Determining Car Driver Interaction Intent through Analysis of Behavior Patterns

Madalina-Ioana Toma, Dragos Datcu

Determining Car Driver Interaction Intent through Analysis of Behavior Patterns

Madalina-Ioana Toma[1] and Dragos Datcu[2]

[1]Transilvania University of Brasov, Product Design and Robotics Department, Romania
[2]Netherlands Defence Academy, Faculty of Military Sciences, The Netherlands
madalina-ioana.toma@unitbv.ro, D.Datcu@nlda.nl

Abstract. The aim of the article is to present preliminary results obtained by analysis of the behavior patterns of various driver subjects, in the context of an intelligent assistive driving system. We determined the parameters which are involved in determining the car driver's interaction intent, and extracted features of interest from various measured parameters of the driver, car, and the environment. We discuss how threshold values can be obtained for the extracted features that can be part of rules to decide on specific interaction intents. The results obtained in this paper will be incorporated in a knowledge base to define the rules of an rule-based expert system that will predict in real-time the driver's interaction intent, in order to enhance the safe driving experience.

Keywords: natural interaction, intention intent, virtual environments, features extraction.

1 Introduction

The skill to drive has become more and more important in the human life since the invention of the car. It is desirable for future intelligent cars to have the capability of foresighted driving, which would ensure that its driver's actions are performed in a safe manner. In this paper, we address the first step of an on-going research project which brings into attention the building of an intelligent co-driver system that recognizes in real time the intention a car driver has, based on the actions of the driver, the state of the car, and information from the environment, and provide proper warnings and alarms to him/her. Moreover, it needs to correct potential mistakes of the driver and to predict the next action for an enhanced safe driving experience. We try to answer to the first question of our project: how can one determine the car driver interaction intent in real time? Answering this question represents a foundation brick of our intelligent real-time rule-based expert system [1] which we want to build. The automatic assessment of driver's intent involves computing the threshold values [2] of various parameters which can say when an event has occurred.

In order to build an intelligent system, we need to analyze the interaction between the following three entities: the driver, the car and the surrounding environment, during the driving activity. In particular, the interactions between each two entities have to be considered. Considerable research that analyzed the driver interaction intent based on a single entity were performed [3], [4]. Prior research tried to determine driver's interaction intent by analyzing only single body parts, e.g. pose estimation [5], gaze detection, or facial expression [6]. Compared with these, we take into account more aspects and actions, such as body postures, gaze direction, and head orientation during interaction with the car and the surrounding environment.

To analyze the driver action, we need use the sensors called smart sensors fusion [7], which track in the transparent way the driver behavior without obstructing activity. A smart sensor is one which consists of a sensing element, a signal processor, and a microprocessor all coupled into a single system and this kind of sensor allows preprocessing acquired data. Smart sensors fusion based on ubiquitous computing can be used to build a robust tracker and for automatic monitoring, allowing a natural interaction between the driver, the car and the surrounding environment [7]. These kinds of sensors fit very well with Virtual Reality Environments (VREs) [8] which simulates a surrounding interaction, so that the sensors and the interaction environment make a ubiquitous interaction space (u-space). In VREs, the driver is immersed in an audio-visually coupled tele-operated environment, where direct interaction and control are achieved in real time [9]. In our system, the behavior of the driver can be assessed by processing data from a KINECT-sensor, which is a smart sensor [10], a system based on smart sensor for gaze-tracker and a camera capturing the head and the body of the car driver. TORCS, a 3D open source car racing simulator is used as the virtual driving environment [11].

The aim of the article is to present preliminary results obtained by analysis of the driver action from tracking behavior of various driver subjects, in order to determine threshold parameter which could denote possible driver interaction intent, in order to build a rule-based expert system doing in real time recognition intention. The driver behavior patterns based on driver action were observed by segmenting collected data during three scenarios: in the first scenario the driver needs to drive with a low speed, in the second one he/she needs to drive and make maneuvers with a high speed and in the last scenario the driver drives normally without constrains from traffic signs.

The paper is organized as follows: Section 2 gives an overview of our solution as a contribution to value creation. Section 3 details the proposed system and methodology used to record drivers' activities and describes the interaction test scenarios. Experimental results are shown in Section 4, and the paper is concluded in Section 5.

2 Contribution to Value Creation

Automotive industry incorporates smart sensors driver interfaces for increasing the safety on board. At the moment, commercial solutions with implemented face analysis are already available as built-in features in cars for assisting drivers in traffic by warning and triggering their attention for obstacles on the road. The intelligent system of our on-going research project is a co-driver system which aims to predict in real-time the driver's interaction intent during driving, based on his/hers movement actions and reaction time for collision avoidance.

In order to design an intelligent system to recognize the car driver intent interaction we need to determine the threshold values of different parameters which control the decision analysis of driver behaviours patterns. Determining an accurate as possible threshold value increases the chance to correctly recognize the driver intent interaction, a crucial feature for a future driving assisting system which has the ability to automatically asses driver’s intent interaction based on driver's actions and reactions in different situations. Successfully achieving this goal by means of automatic systems is equivalent to a tremendous decrease of the economic and social costs implied by the loss of human lives and property. From this reason our system is a promising technology innovation for value creation.

Based on the results presented in this article we will build the rules database of an expert system that can identify in real-time the car driver’s interaction intent. The preliminary test results on one hand, and the low cost proposed system that records the driver action on the other hand, work together to create added value in the design of an intelligent system for future intelligent cars, which recognizes the car driver intent interaction.

3 Proposed System and Methodology

In our paper, the driver intent interaction was determined in laboratory conditions. To build an intelligent system that recognise in real-time the car driver interaction intent for collision avoidance the following steps are necessary to be performed: (1) track the driver’s behaviour and record the collected data; (2) analyse and compute the threshold constrains which code a possible driver interaction event; (3) based on the determined events the rule-based system decides if a driver interaction intent had occurred. Thus, we have built a system based on a set of smart sensors that track the driver’s movements, a driving virtual environment to simulate the driving context and a contextual driver interface consisting of pedals, a gear shifter and a driving wheel. The driver behaviours are tracked using different smart sensors and vision sensors, connected to the TORCS simulator.

The TORCS simulator is one of the most popular 3D open source car simulators [11], written in C++ and available under GPL license. We used TORCS for our purposes because it presents several advantages: (1) it is an advanced fully customizable simulation environment, giving us the possibility to adapt it for our application; (2) it features a sophisticated physics engine as well as a 3D graphics engine for the visualization of the virtual environment for interaction; (3) it has a modular software architecture, hence integrating new controlling and sensing devices is a relative straightforward task.

In order to make a complete analysis of the driver actions, we used several sensor devices to track the driver behaviors. The sensors are used to record the driver upper body posture position, for gaze and head tracking, and finally a set of sensors is used to interact with the virtual environment.

Simulation Apparatus - In our system a set of smart fusion sensors record the driver movements. The upper limb motion is tracked by the Kinect device, developed by Microsoft as a game console input device. This device based on smart sensors tracks the motion of the subject through a combination of hardware and software technologies and achieves a high accuracy tracking of the body [12] with a rate of 30

FPS (frames per seconds. We use in our work the 3D motion capture feature of the device depth sensor based on an infrared light ray which is scattered throughout the scene, with the reflected matrix pattern being captured by a CMOS sensor.

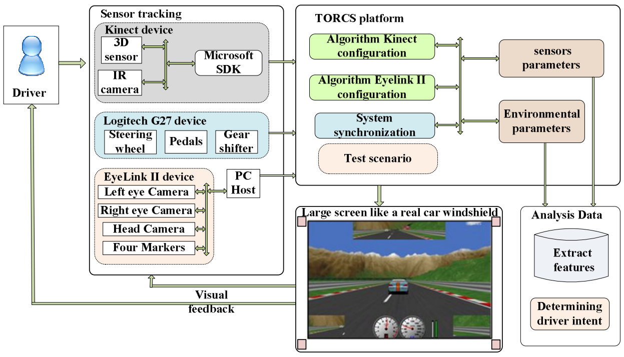

Fig. 1. System Architecture

The eyes and head driver action are recorded in our system by the SR Research EyeLink II device [13]. The EyeLink II system based on smart vision sensors consists of a head-mounted camera system and two PCs for processing data and running experiments. On the head-mounted device, both left and right eye pupil position and the head position relative to computer monitor were tracked. Combining the position of the head with the pupil movement relative to the screen, enabled recording of the gaze direction. Eye tracking data for both eyes was captured at a sample rate of 500Hz (2ms samples) with a resolution of 0.01° visual angle. The EyeLink II device needs a calibration before to be used. Thus, we implemented an algorithm in our application that each subject who involved to a test scenario need do a calibration before to start scenario.

The simulation is controlled via a Logitech G27 controller based on a steering wheel, a gear shifter and three pedals (clutch, break and throttle). The controller simulates a usual basic control system from a cockpit car.

The driving scenes of the scenarios are displayed on a large screen like a real car windshield. As a large screen, we used a TV screen with 56 inches size diameter, and in the four corners of the screen are placed those four optical sensors for head camera of the Eyelink II device.

In our system we implemented algorithms for configuration step by step for each used sensor, because we have access to the source code of the TORCS simulator. Also, we configured the Kinect sensors using Microsoft SDK [12] developed for this device, and the Eyelink II sensors using SR Research[11] that includes libraries that implement the link interface, and support code that makes programming simpler.

Participants - The data were recoded using our above proposed system from 24 (17 male and 7 female) voluntary driver people, who participated in our study. They have the ages 26 to 36 years (Mean age value 29.2), with 6-10 years of drive experience (Mean year value 8.7). They reported driving daily with a mean of 1800 km per month having no accidents.

Test Scenarios - All tests were performed in laboratory conditions in a light good condition, so from this reason we don't have problem with the vision sensors of the system. We designed three different scenarios, in order to track and observe the driver interaction intent through analysed the data recorded during the driver action in different situations.

Each scenario simulates a road shaped as a square, with a length of 4 km and a width of 10 m, with two traffic lanes. The traffic is restricted by traffic signs: in the first scenario the traffic is limited through speed limitations signs, while the second scenario the traffic is kept at a high speed and the third scenario were designed without traffic restriction signs. In all scenarios the other traffic on the road is simulated using 16 car robots from the TORCS simulator.

In the proposed scenarios the driver can be in the one of the next driving states: starting (S), keeping lane (KL), preparing left (PL), turning left (TL), speeding-up (SU), speeding-down (SD), takeover (TO), switching-lane (SU), preparing right (PR), turning right (TR), forward parking (FP), driving back (DB), and stopping (SP) which represent the set of possible intentions. Each driving state is represented by the driver's behaviour pattern in that state, which combines a set of driving actions in order to perform the composite action given by the state name. The transition from one state to another is decided based on events extracted from the captured sensory data. A transition is done if some conditions are met, specifically when some parameters are above their threshold values. This represents the fact that interaction intent had occurred. Determining these thresholds involves a sequence of steps, i.e.,: (1), read sensors and environment parameters every 20 milliseconds during driving scenarios; (2) define the features that can be used to describe driving actions and/or activities; (3) extract the features from the recorded data for every driving scenarios; (4) determine the combined range and the threshold value for every feature from the three proposed scenarios.

The following parameters were recorded, with 20 ms rate, while the subjects drove in our test scenarios: time (t), distance of a car to the starting position (Δd_c), the rotation angle of the steering wheel (ang_{sw}), activate throttle, clutch, brake pedal, gear level, current speed of the car, head orientation, positions (x,y) of the eye on the screen, 3D joint positions (x,y,z) of each part of the upper body (left/right hand, left/right elbow, left/right shoulder), and each traffic car position relative to the starting position. The data files were segmented in order to analyse the driver behaviour step by step. We used in segmentation a constant time window equal with 10 s. In this window we analysed the recorded parameters and we extracted the features summarized in Table 1.

4 Results Discussion

To determine driver interaction intent, we extracted some specific features from the recorded data in the three different scenarios using a constant time window of 10s to segment the recorded data files. We have shown in Table 1 the main features determined in our system. In Table 2 and Fig. 2, we present a statistical evaluation of the obtained values of each feature, i.e., the mean (MD) and standard deviation (SD). In addition, the threshold range, given by the minimum and maximum hard limits values is shown.

Table 1. Feature description.

Features	Characterizations of features extraction (w=10s)
Eye position(x,y) on the screen and parameters	
$\sum p(x,y)$	Number of consecutive gaze focus points when the driver looks in the mirrors, during a given time duration
$\sum \Delta t_g$	The duration of looking time in a mirror
Δt_g	Time interval between two successive gazes
Z_{3x3}	Screen area where a driver focuses – we split the screen in a matrix of 3x3 zones and we take into account just the zone where there are mirrors, such as: z_{31}-rear left mirror, z_{33}-rear right mirror, z_{12}-rear mirror
Upper body position (x,y,z)	
ang_{tf}	The angle between torso and forearm for both female (F) and male (M)
ang_{af}	Angle between right arm and right forearm (female and male)
a_{rh}	Right hand acceleration
Environmental features - Car positions in the traffic	
d_{c-x}	Distance between user car and all other cars
N_c	Number of cars in traffic into a range of x = 50-100 m around the user car
$D_{<c-f\ car>}$	The distance between user car and the car in front
$D_{<c-f\ back>}$	The distance between user car and the car behind
v_c, a_c	Medium speed and acceleration of user car during a segmentation window
hv_{car}	Speed of the fastest car from the traffic
Features of joystick controller	
ang_{sw}	Steering wheel rotation angle
a_{sw}	Steering wheel acceleration computed like the second derivative of steering wheel rotational angle
g_l	Position of gear shift lever
Δp_a	Throttle pedal pressing duration before switching in upper gear
$\sum p_b$	Number of brake pedal presses before switching in lower gear

The visual sense has the most important role in driving a car; hence, it represents the first stage in analysing the driver behaviour. In our proposed scenario, we collected from Eyelink II device the gaze action for each driver. For each scenario the reaction time is different. In the normal speed scenario, the time to take a decision is much higher than in the high-speed scenario and almost equal with the one in the limited-speed scenario. By correlating the threshold parameters obtained from eye gaze with the head orientation, the accuracy of detecting a driver's interaction intent is increased. Moreover, we observed that head action in high-speed scenario is faster than normal and limited scenario, and also the angle orientation of head is greater in high-speed scenario.

The arm actions are another driver part body which we analysed. We focused on right hand motion driver action; because it is involved more in driving manoeuvres. Analysing the right hand posture we observed that a state transition decision depends on the arm angle postures and the time between two consecutive postures. Data recorded with the Kinect sensor give us the Cartesian coordinates (x,y,z) for each segment hand. From this reason, we built a generic algorithm based on forwarding kinematic to compute the angles which can take the hand for each posture. Correlating the hand posture with feature and parameters of the steering wheel and gear shifter we could decide what kind of events took place, e.g. steer left, switch to higher gear. In addition, in our preliminary results, we observed that in the high speed scenario the driver mostly keeps the right hand on the gear shifter, while in the normal and low speed scenario considerably more hand movements between the steering wheel and gear shifter were observed. Events containing feet actions are assessed by correlating the pressure value applied on the pedals with hand posture and gear shifter lever position.

Table 2. Extracted features for hand postures

Features	Right arm feature values			
	MD	**SD**	**min**	**max**
ang $ang_{tf}(°)F$	20.4	2.25	18.2	24.1
$ang_{tf}(°)M$	18.7	1.87	17.2	19.2
$ang_{af}(°)F(sw)$	115.1	1.26	98.2	124.1
$ang_{af}(°)M(sw)$	100.2	2.14	87.6	158.2
$ang_{af}(°)F(gs)$	168.7	2.58	158.3	180
$ang_{af}(°)M(gs)$	152.1	1.81	14.3	139.3

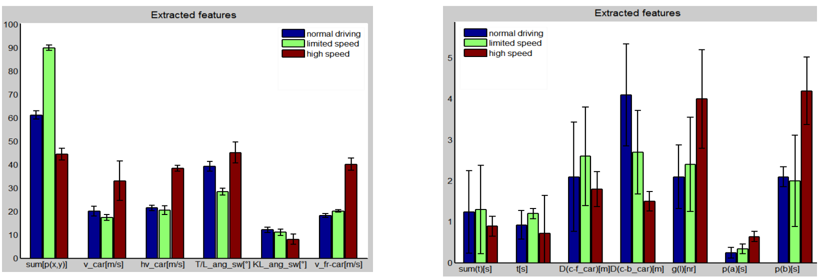

Fig. 2. Features extraction

The influence of the interaction environment can also determine the car driver interaction intent. Traffic signs and others cars in traffic represents the main factors which cause the driver to react. In our three scenarios, we recorded some parameters of other cars in the traffic such as speed, and distance to the subject car. If the driver action is not performed in the vicinity of an environment event, it means that the driver action is only a stimulus which will not determine intent for interaction.

5 Concluding Remarks and Future Works

The presented research from this paper represents early but promising results from our project to build an intelligent driving assisting system that recognizes in real time the intention a car driver has, based on the actions of the driver, the state of the car, and information from the environment. The aim of the present study was to determine the parameters which are involved in determining the car driver's interaction intent. We extracted features of interest from various measured parameters of the driver, car, and the environment. By analysing the results we discuss how threshold values can be obtained for the extracted features that can be part of rules to decide on specific interaction intents.

The preliminary results obtained in this paper will be incorporated in a knowledge base to define the rules of a rule-based expert system that will predict in real-time the driver's interaction intent, in order to enhance the safe driving experience. These results helps us to achieve a new level of understanding of car drivers interaction, and to further train a Bayesian network which can predict a future driver intention and/or action, based on the current recognized interaction intention. Once the expert system is completed and refined, we can apply the same methodology for obtaining features and threshold values into real driving scenarios.

Acknowledgment. This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/88/1.5/S/59321.

References

1. Laffey, T.J., Cox, P.A., Schmidt, J.A., Kao, S.M., Jackson, Y.: Read: Real-Time Knowledge-Based Systems, AI Magazine (1988)
2. Avci, E., Avci, D.: An expert system based on fuzzy entropy for automatic threshold selection in image processing, Journal of Expert Systems with Applications, 3077--3085, (2009)
3 Lefèvre, S., Ibanez-Guzman, J., Laugier, C.: Context-based Estimation of Driver Intent at Road Intersections 2011. IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (2011)
4. Doshi, A., Morris, B., Trivedi, M.: On-Road Prediction of Driver's Intent with Multimodal Sensory Cues, IEEE Pervasive Computing, vol. 10, no. 3, pp. 22--34, (2011)
5. Trivedi, M.M., et. al.: Posture analysis with stereo and thermal infrared video: Algorithms and experimental evaluation. IEEE Trans.Vehicular Technology, 53(6):1698--1712, (2004)
6. McCall, M.T.J., Mallick, S.: Real-time driver affect analysis and tele-viewing system. In Intelligent Vehicles Symposium, Proceedings. IEEE, pp. 372--377, (2003)
7. Inagaki, T.: Smart collaborations between humans and machines based on mutual understanding. Annual Reviews in Control, vol. 32, pp. 253--261, (2008)
8. Lorentzen,T., Kobayashi,Y., Ito,Y.: Virtual Reality Simulation: Integrating Infrastructure Plans, Traffic Models, and Driving Behaviors , Proceedings of the 16th World Congress on Intelligent Transportation Systems and Services, (2009)
9. Ingebretsen, M.: In the News," Intelligent Systems, IEEE , vol. 25, no. 4, pp. 4--8, (2010)
10. Shotton, J., Fitzgibbon, A., Cook, M., Blake, A.: Real-time Human Pose Recognition in Parts from Single Depth Images, (2011)
11. Loiacono, J., Togelius, et al.: The wcci 2008 simulated car racing competition, in Proceedings of the IEEE Symposium on Computational Intelligence and Games, (2008)
12. SR Research, Eyelink II - http://www.sr-research.com/EL_II_scam.html
13. Microsoft SDK for Kinect - http://kinectforwindows.org/