

System for Detecting Urban Areas with High Density of Vehicle Incidents and Issues

Enrique Puertas, Nourdine Aliane, María de la Luz Morales-Botello, Javier Fernández

Dept. Informática, Automática y Comunicaciones

Universidad Europea de Madrid

Villaviciosa de Odón, Madrid, SPAIN

{enrique.puertas, nourdine.aliane, mariadelaluz.morales, javier.fernandez}@uem.es

Abstract—This paper presents an on-board perception and pre-collision system, aimed at detecting traffic risky situations in urban areas, as well as trying to avoid collision with other traffic participant deploying automatic braking and/or evasive steering maneuver. Information related to the detected risky situations is recorded in a vehicles' local database. Local data are then transferred later into a global database that can be exploited for finding out remarkable traffic hotspots. This analysis is aimed at analyzing risky situation patterns, optimize resource allocation, and improve emergency call responses.

Keywords-Data Mining; Event Data Recorder; Black-box; Accident Hotspots; Intelligent Transport Systems; Computer Vision; Road Safety; Traffic

I. INTRODUCTION

In urban areas, risky traffic situations can come up everywhere and in any moment; for example, they may come up driving close to schools or in residential areas, driving through intersections and roundabouts, etc. These hazardous situations can be in form of assaults on pedestrians, cyclists or animals, a collision between two or more vehicles, or a collision with either mobile or fixed objects.

According to Spanish transport agency (DGT), the number of fatalities in road accidents in the last 10 years (2001-2010) has decreased by 44%. However, the reduction of urban accidents in the same period is about 16%. Furthermore, in 2010 urban accidents with victims is about 54% of the total, where 87% of the victims are pedestrians. It is worth mentioning that people of over 75 years is the most affected group reaching almost 22% of the total deaths. According to the same report, in 2010 urban accidents occurring at intersections reach almost 50% with a mortality rate of 37%. In contrast, accidents occurring at through roads are apparently fewer and barely reach 1.3% of the total, but with a mortality rate of 9%. The most common types of urban incidents are running traffic control (red light, stop, or yielding signs), striking vehicles or objects, running off roads, and turning left. It is also worthy to mention that not all of these risky situations should be attributed to drivers, but many of them are due to deficiency in road signs, to inadequate design of urban roads, to lack of visibility, etc.

In order to improve road safety, the European Commission is committed by integrating technology in

cars and road infrastructure through several initiatives such as e-Call, or ERTRAC. Making driving safer has been also field of many research aimed at incorporating to vehicles driver assistance systems (DAS). For example, the use of computer vision focused on identifying traffic signs [1], lane estimation [2], or fatigue and distraction [3, 4], and then warn drivers accordingly. Laser sensors are also used for detecting and tracking pedestrians as well as fixed or mobile obstacles [5]. New DAS support drivers not only during the normal driving, but also in hazardous situations by braking [6], and consequently preventing or at least mitigating eminent collisions. There are many situations where full braking is too late and will not avoid collisions. So, new driver assistance systems provide steering intervention [7, 8 and 9] as an additional option to prevent collisions.

In this paper, our concern is focused on the fact that many risky situations could be detected by using a suitable on-board sensing system and then recorded. Afterwards, by the use of appropriate tools, statistical techniques can be applied to the recorded data to find out hotspots and locating them in a map. Identifying hotspots in urban area, where collisions or risky situations have taken place, is an important concern for local governments for understanding the causes of crashes and to determine effective countermeasures based on the analysis of the causal factors.

In this sense, geographical information system (GIS) technology has been used by many traffic agencies for accident analysis [10, 11, 12]. However, the reporting of accidents still uses the traditional way by using textual reports, and thereby making it difficult to analyze accident data and risky situations [18]. Knowing where risky situations are concentrated helps law enforcement and make better decisions in terms of strengthening the traffic monitoring such as improving signal and sign visibility, increasing sight distances, and reducing vehicle speeds near intersections. It could be helpful to enhance traffic participant protection by improving road infrastructures, optimize resources allocation, and improve emergency call response, and so on.

This paper presents a system aimed at identifying risky traffic situations and then recording some specific vehicles data (speed, GPS location,) as well as a track of surrounding pictures evidencing the circumstances when they take place. The system itself is composed of: (1) a perception subsystem based on a computer vision and laser

technology aimed at obstacle detection; and (2) a pre-collision subsystem with an automatic braking and steering devices allowing the vehicle to stop or change the trajectory depending on the available free space around the vehicle.

The rest of the paper is organized as follows: the concept of risky situations is clarified in section 2. Event Data Recorder is described in section 3. In next, the approach used in the decision engine is described. Sections 5 and 6 discuss hotspots analysis and visualization. Finally, some conclusions are drawn in section 7.

II. RISKY SITUATIONS

In common urban traffic, there are driving circumstances, referred to as risky situations that increase the likelihood of occurring accidents. These risky situations can be classified as follow:

Collision: Any circumstance leading to any contact with moving or fixed objects. These situations are not limited to contact with pedestrians or other vehicles, but includes all sort of impacts between vehicles and roadside objects such as barriers, trees, poles and road signs, etc.

Pre-Collision: Any circumstance leading to an automatic evasive maneuver by the vehicle to avoid a collision. An automatic evasive maneuver can be a breaking, a steering, issuing an alarm, or any combination of previous controls.

III. EVENT DATA RECORDER

Detected risky situation involving vehicle are recorded in a local EDR before sending them to a global database. The local EDR has been implemented on an embedded hardware running SQ Lite as the transactional SQL database engine.

The local database is modeled as a collection of objects, called urban trips, where their main attributes are Paths and Risk events. Path objects are recorded as a set of GPS locations taken at rate of 5 samples per second, and Risk objects associated to risky situations. Risky situations, identified by a decision algorithm, are associated to two entities referred to as Pre-collision and Frame. Pre-collision entity is a set of data associated with automatic maneuver by the vehicle to avoid a collision. An automatic maneuver can be a braking, steering, or issuing light or acoustic alarm, where each attribute is associated to two numbers indicating intensity and duration. Frame entity is a set of data related to detected obstacles and their identification if it is possible such as pedestrian, short obstacles (animal, stroller, cart,...), or tall obstacles (tree, lamppost,...), GPS location where the hazardous situation have taken place, and a track of 10 to 15 pictures of the surrounding. The storage of images taken in urban areas is only performed after an anonymizing process, which consists in removing faces as well as license plates, for preserving privacy of people.

Each vehicle acts as a mobile sensor and periodically relays risky situations information to a global database (server), where statistical studies and data-mining algorithms aimed at extracting hotspots and locating them

in a map are performed. Figure 1 shows the architecture of the global database exploitation architecture. Currently, the updating process is performed after each urban journey in an off-line mode.

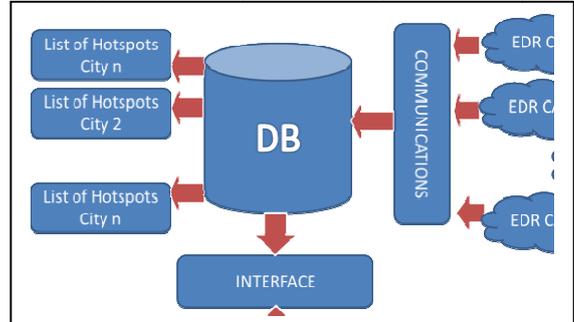


Figure 1. Global Repository Architecture

IV. DECISION ENGINE

The decision engine within the framework is an Expert System that analyzes inputs in real time and decides if current environment conditions can be a risky situation. System also has a software module for deciding what kinds of actions (braking, steering, issuing an alarm, etc.) to be undertaken by the pre-collision system in order to avoid collision (if necessary). It is also responsible for recording data of the corresponding risky situations in the local EDR when they are detected.

The decision engine is designed as a deterministic finite state machine (DFSM), where its behavioral model goes through the different states according to the accomplishment of certain rules driven by a set of inputs (vehicle speed, distance to the nearest obstacles, availability of side space, etc). Transitions between different states are accompanied by triggering actions such as breaking, steering, or simply issuing an alarm as a response to the incoming scenarios. Currently, system's finite state machine is with 4 inputs, 4 outputs, and 6 states. DFSM has been created using knowledge extracted from human experts in traffic and car accidents analysis.

V. DATA MINING AND KNOWLEDGE DISCOVERY

In this section we will comment the process for extracting knowledge from incident data gathered from different vehicles over the last year. As seen in previous sections, the mining process is aimed at outputting a list of hotspots representing areas with risky incidents. In order to make output visually attractive, hotspots are represented as histogram layers on a map, and are color-coded in order to highlight their locations.

A. Pre-processing data

Before applying statistical methods, recorded data are pre-processed for removing inconsistencies such as data with inconsistent GPS location, as well as removing irrelevant features like labeling. Afterwards, an engineering process is conducted in order to transform attributes into more meaningful ones. For example, GPS

date and time attributes are transformed into discrete values as shown in Table I. Therefore, a normalization process is then performed applying a linear scaling transformation.

TABLE I. ATTRIBUTES MAPPINGS

	Date and Time Transformation		
	New Attribute	Type	Value
GPS Date	weekday	Nominal	0..6
	season	Nominal	0..3
GPS Time	quarter	Nominal	0..3

B. Incidents Hotspot mapping

The starting point for any traffic hotspot map is a city map showing the locations of interesting incidents, represented as discrete points. Each point hides a collection of attributes, which can be displayed by clicking on a selected point. However, points in a map still need some visual facilities for good interpretation and make data understandable.

What is really needed is how to visualize and show the density of frequent incident sets of points. Hotspot mapping is dealing with probability surfaces for geographical data. In other words, estimating the probability on the number of incidents is different from the intensity of attribute values at a particular point. Another situation is that white spaces on the map do not mean absence of data for this area, and in the other side, hotspots can exist in point where no incidents have been collected.

VI. TOOL FOR VISUALIZING TRAFFIC INFORMATION

A. Graphical User Interface (GUI)

As part of the visualization tool [19][20] for the analysis and visualization of the traffic information, a graphical user interface has been built in Matlab (Figure 2). This interface allows the user to choose the type of information that he want to visualize. The user (e.g., city councils) can choose between different options: 1) by date, in that case the traffic information displayed corresponds to all routes recorded by the vehicles that are stored in the database for that date; 2) selecting vehicle-route-obstacle, in this case, the interface displays first the routes stored in the database which has been recorded by the selected vehicle, then the interface shows the obstacles (vehicle, pedestrian, non-high obstacle, high obstacle) present in the risk points corresponding to the selected route. After selecting the type of obstacle or after date selection, the interface shows a table with the incidence of each type of obstacle according with the previous choices. Finally, clicking on "DISPLAY MAP" button, a browser is opened with the traffic information visualized on a map.

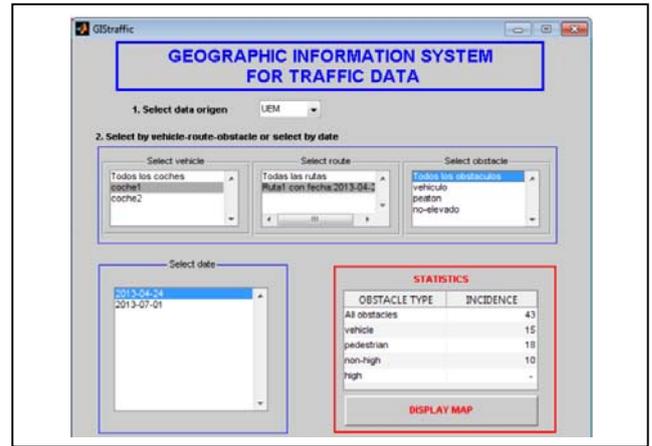


Figure 2. Geographic information System (GIS) for traffic data based on graphical user interface created with Matlab.

B. Traffic information mapping

The visualization of the incidents information is based on layer superposition on Google Maps (Google Maps Javascript API V3). The result is an interactive and easy to use map and quickly understandable visualization. First, the map displays the risk points like a heatmap, where the areas of higher density of risk incidence are colored red and areas of lower risk appears green. A top panel allows to change the radius and opacity and show/hide the heatmap (Figure 3).

A marker represents the obstacle types in each risk point by zooming the map. Moreover, clicking on the risk point brings up an information window with information about this point (image, vehicle velocity, type of obstacle, geolocalization coordinates, date and hour).

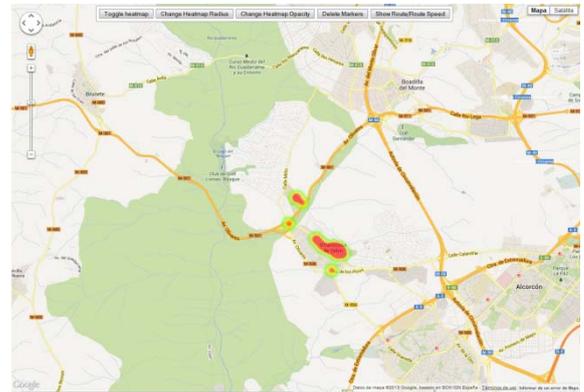


Figure 3. Heatmap of the risk points corresponding to the user selection in the graphical user interface.

The application allows also to visualize at the same time several routes recorded by the vehicle. Each route appears differently colored. If the user selects a single route, then in addition to the route is also displayed a graphic with the vehicle speed in each route point. Overing the cursor on the graphic display automatically a green marker on the corresponding route point, this moves according to the cursor movement.

C. Creating traffic information map

The use of matlab software and javascript language is combined for creating the incidents information map. Matlab allows the access to the database to select the data according the options chosen by the user in the interface. The main selected data are: (1) for each point of the route recorded by the vehicle, geographic coordinates of latitude and longitude and speed vehicle are selected, (2) for each risk point along the route, the latitude and longitude coordinates, vehicle speed, date, risk image and obstacles present in that risk point are selected. The data in the database are expressed in GPS units (GPS 18x, Garmin International, Inc.), so they are transformed before to be used. The latitude and longitude coordinates are transformed to decimal degree and the speed is transformed to Km/h. Once transformed the data, are expressed in javascript language and copied into an html file, which is written in matlab. This html file runs on a browser when the user presses the "DISPLAY MAP" button on the interface.

The first thing that shows the map is a heatmap layer. The heatmap is a visualization used to depict the intensity of data at geographical points. In this representation, every geographical point is represented with a circle with red core and green edge. In heatmap, red color represents high intensity of data and green color represents lower intensity of data. Initially (without zoom), heatmap covers a mayor region, however, incrementing zoom, colored areas goes in accordance with the localization of the risk points. When clicking on a risk point opens an infowindow with information about the point. However, clicking exactly in a risk point can be difficult so, for this reason, infowindows are also opened when clicking close to the risk points. In this case, infowindow corresponds to the risk point nearest to the click point. This distance is calculated as the Euclidian distance, $d(r,c)$, between the latitude and longitude coordinates of the risk and click points:

$$d(r, c) = \sqrt{(Lat_c - Lat_r)^2 + (Lon_c - Lon_r)^2} \quad (1)$$

where Lat_c , Lat_r , Lon_c , Lon_r are the latitude and longitude coordinates of the click and risk points, respectively.

VII. CONCLUSIONS

The main objective of this paper is to show statistical and data mining techniques for accident hotspots identification. We proposed a very effective method that can capture data from traffic states in urban areas and detect risky situations. By using vehicles equipped with sensorial systems and a decision engine. Those incidents are stored in on-board EDR, used to feed a central repository of critical incidents. Information of that global database is used to identify urban areas with a high density of incidents.

The use of Heatmaps over geographical maps allows us to show hotspots as histogram maps, using hot colors

for areas with higher amount of reported incidents. This kind of visualization of hotspots helps can be exploited by city councils, allowing them to address issues in those areas of high concentration of risk situations.

ACKNOWLEDGEMENT

This work was supported in part by the Regional Government of Madrid under the S2009/DPI-1509-SEGVAUTO grant, and the PN I+D+i under the TRA2010-20225-C03-02SAMPLER grant

REFERENCES

- [1] Bahlmann C., Y. Zhu, V. Armes, M. Pellkofer and T. Koehler: "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information". IEEE Intelligent Vehicles Symposium, vol. 6, no. 8, pp.255-260, 2005.
 - [2] McCall, J.C and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation", IEEE Intelligent Transportation Systems, Vol.7, no.1, pp. 20-37, 2006.
 - [3] Yanchao Dong; Zhencheng Hu; Uchimura, K.; Murayama, N.; "Driver Inattention Monitoring System for Intelligent Vehicles: A REVIEW", IEEE Intelligent Transportation Systems, vol.12, no 2, pp. 596 – 614, 2011.
 - [4] Eriksson M., and P. N. Papanikolopoulos, "Driver fatigue: a vision-based approach to automatic diagnosis", Transportation Research Part C: Emerging Technologies, Vol. 9, no.6, pp.399-413, 2001.
 - [5] Huijing Zhao; Shibasaki, R., "A novel system for tracking pedestrians using multiple single-row laser-range scanners," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on , vol.35, no.2, pp.283,291, 2005.
 - [6] Broggi, A.; Cerri, P.; Ghidoni, S.; Grisleri, P.; Ho Gi Jung, "A New Approach to Urban Pedestrian Detection for Automatic Braking," Intelligent Transportation Systems, IEEE Transactions on , vol.10, no.4, pp.594,605. 2009
 - [7] Felipe Jiménez, José Eugenio Naranjo&Óscar Gómez, "Autonomous Manoeuvring Systems for Collision Avoidance on Single Carriageway Roads", Sensors, vol 12, pp. 16498-16521, 2011.
 - [8] Keller, C.G.; Thao Dang; Fritz, H.; Joos, A.; Rabe, C.; Gavrilu, D.M., "Active Pedestrian Safety by Automatic Braking and Evasive Steering," Intelligent Transportation Systems, IEEE Transactions on , vol.12, no.4, pp.1292,1304, 2011
 - [9] Thao Dang; Jens Desens, Uwe Franke et al, "Steering and Evasion Assist", Handbook of Intelligent Vehicles, Chapter-29, Springer-Verlag, London 2012
 - [10] Kwok-Suen Ng, Wing-tat Hung, Wing-gun Wong, "An algorithm for assessing the risk of traffic accident", Journal of Safety Research, Vol. 33, Nº 3, pp. 387-410, 2002.
 - [11] Saffet Erdogan, Ibrahim Yilmaz, Tamer Baybura, Mevlut Gullu, "Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar", Accident Analysis & Prevention, Vol. 40, Nº.1 , pp. 174-181, (2008).
 - [12] Omer M. Soysal, Helmut Schneider, Asim Shrestha, et al. (2012). "Zonal Statistics to Identify Hot-regions of Traffic Accidents", In 9th Int. Conf. on Modeling, Simulation and Visualization Methods, WORLDCOMP'12. (2012).
 - [13] Bolstad, P. (2005) GIS Fundamentals: A first text on Geographic Information Systems, Second Edition. White Bear Lake, MN: Eider Press, 543 pp.
- Greiner, Lynn. Putting Canada on the map: Father of digitized mapping recounts how a stroke of luck led him to develop the world's first geographic information system», The Globe and Mail, 17-12-20