Near-Miss Accidents – Classification and Automatic Detection

Georg Thallinger[1](✉), Florian Krebs[1], Eduard Kolla[2], Peter Vertal[2], Gustáv Kasanický[2], Helmut Neuschmied[1], and Karl-Ernst Ambrosch[2]

[1] DIGITAL,JOANNEUM RESEARCH, Graz, Austria
{georg.thallinger,florian.krebs,helmut.neuschmied}@joanneum.at
[2] University of Žilina, Žilina, Slovakia
kolla@uniza.sk,peter.vertal@uzvv.uniza.sk,gustav.kasanicky@usi.sk,
karl.ambrosch@erachair.uniza.sk

Abstract. In this work, we propose a system that automatically identifies hazardous traffic situations in order to gather comprehensive evidence, allowing timely mitigation of dangerous traffic areas. The system employs optical and acoustic sensors, stores the recorded sensor data to an incident store, and provides an assessment of the causes and consequences of the captured situation. Three main categories of features are used to assess the risk of a traffic situation: (1) key parameters of the traffic participants such as size, their distance, acceleration and motion trajectories; (2) the occurrence of acoustic events (shouting, tire squealing, honking sounds, etc.) which often co-occur with hazardous situations; (3) global parameters which describe the current traffic situation, such as traffic volume or density. An automated detection allows to monitor an intersection for an extensive time period. Compared to traditional manual methods, this facilitates generating significantly more data, which increases the informative value of such an assessment and therefore leads to a better understanding of the hazard potential of the spot. The outcome of such an investigation will finally serve as a basis for defining and prioritizing improvements.

Keywords: Near-miss accidents · Accident detection
Automatic detection · Sensor fusion

1 Introduction

Near-misses, hazardous traffic situations without personal or material damage, are rarely registered by the authorities. Nevertheless, they are important indicators to identify *accident black spots*. According to the Austrian research association for roads, railways and transport (FSV[1]), a junctions or street section up to a length of 250 m is defined as accident black spot, if either at least three similar accidents with bodily injury have occurred within three years, or at least

[1] www.fsv.at.

T. Kováčiková et al. (Eds.): INTSYS 2017, LNICST 222, pp. 144–152, 2018.
https://doi.org/10.1007/978-3-319-93710-6_16

five accidents with bodily or material damage have occurred within one year[2]. If such an accident black spot is identified, the FSV suggests to carry out a road safety inspection with the aim of defusing the black spot. Since December 2010 (implementation of the EU Directive 2008/96/EC), this approach significantly increased the road safety due to its structured data collection methodology.

However, this is a purely reactive approach, where several accidents must first occur in order to bring about an improvement. Obviously, it is desirable to detect black spots *before* accidents happen, by a structured analysis of hazardous situations including near-misses. However, to draw statistically valid conclusions, a spot has to be monitored over a long period of time, which is unfeasible to do manually.

In this paper, we propose a system to analyze the traffic situation *automatically* over a long time period, by detecting hazardous situations with optical and acoustical sensors. Three main categories of features are used to assess the risk of a traffic situation: (1) key parameters of the traffic participants such as size, their distance, acceleration and motion trajectories; (2) the occurrence of acoustic events (shouting, tire squealing, honking sounds, etc.) which often co-occur with hazardous situations; (3) global parameters which describe the current traffic situation, such as traffic volume, density or the states of the traffic lights.

This creates a well-founded basis for defining and prioritizing improvements. In the case, that hazardous situations occur, possible solutions are to change the traffic pattern or to separate users in time. A repeated misbehavior of traffic participants can be detected and might become the subject of campaigns. Data generated through our approach can support road safety audits and give guidance for traffic planners and technicians to mitigate the hazardous potential of an intersection.

As the project is still in progress, we offer in this paper a general overview about the system. A thorough evaluation of the individual components is left for future work.

2 Related Work

Both video and audio sensors have been used for analyzing traffic situations. Below we review the most important approaches.

2.1 Video Analysis

In order to automatically detect hazardous traffic situations by video surveillance, all traffic participants first must be recognized, tracked and classified from the image data. This is a challenging problem due to various lighting situations, shadows, multiple objects, concealments, viewing angles and insufficient image resolution for detecting small traffic participants (people, bicycles, etc.).

Various visually identifiable indicators have been developed to assess the hazardousness of a traffic situation [1]. According to [2], these indicators can be

[2] RVS 02.02.21.

divided into two main categories: In the first group, unusual events are detected by first detecting the motion paths of the tracked objects and then detecting deviations from these movement paths. For this purpose, the speed, speed changes, and directional changes are also used as a basis for comparison (positional deviation from the paths, at a standstill from the stop lines) [3,4]. In the second group, methods for predicting accidents based on the analysis of collision parameters, e.g. vehicle alignment, speed and vehicle distances used by road users [5,6].

Finally, Ki and Lee [14] developed a vision-based system for automatically detecting, recording, and reporting traffic accidents at intersections.

2.2 Audio Analysis

In addition to the visual analysis, acoustic sensors have been used to monitor traffic. Most publications concentrate on determining the strength of the traffic flow [8,10] or detecting traffic jams [9]. The sensors used are either microphone arrays or distributed acoustic sensors (microphones) [7].

The first system for automatically detecting traffic accidents at intersections solely based on acoustic data was proposed by [16]. A neural network was used to classify sound events into crash and non-crash classes. Later, a similar system, the automated incident recording system (AIRS) [13], was proposed. It consists of two directional microphones, which are used to identify incidents and near-misses acoustically and two video cameras which are used to store the accompanying video footage (including four seconds before and after the event) for later analysis. The system was installed at an intersection in Louisville, Kentucky. After a 38 months testing period, a total of 92 incidents and 201 near-misses were recorded by AIRS. A redesign of the intersection, as supported by the AIRS data, was reported to reduce the total number of crashes by 14 percent, measured over a period over three years [17]. Finally, Foggia et al. [18] proposed a system for audio surveillance of roads, which focused on detecting tire skidding and car crash sounds.

3 Process and System Overview

In the project SIMMARC (**S**afety **Im**provement Using Near **M**iss **A**nalysis on **R**oad **C**rossings), we aim at developing a tool for automatically assessing the hazardous potential of a road crossing.

To that end, the selected traffic area will be long-term monitored by various sensors. Incidents are automatically detected, classified and the recording of the relevant video-audio footage will be saved to an *incident store*. In addition to the sensor data, the traffic light signals can be recorded. For the final analysis, a software tool will be developed that carries out a statistical evaluation of the recognized incidents. The tool will offer various ways to manually inspect and analyze the detected situations (incident time-table, playback options, etc.), as well as offer an assessment of the causes and potential consequences. The work flow is illustrated in Fig. 1.

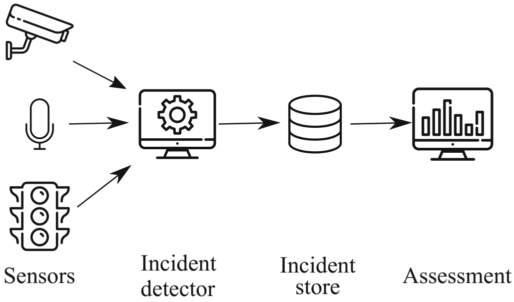

Fig. 1. System overview. (icons from www.flaticon.com)

4 Sensors

The amount and types of the sensors have to be adapted for each site individually. Here, we describe a typical setting for a small crossing in the city center of Graz, Austria. An image of the crossing is shown in Fig. 2.

4.1 Video

An AXIS Q1615 Mk II-Network Camera was installed at about 10 m height onto a mobile pole (see the red circle at the top of Fig. 3). The camera view is shown in Fig. 2.

4.2 Audio

A directional microphone was installed at 5 m height facing the center of the crossing (see the red circle at the center of Fig. 3).

Fig. 2. Crossing at Dietrichsteinplatz, Graz.

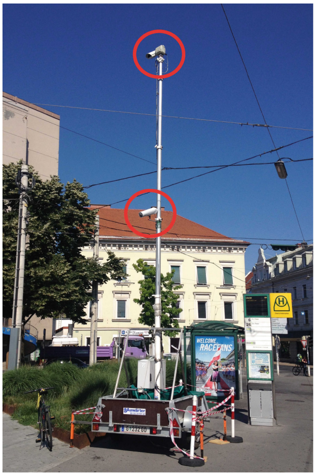

Fig. 3. Mobile pole. (Color figure online)

4.3 Traffic Light State Information

In order to reason about the causes of near-misses, we also record—where available—the state of the traffic lights over time.

5 Incident Detection

5.1 Vision Based

As a first step towards an automatic security assessment, low-level information is extracted from the raw sensor data. For video signals this means extracting the trajectories of objects and distances between objects in order to find those cases where trajectories cross and an acceptable minimum distance is violated. We use a feature based tracking method where distinguishable points in the image will be tracked and the resulting trajectories have to be clustered to assign them to individual objects like cars, bicycles, and pedestrians. The feature tracking is done by using the Kanade-Lucas-Tomasi (KLT) Feature Tracker [11]. Due to perspective changes and occlusions the algorithm has to handle trajectory disruptions and trajectories that end at background points. To handle this problem we use a trajectory clustering method similar to the one described in [12]. Clustering is achieved by constructing a graph of grouped tracking points at each frame. By checking the consistence of the graph structure over time, it can be detected, if the graph points belong to the same object or not. Additionally, we use a Kalman filter to predict object motion at occlusion image areas where no KLT feature points of an object can be extracted. Further, tracking errors are detected by continuously updating a background model, holding the KLT points which belong to the background area. This helps to detect trajectories which end at background points.

In the next processing steps we classify the road users based on the object size, speed, detection position and the similarity of their trajectories with the main moving directions of the different road users (obtained by manual specification).

To detect near-misses we also define the *safe space* of a participant (see Fig. 4 for an example), which is a function of kinematic variables:

$$S_i = f(d_{front_i}, d_{lat_i}, d_{rear_i}) = g(\mathbf{v}_i, \mathbf{a}_i, t_{r_i}) \quad (1)$$

where $\mathbf{v}$ is the velocity vector, $\mathbf{a}$ is the acceleration (deceleration) vector, t_r is the driver reaction time, and $g()$ is a function that can be designed by the user. A near-miss occurs if a vehicle enters the safe space of another vehicle without touching the other vehicle.

Finally, additional statistical data like traffic amount, turning movement counting, and average crossing speeds are collected to support the analysis of the traffic situation.

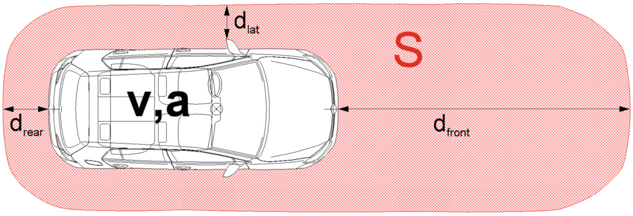

Fig. 4. Safe space around vehicle.

5.2 Audio Based

One of the aims of this project is to investigate which kind of audio features/events coincide with hazardous traffic situations and how to detect these robustly. In a first step we will therefore detect audio events that are suspect to accompany hazardous situations such as honking of car horns, braking sounds, glass breaking, crash sounds and yelling. Later, we will also look for untypical traffic related sounds and investigate their relation to hazardous situations. The detection of sound events in a traffic scenario is challenging due to the great number of sound sources which are usually present at a crossing. For example, Chee and Brisbane [13] report a very high number of false alarms (≈91%) due to disturbing background noise such as traversing of an uneven manhole cover, loud vehicles, tire noises, and sirens. Therefore, we aim at exploiting recent advances in machine listening to alleviate these issues. Precisely, we plan to use deep neural networks to detect the audio events, which has proven to be excellent for audio event detection [15].

Acoustic monitoring will further provide additional information such as the ambient noise level, traffic density, and the composition of the traffic participants (e.g., bicycles, trucks, pedestrians, etc.).

Given a certain distribution of recording microphones, we can also derive spatial information of an acoustic event. This may be important for more complex road crossings.

6 Assessment

In the last stage of the system the recorded traffic situations are evaluated. Incidents are classified, and possible reasons and consequences are inferred automatically to create statistics that can be interpreted and acted on by a human traffic planer.

6.1 Incident Classification

Initially, a classification of hazardous incidents has been defined and risk factors leading to such incidents were identified which will then be used by the tools developed in the project.

Conflict situations between motor vehicles:

- Conflicts in the interlacing of favored and preferred vehicles. Risk factor: Inadequate or too short acceleration and interlacing distances.
- Unauthorized driving strip changes. Risk factor: unclear route signposting. Too short distances between connecting points in the high-level road network often do not allow the necessary pre-announcements to be made and lead to directions where the probability of a misinterpretation increases.
- Priority traffic is forced to perform braking maneuvers even if the speed limit is complied with. Risk factor: low sighting distance, especially at junctions.
- Risky overtaking maneuvers. Risk factor: Lack of overtaking possibilities on long road sections encourage drivers of faster vehicles behind slow-moving vehicles to carry out risky overtaking maneuvers even if the visibility is not sufficient.
- Inharmonious traffic flow increases the probability of rear-end collisions. Risk factor: Large variance in the speed distribution caused, for example, by slow-moving heavy vehicles on uphill routes.
- Enlarged breaking distance. Risk factor: Deficient drainage of the road surface due to incorrect transverse and longitudinal inclination conditions leads to water accumulation with an increased risk of aquaplaning or ice formation.

Risk factors for traffic situations with motorized and non-motorized road users:

- Especially in intersection areas, there are many points of conflict (crossing, threading out, grading) for the traffic flows. A non-signal-controlled four-armed intersection, at which all driving relations are permitted, has 32 points of conflict for motor vehicle and bicycle traffic alone.
- Pedestrian crossings or additional traffic flows (e.g. bus lanes) in the intersection area considerably increase the number of possible conflict points. Unclear traffic flows, poor visual relationships, scarce or ignored clearance signaling lead to an increase in the likelihood of conflict in these sensitive areas.
- Pedestrian crossings represent an additional potential for conflict, especially on open roads. Marked, but not signal-controlled pedestrian crossings often convey a false sense of safety for pedestrians. The lengths required for stopping the vehicles are wrongly estimated and thus increase the probability of conflicts.

Further risk factors:

- Strong changes of light and shadows impair the perception of road traffic and increase the probability of conflicts.
- Advertising near the traffic area distracts the driver's attention from traffic and increases the probability of conflicts.

6.2 Incident Simulation

In addition to the detection and categorization of incidents, we can also use simulation to reason about the causes and consequences of an incident.

The simulation is based on a fusion of the video recording and a kinetic numerical simulation using the PC Crash v11.1 software[3]. To that means, a kinetic mapping is created from the movement of a real vehicle (in a video recording) to the movement of a simulated vehicle, based on volume matching. Computing such a mapping has two main advantages:

1. The mapping allows to infer hidden, underlying kinetic variables. For example, it enables computing the forces that act on the human body, or obtaining information about the probable driver behaviors (e.g., how he/she was manipulating the steering wheel).
2. The kinetic mapping can be used to increase the temporal resolution by model-based interpolating between video frames. This can be important for very rapid movements such as a sudden braking, additionally this enables using video of lower quality.

In a first test run, a vehicle was equipped with a Corrsys Datron Microstar measuring system[4] that provides an accurate measurement of the vehicle speed exploiting the Doppler effect. The recorded trajectories were then matched to a video recording. Preliminary results indicate that the trajectories computed from the video recording are very close to the measured ones.

7 Conclusions

We proposed a system to automatically detect, document and assess potentially hazardous near-miss accidents. As the system runs automatically, it is possible to capture a comprehensive amount of situations, which is important to draw statistically valid conclusions. Using such an approach allows to proactively analyze and defuse potential accident black spots, and therefore facilitates the work of traffic planners and helps to avoid accidents. In future work, we plan to evaluate the system using real data that were recorded at test sites.

Acknowledgments. Work leading to this paper is done in the project SIMMARC - “Safety IMprovement using near Miss Analysis on Road Crossings”, partially funded by the Austrian Research Promotion Agency.

References

1. 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, Academic Thesis (2004)
2. Er, U., Yuksel, S., Akoz, O., Karsligil, M.E.: Traffic accident risk analysis based on relation of common route models. In: 21st International Conference on Pattern Recognition (ICPR) (2012)

[3] www.dsd.at.
[4] www.datron.sk.

3. Junejo, I.N., Javed, O., Shah, M.: Multi feature path modeling for video surveillance. In: Proceedings of the International Conference on Pattern Recognition, pp. 716–719, August 2004
4. Morris, B., Trivedi, M.: A survey of vision-based trajectory learning and analysis for surveillance. IEEE Trans. Circuits Syst. Video Technol. **18**(8), 1114–1127 (2008)
5. Saunier, N., Sayed, T.: A probabilistic framework for the automated analysis of the exposure to road collision. J. Transp. Res. Board 96–104 (2008)
6. Shirazi, M.S., Morris, B.: Observing behaviors at intersections: a review of recent studies & developments. In: IEEE Intelligent Vehicles Symposium, Seoul, Korea (2015)
7. Barbagli, B., Magrini, I., Manes, G., Manes, A., Langer, G., Bacchi, M.: A distributed sensor network for real-time acoustic traffic monitoring and early queue detection. In: 2010 Fourth International Conference on Sensor Technologies and Applications (SENSORCOMM), pp. 173–178, 18–25 July 2010. https://doi.org/10.1109/SENSORCOMM.2010.102
8. Borkar, P., Malik, L.G.: Acoustic signal based traffic density state estimation using adaptive neuro-fuzzy classifier. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ), pp. 1–8, 7–10 July 2013. https://doi.org/10.1109/FUZZ-IEEE.2013.6622444
9. Sen, R., Siriah, P., Raman, B.: RoadSoundSense: acoustic sensing based road congestion monitoring in developing regions. In: 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), pp. 125–133, 27–30 June 2011. https://doi.org/10.1109/SAHCN.2011.5984883
10. Stadtschnitzer, M., Rainer, H., Graf, F., Fellner, M.: Improving video surveillance systems using acoustical source localization. In: 1st EAA-EuroRegio, Ljubljana (2010)
11. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: International Joint Conference on Artificial Intelligence (1981)
12. Saunier, N., Sayed, T.: A feature-based tracking algorithm for vehicles in intersections. In: Proceedings of the 3rd Canadian Conference on Computer and Robot Vision, CRV 2006 (2006)
13. Chee, D., Brisbane, G.: Automated incident recording system (AIRS) – trial study. In: Road Safety research, Policing and Education Conference, Sydney, New South Wales, Australia (2003)
14. Ki, Y.K., Lee, D.Y.: A traffic accident recording and reporting model at intersections. IEEE Trans. Intell. Transp. Syst. **8**(2), 188–194 (2007)
15. Çakir, E., Parascandolo, G., Heittola, T., Huttunen, H., Virtanen, T.: Convolutional recurrent neural networks for polyphonic sound event detection. In: IEEE Transactions on Audio, Speech and Language Processing, Special Issue on Sound Scene and Event Analysis (2017)
16. Harlow, C., Wang, Y.: Automated accident detection system. Transp. Res. Record J. Transp. Res. Board Nr. **1746**, 90–93 (2001)
17. Green, E.R., Agent, K.R., Pigman, J.G.: Evaluation of Auto Incident Recording System (AIRS) (2005)
18. Foggia, P., Petkov, N., Saggese, A., Strisciuglio, N., Vento, M.: Audio surveillance of roads: a system for detecting anomalous sounds. IEEE Trans. Intell. Transp. Syst. **17**(1), 279–288 (2016)