

# Road Traffic Conflict Analysis from Geo-referenced Stereo Sequences

Sebastiano Battiato<sup>1</sup>, Stefano Cafiso<sup>2</sup>, Alessandro Di Graziano<sup>2</sup>,  
Giovanni M. Farinella<sup>1</sup>, and Oliver Giudice<sup>1</sup>

<sup>1</sup> Image Processing Laboratory, Dipartimento di Matematica e Informatica

<sup>2</sup> Dipartimento di Ingegneria Civile e Ambientale  
University of Catania, Italy

**Abstract.** In this paper an imaging system for road traffic conflict analysis is proposed. The system exploits geo-referenced stereo sequences and tracking procedure to compute traffic conflict measures which can be analysed by experts. Using the potentiality of the traffic conflict technique as a surrogate safety measure could constitute an effective tool in understanding how the driver interacts and adapts its behaviour with respect to the vehicle, the road characteristics, the traffic control devices and environment. Experiments performed on real data acquired in urban environment confirm the effectiveness of the system which makes simple and fast for the experts the understanding of the driver behaviour.

**Keywords:** Traffic conflict analysis, Stereo system, Tracking.

## 1 Introduction

With the increasing number of vehicles on the roads, accidents have become more and more frequent today. There are different causes of road accidents: distracted driving, speeding, unsafe lane changes, etc. Many people die as a result of road accidents, and many more have unwanted consequences, including serious injuries. The huge social, human and also economic costs induced by road accidents make the topic of road safety really important in nowadays society.

The objective of the traffic conflict techniques is the evaluation of traffic safety and the prediction of road accidents. By automatically understanding and predicting the risk of a road accident some actions can be taken in time by systems (e.g., Driver alerting) in order to avoid the conflict. Different studies have been done in the traffic safety context to build appropriate tools and measures for traffic conflict [1,2]. The major problem when using the Traffic Conflict Technique (TCT) is the complexity in data acquisition and processing: everything that happens in front of the vehicle needs to be taken into account in terms of time and space variables.

In this paper we propose a system which is useful for both, data acquisition and processing in the context of traffic conflict monitoring. The proposed system makes use of geo-referenced stereoscopic vision for the acquisition of data which are then analysed with a semi-automatic pipeline acting on the acquired stereoscopic video in order to obtain the traffic conflict risk measures.

The paper is organized as follows: we introduce the reader to the Traffic Conflict Technique fundamentals in Section 2. The hardware used in proposed system is summarized in Section 3, whereas the software components employed for data processing and traffic conflict analysis are described in Section 4. Experiments and results are reported in Section 5. Finally, conclusion and hints for future works are given in Section 6.

## 2 The Traffic Conflicts Technique

The Traffic Conflicts Techniques (TCT) are based on the “Heinrich Triangle” theory [3] founded on the relationship that “no-injury accidents” precede “minor injuries” (i.e., events closer to the base of the triangle precede events nearer the top). Application of this theory assumes that the appropriate traffic conflict factors can be defined as measures of near-crash events. A traffic conflict could be defined as “an observable situation in which two or more road users approach each other in space and time to such an extent that there is risk of collision if their movements remain unchanged” [4]. Traffic conflict measures, such as “time to collision” (TTC), address the first condition of surrogate measures, namely the common factors that are shared with safety highlighting advantages and limits of the TCT [5]. The TCT are based on the measurement of both spatial and temporal variables which describe the interactions between two road users involved in a critical event for safety. The traffic conflict measures considered in this paper are connected to the severity of the conflict that is a combination of accident probability and severity of injuries. The Time To Collision (TTC) is the indicator used by TCT and it is calculated for the vehicles ( $TTC_v$ ) and an obstacles ( $TTC_o$ ). The TTC is defined as follows:

$$TTC = \frac{D}{V} . \quad (1)$$

where  $D$  is the distance between the vehicle and the conflict point, and  $V$  is the velocity of a generic obstacle (e.g., another vehicle, a pedestrian, etc.).

In case the obstacle is a pedestrian crossing the street, the TTC measures at time  $i$  for the two actors could be defined as follows [6]:

$$TTC_v(i) = \frac{D_v(z_i)}{V_v(i)} , \quad (2)$$

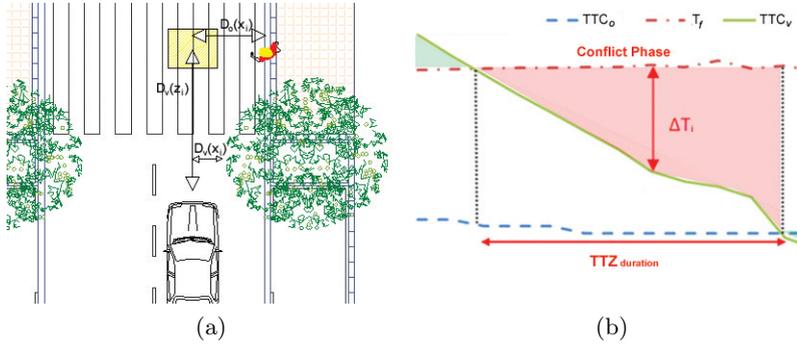
$$TTC_o(i) = \frac{D_v(x_i) - D_o(x_i)}{V_o(i)} . \quad (3)$$

where, by considering a 3D reference system centered on the vehicle,  $D_v(z_i)$  represents the distance between the vehicle and the conflict point along the  $Z$  axis, and  $D_v(x_i) - D_o(x_i)$  represents the distance between the vehicle and the obstacle along the  $X$  axis (see Fig. 1a).

Let be  $T_f$  the time required for the vehicle to stop at time  $i$  defined as:

$$T_f(i) = T_r + \frac{V_v(i)}{d} . \quad (4)$$

where  $T_r$  is the reaction time,  $V_v(i)$  is the vehicle velocity at time  $i$ , and  $d$  is the deceleration during braking.



**Fig. 1.** (a) Distances involved in a TTC computation. (b) Temporal trend of the quantities involved in the TTC. The area where  $TTC_o(i) < TTC_v(i) < T_f(i)$  is related to the conflict.

It is possible to consider the following cases, which determine the areas as in Fig. 1b:

- i)  $TTC_v(i) > T_f(i)$ : vehicle may stop before conflict area,
- ii)  $TTC_o(i) > TTC_v(i)$ : the vehicle passes the area of conflict before the pedestrian reaches it,
- iii)  $TTC_o(i) < TTC_v(i) < T_f(i)$ : there is the conflict and it is possible to evaluate the  $TTZ_{duration}$  that is the duration of the conflict.

Given the  $TTZ_{duration}$  it is possible to compute the severity of the conflict as Risk Impact (RI) at time  $i$  according to the following equation:

$$RI(i) = \sum_{i \in TTZ_{duration}} V_v(i)^2 \times \Delta T(i). \tag{5}$$

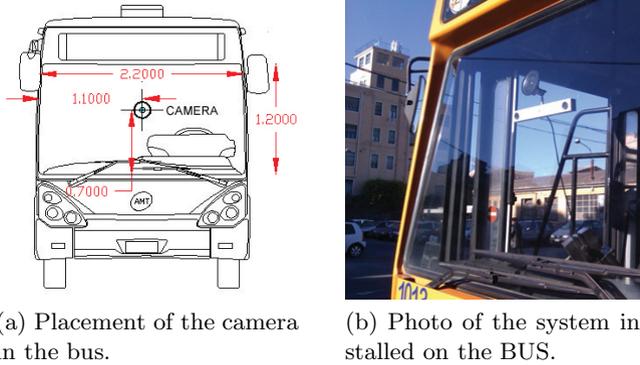
where  $V_v(i)$  is the velocity of the vehicle at time  $i$  and  $\Delta T(i) = T_f(i) - TTC_v(i)$ . The RI for the entire conflict is given by the following formula:

$$RI = \frac{\frac{1}{N} \sum_i RI(i)}{TTZ_{duration}}. \tag{6}$$

In the case of an obstacle moving along the direction of the vehicle, as in a “car following” case, the  $TTC_o$  will be 0 and therefore the conflict will occur only if  $TTC_v(i) < T_f(i)$ .

### 3 Hardware Components

The hardware employed in the proposed system is composed by a number of devices connected together. Specifically, a TYZX DeepSea G3 Embedded Vision System (EVS) [7] is employed to acquire stereo images and to obtain a depth map of the scene in front of the vehicle. This piece of hardware is positioned in the front of the vehicle (i.e., a bus in our case) as depicted in Fig. 2.



**Fig. 2.** The proposed system on a urban BUS during the experimental phase

The EVS implements a Census-Based stereo algorithm [8]. As the input pixels enter the EVS, the Census transform is computed at each pixel based on the local neighborhood, resulting in a stream of Census bit vectors. At every pixel a Hamming distance is used to compare the Census vectors around the pixel of one view (i.e., left image) to those at 52 locations in the other view (e.g., right image). These 52 comparisons are performed simultaneously making the stereoscopic system very fast at subpixels precision. The EVS processor converts the pixels disparity map to metric distance measurements using the stereo camera's calibration parameters and the depth units specified by the user [7]. In our setting a stereo cameras with a baseline of 33cm and an 83 degrees HFOV lens has been used. This configuration allows the overall system to work with distance in a range between 2.5m and 50m. All the specifications of the EVS used for the experiments are reported in Table 1.

**Table 1.** Datasheet of the EVS hardware used in our experiments

G3 Embedded Vision System Specifications	
<b>Size</b>	3.8cm x 18.7cm x 14.5cm
<b>Power</b>	11W typ.: 12vdc or PoE class III
<b>Imagers</b>	Aptina MT9V022 color, up to 60fps, 752x480 pixels
<b>Lens</b>	83 degrees horizontal FOV
<b>Stereo Baseline</b>	33 cm
<b>CPU</b>	Freescale PowerPC 8347 @400MHz
<b>Memory</b>	256 Mbytes
<b>Operating System</b>	Linux 2.6 Kernel

In the proposed system, the EVS is integrated with a series of components: a router, a GPS module and a notebook. The router solves the EVS data access problem. We used a gigabit Ethernet WLAN router to create a LAN in which the EVS can be accessed by a notebook. A client can access the EVS by connecting to a TCP service that sends each captured image and its computed depth map. For the application described in this paper, image and depth information

must be geo-referenced. To achieve this requirement, we connected the EVS to a Bluetooth GPS receiver module. The choice of a wireless device was mandatory to solve the signal satellite problem: the specific shape of an urban BUS forced us to put the GPS module placed on the external roof of the bus to have the best possible number of satellites in sight. We connected the Bluetooth GPS module to the notebook communicating with the EVS through LAN network. Some ad-hoc routines have been developed to synchronize the GPS information with every single frame acquired by the EVS. Finally we have a notebook connected to the LAN running the software layer which coordinates the different hardware. The SW also performs the acquisition and analysis of the data. It gives commands to the EVS, receives data both from the EVS and the GPS module. It synchronizes every single EVS information making some computation on it and finally it gives a single output as archive files that can be opened and analyzed by a specific module, after the data acquisition phase.

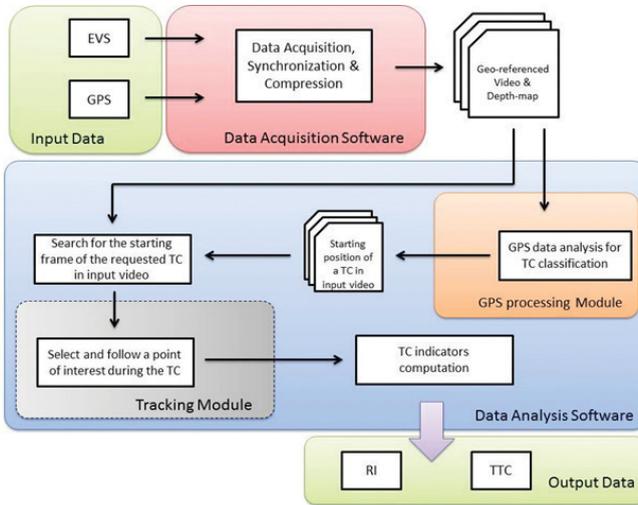
After mounting the hardware on the vehicle (i.e., the bus), a dedicated calibration procedure must be done. This is because the measures of the indices of traffic conflicts depend on the reference system that must be integrated to the vehicle. The calibration procedure consists in placing outside and in front of the vehicle two targets orthogonal to the axis of the vehicle at 3 meters away from the camera. The targets are placed using an high precision laser rangefinder positioned on vehicles vertices and pointed to the targets. At this point, with the help of the data acquisition software in its initialization phase, the system evaluates the distances of the two targets and records the distance difference between the right target and the left one. This difference is due to the inclination of the camera related to the axis of the vehicle and the related values will be used for subsequent analysis as will be described in next section.

## 4 Software Components

As detailed in the following sub-sections, the software layer is composed by two main modules: one for data acquisition and the other for data analysis. The schema of the different components involved in the software layer are reported in Fig. 3.

### 4.1 Data Acquisition Module

This module takes in input the data coming from EVS and GPS as reported in Fig. 3. The EVS outputs are the color image sequences taken from the left camera, and the related depth-maps. The images are compressed according to the MPEG-4v3 codec and stored on a dedicated device. The depth-maps are represented with a 16 bits precision. From the depth-map stream a video file is obtained (the lossless compression codec FFV1 is used in this case). In this configuration, the software produces an amount of data of about 10GB every 2 hours of recording. Since the context of use of our system is related the urban buses traffic, we decided to stop the recording when the bus is not moving in order to reduce the amount of stored data of about 30%.



**Fig. 3.** The software layer of the proposed traffic conflict analysis system

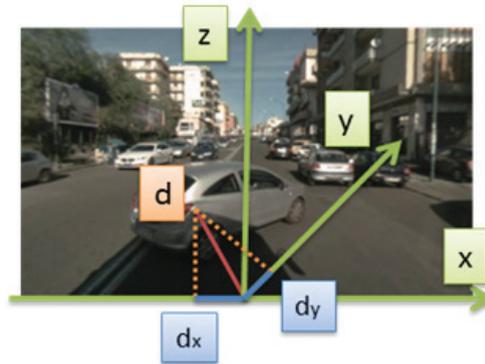
As previously mentioned, the images (and the corresponding depth-maps from the EVS) need to be geo-referenced through an external GPS. The data acquisition module deals with this issue by solving the problem of the difference in the sampling frequency between the video information (20-30Hz) and the GPS (10Hz). For this purpose, the software operates as follows:

- i) When a new video information arrives (i.e., a frame) the software operates a sample and hold mechanism on the GPS information by associating the last GPS information to each video frame. In this way there will be some frames corresponding to the same GPS information;
- ii) When a new GPS information arrives, the acquisition module corrects all the previous informations with the same data interpolating the position and velocity values according to the data of the newly arrived GPS information.

Finally, the acquisition module produces a textual output that allows to bind each frame and depth-map to the corresponding GPS information. Since both the GPS and the EVS are connected to the computer via Bluetooth and LAN respectively, in the initialization phase, the module makes an evaluation of the latencies on a single packet transmission. This operation allows the acquisition module to know the arrival delay from the generation of each individual package to correct the output. In this way it's possible to obtain a more precise synchronization between each frame and the related GPS information.

## 4.2 Data Analysis Module

The data analysis module allows the expert to operate on the geo-referenced frames and depth-maps to obtain the traffic conflict indicators used for the analysis (see Section 2). The analysis starts from the first frame of a conflict which is selected



**Fig. 4.** Reference system used by the data analysis module

through a manual search in the video, or in an automatic way by processing the information provided by the GPS (i.e., vehicle speed, vehicle trajectory, acceleration, jerk and angular deviation rate) [11]. Once the initial frame of the traffic conflict is detected, the obstacles that may have generated the conflict is identified as the closest point to the vehicle. The obstacle is then tracked in the successive frames in order to obtain the traffic conflict measurements. For tracking the obstacle we have used a modified version of the well known mean-shift tracking algorithm [9]. The algorithm has been extended in order to take into account information provided by the depth-map and considering the speed information provided by GPS. Both depth and speed of the vehicle are useful to perform a fast binary segmentation and hence remove information which are not related to the obstacle, making the tracker more robust to clutter background (e.g., if the obstacle is a vehicle, the other vehicles or objects similar with respect to visual information are removed by using the distance given by the depth-map information).

Hence the closest point of the obstacle is tracked exploiting the data coming from the depth-map is possible to fully reconstruct the behavior of the vehicle and the target obstacle involved in the conflict in the space in front of the vehicle itself. The data analysis module hence computes speeds and distances of both vehicle and obstacle in X and Y axis components by considering the reference system synthetically sketched in Fig. 4. The quantities thus obtained are then used to compute the traffic conflict indicators as described in Section 2.

## 5 Experimental Phase

The described system has been used to analyze the behavior of bus drivers in urban context. The experiments have been performed on real traffic condition in the city of Catania, Italy. The system has been mounted on a urban bus and four different acquisition sessions of about 2 hours each have been carried. Using the acquired data, traffic conflict analysis was carried out. The results provided by the analysis are useful to understand the traffic behaviour and hence take



**Fig. 5.** Five frames and the related depth-maps of a pedestrian crossing the street hence generating the conflict

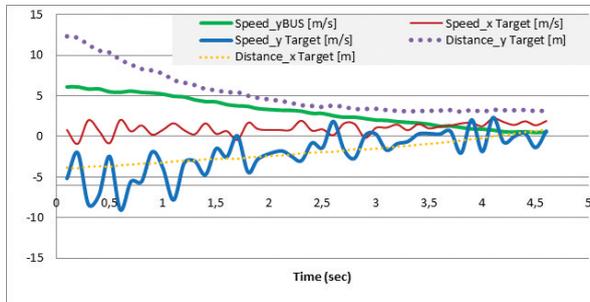
actions (i.e., installation of traffic lights) in order to prevent road accidents. In the following we present a case study, as well as the results obtained from the analysis performed with the system, related to a traffic conflict in which a pedestrian is involved. Among the other traffic conflicts (e.g., “car following”) in the acquired data (about 8 hours of recording), the case here discussed is the most frequent and representative. Similar results have been found in the other traffic conflict cases. Due space limit, the other traffic conflicts examples (e.g., car following) will be presented and discussed in detail at conference time.

The conflict detailed in the following is generated from a pedestrian crossing the street. In this case the pedestrian decided to cross at the very last moment forcing the bus to make a stop and generating a conflict. Fig. 5 reports a sequence of frames and the related depth-maps which show the conflict tracked and analysed by the system. The data analysis module computes the behavior of the actors involved in the conflict, in terms of speed and distance on the two axes components, according to the reference system. The computed measures are reported in Fig. 6. The quantities extracted from the data analysis module are then further processed to product the measures of a TC. The obtained  $TTC$  for vehicle and pedestrian and the  $T_f$  are shown in Fig. 7. As shown in Fig. 7 the TC exists because there’s an area in which  $TTC_o < TTC_v < T_f$  (see Section 2). After obtaining traffic conflict indicators the data analysis module calculates the parameters indicating the severity of a conflict (RI) (see Equations (5) and (6)). The results of the RI computed in the case study presented here are reported in Table 2.

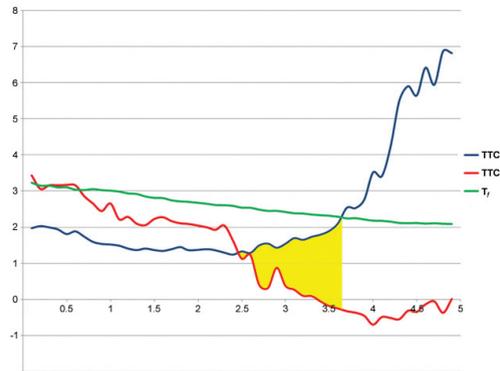
Given the variability and complexity of driver behaviors and performance, the random and rare nature of crashes and lack of adequate pre-crash data in the current crash record, it is becoming increasingly apparent that the necessary human factor data collection should be obtained from a “naturalistic” setting.

**Table 2.** Risk Impact results for the presented traffic conflict

	MAX RI	MEAN RI
<b>Pedestrian</b>	15.5613	9.2091



**Fig. 6.** Trend of speed and distance of vehicle and target in X and Y components



**Fig. 7.** Traffic conflict measures over time. The conflict phase is highlighted in yellow.

These approaches fill voids in the existing driving safety research methods. Specifically, the proposed method provides much greater information regarding pre-crash and crash events than what is currently available, even after a detailed crash investigation. Furthermore, the data provides much greater external validity relative to the larger context of driving than that of empirical methods, such as test tracks or simulators. For instance if the two types of conflicts are recurrent events in the same points a better regulation of the crosswalk/bus line could be analyzed.

## 6 Conclusions and Future Works

The main aim of traffic conflict analysis is the objective evaluation of traffic safety and the prediction of fatal crashes. In this paper we have presented a system able to make measurements of the traffic conflicts by exploiting geo-referenced

stereo sequences and tracking procedure. The effectiveness of the system has been demonstrated on real urban environment. Future works will be devoted to exploit advanced tracking techniques (e.g., TLD [10]) combined with depth information to better deal with appearance changes (i.e., pose, occlusion, etc.) and to allow re-detection of obstacles which disappear in some frames of the conflict. Also, some prior information about the context could be used to refine the depth map [12]. Moreover, the ability to detect and discriminate some classes of obstacles (e.g., pedestrians, cars) will be included in order increase accuracy of the computation of the traffic conflict risk. Finally, we plan to build a labeled dataset involving different classes (e.g., crashes, near-crashes, other “accidents”) and information coming from video, GPS, etc., as well as information about the driver behavior (e.g., head pose). The dataset will be used to understand the nature of the conflict events and to determine if the near-crashes could be quantitatively and automatically identified.

**Acknowledgments.** The authors wish to thank the Italian Ministry of Economic Development for the financial support of this research within the program “Industria 2015” and the Public Transport Company AMT Catania for the kind availability of the bus.

## References

1. Chin, H.C., Quek, S.T.: Measurement of Traffic Conflicts. *Safety Science* 26(3), 169–187 (1997)
2. Migletz, D.J., Glauz, W.D., Bauer, K.M.: Relationships between Traffic Conflicts and Accidents. Report No: FHWA/RD-84/042. US Department of Transportation, Federal Highway Administration (1985)
3. Heinrich, H.W.: *Industrial Accident Prevention*. McGraw-Hill, New York (1932)
4. Hyden, C.: The Development of Method for Traffic Safety Evaluation: The Swedish Traffic Conflict Technique. Bulletin 70. Lund Ins. of Technology, Sweden (1987)
5. Songchitruksa, P., Tarko, A.P.: Extreme value theory approach to safety estimation. *Accident Analysis & Prevention* 38, 811–822 (2006)
6. Cafiso, S., Garcia, A.G., Cavarra, R., Romero Rojas, M.A.: Crosswalk safety evaluation using a pedestrian risk index as traffic conflict measure. In: *The 3rd International Conference on Road safety and Simulation* (2011)
7. Woodfill, J.I., Gordon, G., Buck, R.: Tyzx DeepSea High Speed Stereo Vision System. In: *IEEE Workshop on Real Time 3-D Sensors and their Use, IEEE Conference on Computer Vision and Pattern Recognition* (2004)
8. Zabih, R., Woodfill, R.: Non-parametric Local Transforms for Computing Visual Correspondence. In: Eklundh, J.-O. (ed.) *ECCV 1994*. LNCS, vol. 801, Springer, Heidelberg (1994)
9. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-Based Object Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(5) (2003)
10. Kalal, Z., Mikolajczyk, K., Matas, J.: Face-TLD: Tracking-Learning-Detection applied to faces. In: *IEEE International Conference on Image Processing* (2010)
11. Cafiso, S., Di Graziano, A.: Automated in-vehicle data collection and treatment for existing roadway alignment. In: *Efficient Transportation and Pavement Systems: Characterization, Mechanisms, Simulation, and Modeling - International Gulf Conference on Roads*, pp. 785–797 (2008)
12. Battiato, S., Curti, S., La Cascia, M., Scordato, E., Tortora, M.: Depth Map Generation by Image Classification. *SPIE Electronic Imaging* (2004)