VIDEO BASED ON-ROAD EVENT RECOGNITION TECHNOLOGIES

Urban Burnik, Jurij F. Tasič
University of Ljubljana, Faculty of Electrical Engineering

Abstract: *In recent years, on-road object and event based video analysis has attracted special attention. For the identified objects and events the semantic representations can be provided. In-time recognition and proper response to objects, especially moving ones such as pedestrians, animals, random obstacles and other vehicles etc., play an important role in traffic safety. This paper provides a representative overview of progress in object and event based traffic scene (video) analysis. It specifically focuses on moving object detection and recognition including face, body, vehicle, event analysis of video and abnormal events detection in traffic. It also gives an insight in future vision-based traffic systems exploiting extensive broadband intercommunication among active subjects in road traffic.*

Keywords: *Road vision, traffic assistance, image processing*

1. Introduction

Up until recently, the way of driving hasn't change significantly since the invention of a car. Drivers were supposed to be independent subjects in traffic; neither they nor the vehicles have been interconnected in a system. All the signals, directions, including situation on the road, were communicated to a driver by means of his eyes, and the drivers responded to a current situation as good as they have managed to, subject to many objective and subjective factors.

In recent years, the situation is being changed dramatically, and is continuing to evolve to an even greater extend, in which vehicles are expected to be interconnected, actively communicating with their surroundings. The driving of the future is supposed to bring less effort to the driver, more comfort to passengers and better management of traffic to the society.

Considering the level of automated driving assistance systems one shall not expect a sudden change in traffic technology, leading to an immediate introduction of autonomous driven vehicles as at current stage collisions with other non-compatible traffic subjects would probably be unavoidable. Realistically, the progress in this area is passing more generations, which basically provide functions of informing (GPS navigation, traffic warning systems, road sign reporting), supporting with correction facilities (lane departure warnings, break assistance, active cruise control), single lane auto-piloting (autonomous driving in a controlled environment), cooperative driving (flexible platooning of interconnected vehicles) to final, fully autonomous driverless vehicles.

Nowadays we are facing dramatic improvement using a range of sensors and actuators providing certain level of assistance to the driver in a form of commercially available solutions. These include (1) satellite navigation techniques complemented by radiofrequency broadcasting providing navigation, traffic condition information and even road sign information based on preloaded roadmap database, (2) sensor and actuator based corrective mechanisms compensating drivers' misbehaviour (ABS, ASR, lane departure detectors,...), (3) active and passive speed control systems (radar-based cruise control systems, break assistants,...) to name just a few.

As the mentioned assistance technologies use several different sensing systems visual information itself retains its importance in road traffic. As already mentioned, visual information traditionally provides the driver with all the traffic information including traffic conditions, road signs, surrounding traffic subjects etc required for safe driving. Therefore it is expected that any intelligent vehicle system should be based on or at least complemented by some form of video based scene analysis. These automotive vision systems may further benefit from interconnectivity among neighbouring subjects, another feature to be exposed in future intelligent transport systems. Exchanging information when necessary should not only be used to inform subjects on the surrounding actions but also to learn how to respond in critical situations in order to improve response in safety critical situations.

2. State of the art in road vision systems

One of the most important issues in safer car development is to improve the drivers' abilities to react in critical driving situations. As already mentioned, the complex goal might be more easily achieved by employing advanced automated vision methods and techniques. It is obvious that the quality of visual information gets seriously affected by weather conditions, time of the day and by the presence of obstructive objects. Furthermore, human ability to process the incoming information varies and the probability of misinterpretation of a current situation is not to be neglected.

Generally, visibility plays an important role in the following types of traffic accidents:

- vehicles that collide with slowly moving or stationary traffic subjects (vehicles, pedestrians), especially in bad visual conditions,
- angled or head-on collisions,
- rear-end collisions that occur in poor weather conditions.

Poor visibility of pedestrians and vehicles represents one of the most serious risk hazards. The mix of motorized and non-motorized traffic, together with poor street lighting, increases the risk of unprotected road users not being seen. European research found that one third of pedestrian casualties had difficulty seeing the vehicle that had struck them, while two fifths of drivers had difficulty seeing the pedestrian. Besides, the pedestrian is hard to detect as objects off the direct line-of-sight are hard to recognize. A better chance for the driver to see the pedestrian occurs when the car is still far away. As the car approaches, the pedestrian recedes further into the driver's peripheral vision. At larger distances, the pedestrian is closer to the driver's line of sight and is more detectable.

Ambient vision is used for determining location in space and orientation in the environment and to perform tasks such as steering a car. It operates out in the visual periphery and needs only detect faint large shapes. Drivers can steer the car just as well at night as during the day and feel little need to slow. They do not realize that their ability to recognize objects on the road has been greatly reduced. Driver age is another important handicap. A group of older drivers detected the pedestrians at only 60% of the distance of a younger group. Older drivers travelling 80 km/h would hit almost all dark-clad pedestrians.

These critical traffic situations, safety may significantly improve by enhancing visibility, for example by using digital cameras to reduce dead-angle situations, enhance night vision or to make a wide-angle traffic overview by merging data from several imaging sensors. Additionally, some more advanced computer vision based methods may analyze the presence of objects within the scene; detect traffic situations and alert driver accordingly. The driver's visibility of the objects on the road could be improved by automated vision and object recognition system. Such systems improve visibility at night and increase the driving security and comfort.

It is also known that drivers' detection ability is impaired by at least 0.5 s in terms of brake reaction time when they talk over a mobile phone [1] or even of a couple of seconds when they are tired and sleepy [2]. There are numerous project studies addressing the critical safety issue of drowsiness while driving, which is known to cause accidents with most critical outcomes. The methods in these areas vary from using vital functions sensors (e.g. heart rate monitoring) however at this point we may emphasize on both passive (lane detection warning systems) and active (face recognition, driver behaviour, eye movement analysis) vision-based systems [7][8].

The systems utilizing vision enhancement, automated processing of traffic conditions, supported by cameras, semantic recognition of objects, subjects and events in traffic can further benefit from additional modality sensors. It is foreseen that advanced wireless communication systems, enabling interchange of information among active traffic subjects will bring further advances in traffic safety and may represent the most important step towards full automated road traffic systems.

2.1 Commercial examples in road vision systems

Head-up display systems

There is a problem of safe and ergonomically suited displaying of driver-assistance information, like current speed and status of vital automotive parts, but recently also traffic navigation instructions or processed results provided by advanced driving-assistance systems. Traditionally, these data are being displayed far from a presumed line-of sight of a driver. Modern automotive suppliers are offering head-up displays as a convenient alternative. BMW, for example is offering one of the first head-up display systems as a marketable driver assistance system in the new 5 series [9]. It is claimed to offer a significantly improved ergonomics and driver safety. This system projects important information such as speed, navigation guidance, and status indicators directly onto the windshield. The obvious benefit is that drivers can keep their eyes on the road and react more quickly to unexpected situations. According to BMW, head-up displays reduce the time needed to absorb information by half, compared with conventional instruments, representing a further safety bonus for drivers.

Figure 1: BMW's head-up display in 5' series.

Vision enhancements

Many research centres are offering driver-supporting vision systems that operate at different levels, like vision enhancement, scene analysis and display of supportive information. The systems are being developed to support car driver's vision, especially in conditions of bad visibility and to warn the driver on objects on the road or objects approaching the road. In heavy rain and during the night it is often difficult to recognize objects and to optimally see the drive path ahead therefore such systems are of a great support for the driver. Some examples of these are listed below.
To relieve the driver during dusk and in the dark and to significantly increase his driving safety and comfort, several companies are developing advanced driver assistance systems based in infrared vision. The so-called "night vision" systems are offering infrared-based imaging to increase visibility in the dark and to highlight lane barriers and pedestrians.

Figure 2: Siemens VDO's Night Vision.

Some examples include, but are not limited to BMW [9], and Siemens VDO [11].

Road sign recognition and management

Drivers know the problem of identifying state of traffic rules imposed by the road signs, especially current speed limit. This question can lead to insecurity especially on unknown roads. Even a town sign indicating the entering of a town and which also represents a speed limit can be quickly overlooked. While the task is nowadays easily managed by advanced GPS navigational systems a portable database of traffic speed limits may not always be up-to-date. In order to have a relaxing trip at any time, it is good to know the speed limit for the road one takes, and image recognition based traffic signs recognition and management system may be the most reliable idea to perform the task. The "Traffic Sign Recognition" system offered by Siemens VDO [14] and University of Koblenz [15] provides the driver with such information. These systems may display the speed limit permanently while not interfering with the driving behaviour. Moreover, the driver can reset the speed display at any time if he no longer needs it.

Critical event recognition systems

Also the information about the lane environment is necessary for the safe driving. On the market, there exist more or less sophisticated vision-based techniques that observe the driving environment and take actions accordingly. Such systems use information from several environment sensors and enable situation-based driving tips for more comfort and safety. Such advanced driver assistance system, can significantly increase safety and comfort in majority of situations. It gives recommendations only when the situation is appropriate. For the example of the lane change it means that such a system warns of a lane change as long as there is a car in the blind spot. Some less sophisticated systems may reduce the warning to periods when line markings are being crossed without using indicators. Such a system is being advertised as Sensitive Guidance by Siemens VDO [12]. A less-sophisticated vision system is offered by Citroen as "Lane Departure Warning System - LDWS" [13].

Figure 3: Siemens VDO's Sensitive Guidance.

A technology provider Mobileye Vision Technologies from Israel [16] acts a supplier for BMW, Volvo and GM, as well as a manufacturer of retrofittable consumer products offering advanced image-processing based critical event recognition systems. The products are based on SoC automated vision systems of their own brand and are complemented in OEM products with additional sensors such as radar systems, accelerometers etc to ensure more reliable operation. The functions cover a wide range of event detections including lane departure warning, forward collision warning and headway monitoring, with more sophisticated upgrade capabilities like pedestrian detection and intelligent event video recording.

Figure 4: Mobileye Vision System in Volvo cars.

3. Object recognition in traffic

The goal of object detection on roads is to locate objects in a scene. There are several important questions that relate to this task. What is defined as an object? What needs to be done to recognize something as an object? The ultimate goal of computer vision is to design a system that would be capable of analyzing a scene and determining which items in a scene were relevant objects. Realistically, the designer of the system has to define what an object is. Whether that object is defined to be a car, a boat, or a face must be defined by the system designer. In order to find these objects the system must apply some techniques which are designed to look for particular objects. Because these techniques general have to be applied one at a time the computational power of the system can have a large impact on how many different types of objects can be detected.

Edge detection is one of the most commonly used operations in image analysis. The reason is that edges form the outline of an object. An edge is the boundary between an object and the background, and indicates the boundary between overlapping objects. Edge detection represents the first part of a process called segmentation - the identification of regions within an image. Technically, edge detection is the process of locating the edge pixels, and edge enhancement will increase the contrast between the edges and the background so that the edges become more visible. The significance of edges comes from practical, biological points of view. Compared to basic image elements - pixels, edges are much more stable, less susceptible to noise and changes in lighting conditions. This makes them good candidates for image analysis and representation.

Let us assume that the object that we want to recognize is a car and that it is represented as a collection of edges of different sizes and orientation arranged at specific locations with respect to each other. After reliably extraction of edges grouping process of the extracted edges must be started such that they represent a car. The image of the car can be used as a starting point and where one of the gradient operators is used to identify the edges of the car. The points above a specific threshold value were taken as edge points and then edge sizes and the locations of their centres were estimated. The same process can be repeated for the different edges thresholds. All so identified edges are displayed over a difference image for the purpose of clarity, although the edges themselves were extracted not from the difference image but from the original image.

The basic idea of deformable contour technique is that there exists a contour model or template of the object which we wish to detect. In the picture below it is the contour of a car on the lane. This closed contour template defines a series of points. The set of points in this template is compared to other closed contours inside the picture. If these contours are inside off predefined and acceptable deformation of the template contour then there is a match. The pictures below represent some examples of this algorithm working and the contour that it found acceptable. In the second and in fourth image, it is evident that the algorithm has problems when the background in the scene is of similar intensity of the object itself.

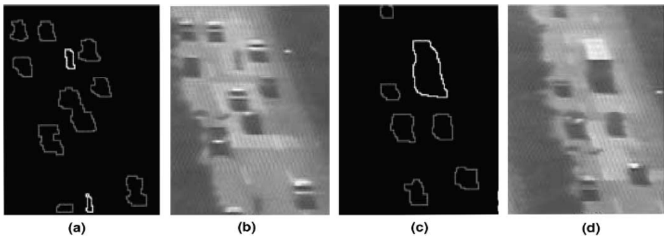

Figure 5: Deformable contour technique

Better results can be achieved by semi-automatic methods, where a user controls the grouping process of edges and provides initial object segmentation. The initial segmentation is tracked through image sequence until major object distortions occur. The proposed object extraction scheme uses initial object outline provided by the user and performs object extraction by tracking and approximating its edges. On the basis of the given object outline an initial object shape model is generated using B-spline approximation scheme. An example of the shape model is given in Figure 5.

A colour model of the video object is generated by averaging local colour of the object along the object's boundary separately for hue, saturation and intensity components of the image. Object boundary estimation is based on image edges detected in the neighbourhood of the compensated shape model. Because we want to use all available information present in colour images, intensity edges and edges of two colour components of the image are combined into a single edge map.

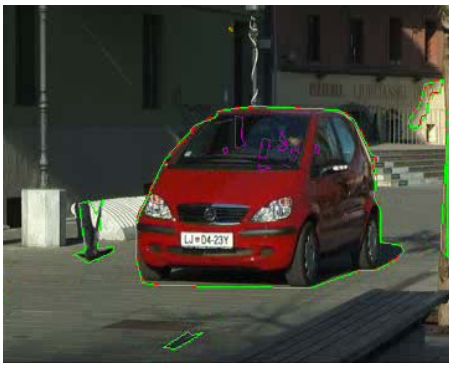

Figure6: The initial shape of a car.

Selection of the object boundary from the edge map is based on a colour model of the object. For each edge point in the neighbourhood of the compensated shape model average colour components (hue, saturation, intensity) on the line perpendicular to B-spline shape model are calculated. Colour similarity measure in HSV colour space is used for selecting edge points, which will constitute object boundary. The selected boundary points are then filtered to remove large random deviations from the initial shape model. Having determined object boundary in the current frame, the initial B-spline shape model is adjusted by performing optimization procedure and knot insertion if necessary.

Finally the colour model is adjusted by averaging object colour in the neighbourhood of object boundary over m latter frames. The choice of the number of frames m is important because it represents the trade-off between flexibility of the model and its robustness. Larger m increases robustness to small occlusions and false edge selection, but it is not suitable for objects, which change quickly. The colour model ensures temporal coherence of the extracted video object.

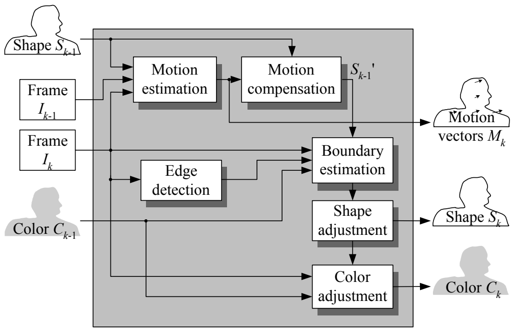

Figure 7: Block diagram of the object tracking algorithm.

As it was shown the approach to the object model applied to recognition process and to detection of traffic situation from real-time video streams is important task for drivers alert. Model or an object is represented as a collection of features of specific classes. In the car detection process the features can be the edges of different orientations and sizes. We further assume that this collection of features cannot be made on any point in the image but only on edges, more specifically on their centres. A presence of other edges of an object is used to put an edge in context, and we want to know how to measure their influence. Due to the fact that camera can acquire the images of the object from different viewpoints, shapes and sizes, the locations of objects' edges are not fixed with respect to the location of a chosen central edge. In such case a given central edge location is arbitrary, but together with other neighbourhood edges is located inside region of interest.

4. Future steps in on-road vision systems

In Laboratory for digital signal, image and video processing, we dedicate our research to both fundamental research in image and video processing as well as to intelligent, self configurable systems. As members of Telecommunication Department we have always been devoted to implement the advanced knowledge in modern communication services. Communication technology itself is breaking the barriers in wireless communication capabilities, and the technologies like ad-hoc sensor networks, serverless peer-to-peer communication networks, combined with the true power of wireless broadband offered in the frequency bands recently established at frequencies above 5GHz are offering a new dimension to road traffic vision-based safety and management systems.

By the time we are witnessing the vision-based drivers' assistance systems entering the automotive market not only in a form of passive driving aids (reverse cameras, cruise controls, head-on displays, lane departure detectors, night vision), but are getting active role in driving (active cruise control with distance measurements, collision avoidance systems with automated braking,...). We are now reaching the point at which traffic subjects should not be seen as independent individuals but as a whole traffic body. In new road traffic automation, to get to a new step active communication and data interchange based techniques such as multiagent technology should combine the knowledge gathered by sensors located in individual vehicles. Using broadband mesh interconnections in high frequency bands (e.g. 60GHz unlicensed spectrum) links among traffic participants could be established which would provide enough bandwidth at a low latency to combine and to process vision information gathered by individual vehicles as an integrated traffic situation.

Multiagent systems have in recent years become a popular and promising technology, first of all because of a new paradigm for conceptualizing. This promise is particularly attractive for their application in distributed, connected and open environment, such as the Internet. Connected vehicles have become interesting because of their integration into global network of traffic subjects. This is also the way to integrate new safety, ecological and comfort functionalities, to make them more attractive to the user. We are discussing, in particular, LORA (Logic Of Rational Agents) as multimodal logic, which presents a logical framework to represent the properties of rational agents and reasons about them in an unambiguous, well-defined way [21].

Such systems should not only be able to forward current traffic conditions among the following vehicles and interchange critical safety information in real time. In the next phase, the integrated mesh of independent vehicles could learn from previous experience and force safe response behaviour in case similar critical conditions occur.

5. Conclusions

In the paper we present trends as well as some commercially available on-road vision systems, which extend from simple drivers' assistance like enhancement of vision, coverage of dead angles all the way to active devices recognising traffic conditions, traffic subjects (pedestrians, animals, vehicles, road signs) empowered not only to inform but also to react in critical conditions. The intelligent transport technologies are now reaching a step further towards driverless vehicles. As first attempts towards this goal have been registered in commercial transport we may soon reach the era of fully automated traffic, which will be based on automated vision technologies combined with powerful interconnection networks, supported by intelligent agent technologies.

References

[1] Green, M. (2001b). Do Mobile phones pose an unacceptable risk? A Complete look at the adequacy of the evidence? Risk Management, November, 40-48.

[2] Green, M. (2000). 'How long does it take to stop?' Analysis of brake reaction times," Transportation Human Factors, 2, 195-216.

[3] Green, M. (2001a). The psychology of warnings. Occupational Health and Safety Canada, 30-38, October/November, 30-38.

[4] Crawford, B. (1947). Visual adaptation in relation to brief conditioning stimuli, Proceedings of the Royal Society of London B, 134, 283-302.

[5] Green, M. (1981). Spatial Frequency Effects in Masking by Light, Vision Research, 21, 801-806.

[6] Green, M., 1984, Masking by Light and the Sustained Transient Dichotomy. Perception and Psychophysics, 34, 617-635. IESNA, 1983, Recommended Practice For the Calculation of Daylight Availability.

[7] Parmar, N. Drowsy driver detection system, design project, Dept. Of Electrical and Computer Engineering, Ryerson University, 2002.

[8] Esra Vural, Mujdat Cetin, Aytul Ercil, Gwen Littlewort, Marian Bartlett and Javier Movellan, Intelligent Monitoring System for Driver's Alertness (A Vision

Based Approach), in Knowledge-Based Intelligent Information and Engineering Systems, Pages 471-477, Springer Berlin / Heidelberg, 2008, ISBN 978-3-540-74817-5
[9] http://www.bmw.de/de/produkte/automobiles/5er/limousine/index.html?content=http://www.bmw.com/generic/de/de/products/automobiles/5series/sedan/ergonomics_01.html, accessed Nov. 2008.
[10] http://www.carpages.co.uk/bmw/bmw-night-vision-22-07-05.asp, accessed Nov. 2008.
[11] http://www.usa.siemensvdo.com/media/releases/2006/ii/0118.htm, accessed Nov. 2008.
[12] http://www.siemensvdo.com/topics/adas/sensitive-guidance, accessed Nov. 2008.
[13] http://www.citroen.com/CWW/en-US/TECHNOLOGIES/ SECURITY/AFIL, accessed Nov. 2008.
[14] http://www.siemensvdo.com/topics/adas/traffic-sign-recognition/, accessed Nov. 2008.
[15] http://www.uni-koblenz.de/~lb/lb_research/ research.tsr.html, accessed Nov. 2008.
[16] http://www.mobileye.com, accessed Nov. 2008.
[17] Object Recognition in Image Processing, Website Authored by: Jared Knutzon and Bryan Walter
[18] Amit, Yali; 2D Object Detection and Recognition Models, Algorithms, and Networks; The MIT Press, Cambridge Massachusetts, 2002
[19] Ritter, Gerhard & Wilson, Joseph; Computer Vision Algorithms in Image Algebra; CRC Press LLC, Bocca Ratton Florida, 2001
[20] Grigorescu, C.; Ptekov, N.; "Distance Sets for Shape Filters and Shape Recognition," IEEE Transactions on Image Processing, vol. 12, no. 10, pp. 1274-1286, October 2003.
[21] Steblovnik, K. Model večagentne dinamične strukture kot zasnova razumnega hišnega pomočnika : PhD Dissertation, Maribor, 2007.

Sadržaj: *Video analiza objekata i događaja u drumskom saobraćaju je postala veoma značajna tema poslednjih godina. Pri tome se koristi semantička analiza prepoznatih objekata i događaja. Blagovremena identifikacija i reakcija na objekte u pokretu kao što su pešaci, životinje, slučajne prepreke i ostala vozila, od velikog je značaja za bezbednost u saobraćaju. Ovaj rad pruža reprezentativni pregled razvoja u prepoznavanju objekata i događaja u saobraćaju pomoću analize video signala. Sa posebnom pažnjom razmotreno je prepoznavanje objekata u pokretu kao što su lice, telo, vozilo, promene u saobraćaju i detekcija neočekivanih događaja.U radu je dat i predlog kako će se u budućnosti koristiti direktna širokopojasna komunikacija između aktivnih korisnika u drumskom saobraćaju, u svrhu povećane bezbednosti.*

Ključne reči: *drumski saobraćaj, asistencija u saobraćaju, obrada slike*

TEHNOLOGIJE ZA PREPOZNAVANJE DRUMSKIH DOGAĐAJA NA BAZI VIDEA

Urban Burnik, Jurij F. Tasič