006. Annual Report.

[6] Milam, Ron and Mitchell, Chris., "Conventional Level of Service Analysis, Thresholds, and Policies Get a Failing Grade." 2008.

[7] , *Pedestrian Research Circular 2005.* Washington, DC : Transportation Research Board, 2005. p. 6. Developing a Method for Conducting Pedestrian Counts.

[8] Pulugurtha, Srinivas Subrahmanyam and Repaka, Sudha R., "Assessment of Models to Measure Pedestrian Activity at Signalized Intersections." s.l. : Transportation Research Board, 2008.

[9] Shankar, Venkataraman N., et al., "Modeling crashes involving pedestrians and motorized traffic." Safety Science, 2003, Vol. 41, pp. 627-640.

[10] Greene-Roesel, R., et al., *Effectiveness of a Commercially Available Automated Pedestrian Counting Device in Urban Environments.* Institute of Transportation Studies. UC Berkeley Traffic Safety Center. 2008.

[11] Mathers, C. and Loncar, D., *Updated projections of global mortality and burden of disease, 2002-2030: data sources, methods and results.* World Health Organization. 2005.

[12] An., *The Global Burden of Disease: 2004 Update.* World Health Organization. 2004.

[13] Ross, A., "Roads & Highways Topic Note RH-2." *Road Accidents: A Global Problem Requiring Urgent Action.* [Online] June 1999. http://www.worldbank.org/html/fpd/transport/topnotes/rh-2.htm.

[14] Vodden, K., et al., *Analysis and Estimation of the Social Cost of Motor Vehicle Collisions in Ontario.* Transport Canada. 2007. Final Report.

[15] An., *The Cost of Urban Traffic Congestion in Canada.* Transport Canada, Environmental Affairs. 2006.

[16] Chin, H.C. and Quek, S.T., "Measurement of traffic conflicts." Safety Science , s.l. : Elsevier, 1997, Vol. 26, pp. 169–185.

[17] Perkins, R. and Harris, J.I., "Traffic conflict characteristics: accident potential at intersections." Highway Research Record, 1968, Vol. 225, pp. 35–43.

[18] Svensson, Å. and Hydén, C., "Estimating the severity of safety related behaviour." Accident Analysis and Prevention, 2006, Vol. 38, pp. 379-385.

[19] Glauz, W. and Migletz, D., "The Traffic Conflict Technique of the United States of America." International Calibration Study of Traffic Conflicts, 1984, Vol. NATO ASI Series Vol. F5, pp. 41-58.

[20] Hayward, J., "Near-miss determination through use of a scale of danger." Highway Research Record, 1968, Vol. 384, pp. 24-34.

[21] Ismail, K., et al., "Automated Analysis of Pedestrian-vehicle Conflicts using Video Data." Washington, DC., January 2009 : In Transportation Research Board Annual Meeting Compendium of Papers, Accepted for publication in the Transportation Research Record, 2009. 09-3743.

[22] Abrams, C. and Smith, S. A., "Selection of Pedestrian Signal Phasing." Transportation Research Record: Journal of the Transportation Research Board, 1977, Vol. 692, pp. 1-6.

[23] Vaziri, B., "Exclusive Pedestrian Phase for the Business District Signals in Beverly Hills: 10 Years Later." 1989. Institute of Transportation Engineers, Washington,D.C.,.

[24] Zegeer, C., Opiela, K. and Cynecki, M., *Pedestrian Signalization Alternatives. Final Report.* Goodell-Grivas Incorporated and Federal Highway Administration. 1985. Final Report.

[25] Gårder, P., "Pedestrian Safety at Traffic Signals: A Study Carried Out with the Help of a Traffic Conflicts Technique21." Accident Analysis and Prevention, 1989, Vol. 21, pp. 435-444.

[26] Kim, K. and Teng, H., "Time Separation Control between Pedestrians and Turning Vehicles at Intersections." 2004. Transportation Research Board.

[27] Bechtel, A., MacLeod, K. and Ragland, D., *Oakland Chinatown Pedestrian Scramble: An Evaluation.* UC Berkeley Traffic Safety Center. 2003. Final Report.

[28] Acharjee, S., Kattan, L. and Tay, R>., "Pilot Study on Pedestrian Scramble Operations in Calgary." 2009.

[29] Hayward, J. Ch., "Near-miss determination through use of a scale of danger." Highway Research Record, 1968, Vol. 384, pp. 24–34.

[30] Cooper, J., "Experience with traffic conflicts in Canada with emphasis on post encroachment time techniques." 1984. Vol. vol. F5, pp. 75–96.

[31] 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.

[32] Ismail, K., Sayed, T. and Saunier, N., "Automated Collection Of Pedestrian Data Using Computer Vision Techniques." Washington, DC : Transportation Research Board Annual Meeting , 2009.

[33] Bechtel, A., MacLeod, K. and Ragland, D., "Pedestrian Scramble Signal in Chinatown Neighborhood of Oakland, California: an Evaluation." Transportation Research Record: Journal of the Transportation Research Board, 2004, Vol. 1878, pp. 19-26.

[34] 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.

[35] Saunier, N. and Sayed, T., "A feature-based tracking algorithm for vehicles in intersections." s.l. : IEEE Computer Society, 2006. Computer and Robot Vision, 2006. The 3rd Canadian Conference on. pp. 59- 59.

[36] Saunier, N., Sayed, T. and Ismail, K., "An Object Assignment Algorithm for Tracking Performance Evaluation." 2009. Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2009).

[37] 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.

[38] Amundsen, F. and Hydén, C., 1977. Lund Institute of Technology.

[39] Masoud, O. and Papanikolopoulos, N. P., "Using geometric primitives to calibrate traffic scenes." Transportation Research Part C, s.l. : Elsevier, December 2007, Issue 6, Vol. 15, pp. 361-379.

[40] J.Weng, P. Cohen, and M. Herniou., "Camera Calibration with Distortion Models and Accuracy Evaluation." IEEE Trans. Pat. An. Mach. Int., 1992, Issue 10, Vol. 14, pp. 965-980.