Assisted/Autonomous vs. Human Driving Assessment on the DiM Driving Simulator Using Objective/Subjective Characterization

A.Affanni[(1)], P.Zontone[(1)], R.Fenici[(2)], D.Brisinda [(2)], D. Bacchin[(3)],L. Gamberini[(3)], P.Pluchino[(3)], M.Bruschetta[(4)], C.Savorgnan[(5)], F.Formaggia[(5)], M.Minen[(5)], D.Minen[(5)]

[(1)]DPIA, University of Udine, Italy

[(2)] BACPIC, UCSC Roma, Italy

[(3)] HIT University Padova, Italy

[(4)] DEI University of Padova, Italy

[(5)] VI-grade

Abstract

More and more sophisticated assisted/autonomous vehicles are becoming available in the market. Automation levels 2 and 3 have been given for settled just a couple of years ago, and the path to fully autonomous car seemed to have no obstacles. In reality, OEMs started recently realizing that the impact of semi and/or fully robotized cars on drivers as well as on passengers is all but predictable. VI-grade has more than ten years experience with developing turn-key solution driving simulators, and has been working for more than five years on a research project to collect meaningful bio-signals from the driver during simulator sessions, in collaboration with the BACPIC of the Catholic University of Sacred Heart and the DPIA of the University of Udine. Recently, a collaboration with the Human Inspired Technology Research Center of University of Padova allowed the extension of the assessment at physio-emotional level.

Introduction

There is increasing interest from major OEMs for technologies providing real-time monitoring of the driver psycho-physiological reactions to vehicle dynamics under direct or robotized control: a) to study the robotic acceptance; b) to evaluate the difference between self and assisted/autonomous driving in identical scenarios; c) to evaluate the impact of on-the-fly setup variations of the vehicle when autonomously driven. The DiM150 driving simulator is ideal to replicate, in a fully immersive and deterministic environment, any kind of real-time vehicle dynamics under any kind of driving conditions. A novel wearable device to record bio-signals and a special software to analyse and monitor in real-time multiple physiological parameters have been developed. All the bio-signals are synchronized with vehicle’s ones. Driving-induced transient variations of autonomic nervous system (ASN) modulation derived from heart rate variability (HRV) and skin potential response (SPR) are collected letting the driver on the simulator drive through a series of tasks (separated by recovery intervals) on a variable highway scenario. Aim of this study was: a) to identify even minor human reactions differences induced by tasks during active and passive driving; b) to investigate possible correlations between the analysed bio-signals and the driver's subjective experience, including the perception of danger, the sense of presence, the in-vehicle perceived comfort and the quality of the driving experience, the level of engagement.

The present work represents a qualitatively investigation of a methodology, based on Driving Simulator session, which could give interesting indications to the vehicle development teams for autonomous and assisted driving about the efficiency/comfortability of the driving intelligence in realistic highway driving scenarios. To prove the methodology, we have preliminary performed a test with 13 participants. The number is still not statistically significative, and the result of the present work are to be intended for a) investigation and tuning of the methodology and b) for the researchers to understand if, applying signal processing developed in previous works [1 – 5] and given a bigger number of participants, the approach could scientifically provide innovative and quantitative indexes for classifying autonomous driving algorithms.

Sensor description

The instrument block diagram is shown in Figure. 1. The ECG is acquired on three channels by amplifying and conditioning the differential voltages V_5-V_1, V_4-V_3 and V_3-V_2. These voltages are picked using a commercial vest having wet electrodes positioned accordingly to Figure 1. The SPR signal is acquired by amplifying and conditioning the differential voltage picked between the palm and the back of each hand using Ag/AgCl electrodes. The reference voltage V_{REF} is applied on the chest and on the wrists.

As shown in Figure 1 each differential voltage (3 ECG and 2 SPR channels) is properly amplified and filtered; a Digital Signal Processor (DSP) acquires the amplified analog voltage and transmits through Universal Asynchronous Receiver Transmitter (UART) the acquired data to the Wi-Fi module that converts UART data into IEEE 802.11 wlan protocol. The device is supplied with one Li-Ion cell: a 3.7 V 1000 mAh lithium polymer battery is used. The device current consumption during transmission is 200 mA, allowing 5 hours of continuous data acquisition. A power supply section of the circuit reduces the battery voltage to +3.3 V using a buck DC-DC converter; the voltage V_{REF} = 1.65 V for the reference electrodes and for the instrument is provided by a linear voltage reference.

SPR circuit design

The voltages on the palm and back electrodes are high-pass filtered by a couple of passive first order filters having cutoff frequency of 0.08 Hz, with the aim of removing the common mode DC voltage that may be present on the skin; the input impedance of the instrument is 100 MΩ, in order to assure a load error lower than 1%, assuming the skin impedance in the order of 1 MΩ. The filters are connected to the instrumentation amplifier which amplifies the differential voltage between palm and back of the hand with a gain G = 160, since the expected maximum SPR pulse is in the range of ±10 mV and must be amplified into 3.3 Vpp. Finally, the anti-alias filter has been designed as a third order low-pass filter with cutoff frequency of 40 Hz.

ECG circuit design

The circuit topology of each channel of the ECG circuit is similar to the SPR with different gain and bandwidth. As in previous subsection at each channel input there are two first order high pass passive filters. The input impedance is, as in previous subsection, 100 MΩ in the pass band. The filters outputs are connected to the instrumentation amplifier, which amplifies the differential voltage relative to the channel (i.e. V_5-V_1, V_4-V_3 or V_3-V_2). The maximum range of the input differential voltage is supposed to be ±4.5 mV, this yields to a gain G = 370 if it must be amplified into a signal with maximum amplitude 3.3 Vpp. The frequency behavior of the sensor is a band pass system with gain 370 in band center, a lower cutoff frequency at 0.022 Hz with a slope of 40 dB/dec and an upper cutoff frequency at 170 Hz with a slope of -60 dB/dec.

DSP and WiFi data transfer

The conditioned analog signals are sent to an analog input of the DSP. The chosen DSP has an on board 12 bit A/D converter operating at 8 Mega Instructions per Second; the sample rate has been set to 1 kSa/s. The converted data are sent via UART protocol to the WiFi module. The Baud Rate for data transfer has been set at 115.2 kbps in order to allow the data flow without crowding the channel.

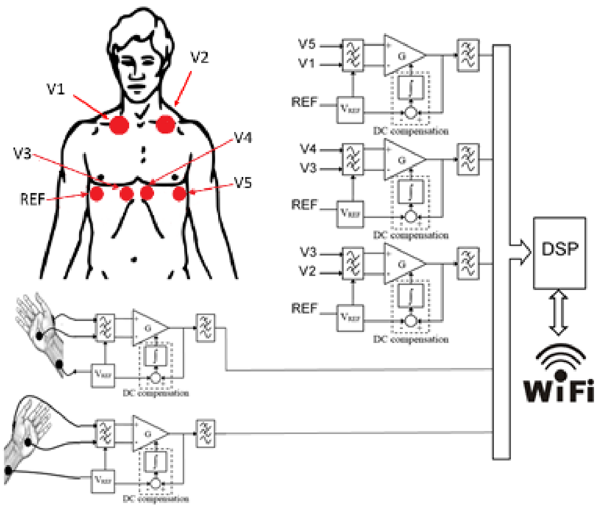

Figure 1: Sensor block diagram

Acquisition software description

For data acquisition, the WINTAX 4 PRO software by Magneti Marelli has been used. It is a suite of data analysis tools developed for motorsport and it provides highly advanced real time analysis functions as well a standard interface to team's proprietary software applications. Specific characteristics of the program can be found in [6]. The advantage of this choice is the capability of monitoring, in real time and synchronized both the vehicle telemetry and the biotelemetry data. Moreover, we can add real time processing of the data channels implementing complex functions like the features used in heart rate variability analysis. In Figure 2 an example is shown of real time monitoring and analysis of ECG, SPR and HRV parameters dynamic time-varying changes.

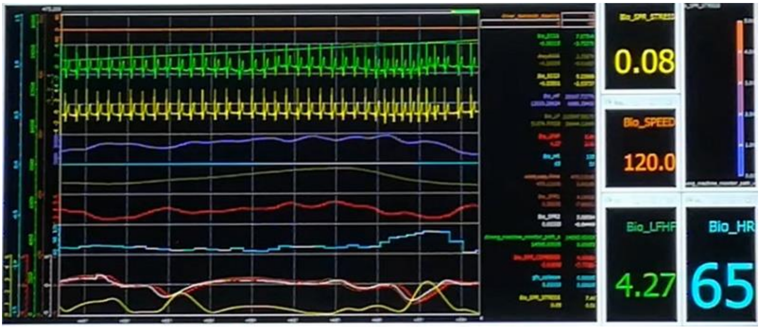

Figure 2: Real time monitoring of multiple physiological parameters using WINTAX 4 PRO

Motion artifact removal on SPR signals

The proposed algorithm is based on the assumption that the interferences due to motion cause an increase of the energy of the measured signals, because a first energy contribution is provided by EDA pulses and a second energy contribution is related to hands movement. On the basis of this assumption, we first evaluate the RMS_1 and RMS_2 values of the spr_1 and spr_2 signals on a moving time window of duration of 1 s. At the i^{th} sample, $RMS_{1,2}$ results:

$$RMS_{1,2}(i) = \sqrt{\frac{\sum_{n=i-N+1}^{i} spr(n)_{1,2}^{2}}{N}} \tag{1}$$

We consider the function:

$$f(x) = \frac{1}{1 + e^{-2(x-1)}} \tag{2}$$

$\alpha(i)$ used to combine spr_1 and spr_2. The i^{th} sample of $\alpha(i)$ is obtained as

$$\alpha(i) = \begin{cases} f\left(\frac{RMS_1(i)}{RMS_2(i)}\right) & if\ RMS_2 \neq 0 \\ 1 & if\ RMS_2 = 0 \end{cases} \tag{3}$$

From (2) and (3), it is clear that $\alpha(i) \rightarrow 0$ when $RMS_1 << RMS_2$ and $\alpha(i) \rightarrow 1$ when $RMS_1 >> RMS_2$. The output is finally obtained as a linear combination of spr_1 and spr_2:

$$OUT(i) = \alpha(i) \cdot spr_2(i) + [1 - \alpha(i)] \cdot spr_1(i) \tag{4}$$

Intuitively, the input signal (whether spr_1 or spr_2) with lower energy content (within the moving window of duration 1 s) is considered more reliable for the output in (4).

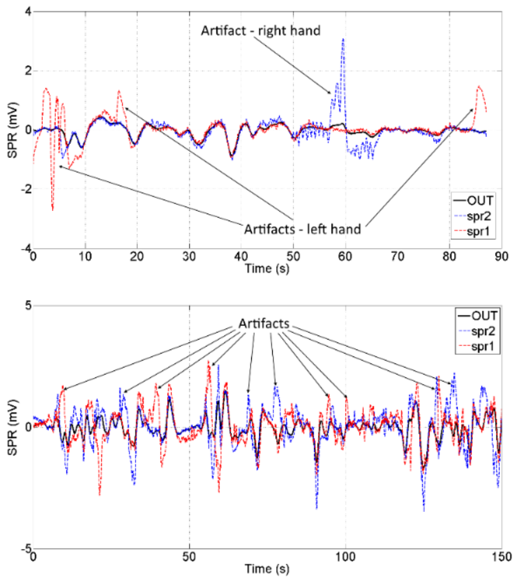

Figure 3: motion artifact removal: in lab experiments (top) and on simulator (bottom)

Figure 3 shows the results obtained from the motion artifact removal algorithm. In top graph the experiments are initially conducted in laboratory, just moving the hands and disturbing the electrodes during acquisition; in bottom graph we acquired the signals during a lap driven on a driving simulator. The simulated circuit is Jerez de la Frontera, the driver never drove on it before. It is evident that the algorithm follows the input with minimum energy when there is discordance between inputs. In both charts, it is possible to see that the motion artifacts are perfectly removed by the proposed algorithm.

Test description

A group composed by 13 healthy volunteers (age 31.4 ± 9, 8 males and 5 females) drove on simulator in a highway scenario in three different phases. The first phase had the aim of training people and getting familiar with the simulator driving with very low traffic in a highway for 5 minutes, while their baseline bio-signals were acquired. In the second phase participants had to manually drive in a highway for a distance of 40 km (corresponding approximately to 20 minutes). On the path, at specified positions, where posed four tasks: 1) **overtaking** another car which is undecided on which lane to keep; 2) brake maneuver because two **trucks** ahead are overtaking each other; 3) lane **narrowing** and mandatory shift due to road works and 4) unexpected lateral **wind** gusts. Figure 4 shows the screenshots of the tasks.

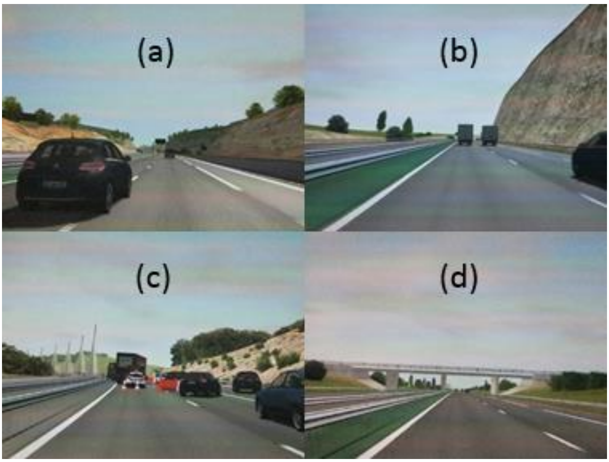

Figure 4: the tasks positioned on the road. (a) overtake of undecided car; (b) trucks overtaking; (c) lane narrowing; (d) wind gusts.

During the third phase, the subjects were on the same highway with an autonomous car which had to cope with the same tasks. In order to avoid data bias, one half of the volunteers did the manual drive before the autonomous, and the other half vice versa.

As an example, Figure 5 shows the typical behavior of heart rate and SPR signal during the tasks.

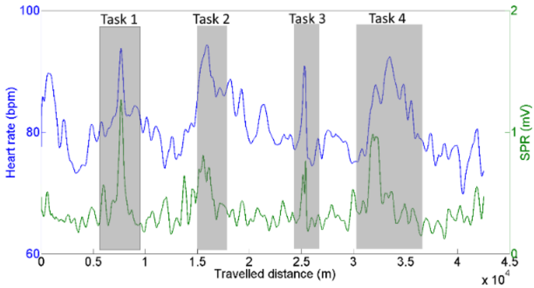

Figure 5: heart rate and SPR signal versus the travelled distance. The effect of the tasks is clearly visible in both traces

As it can be seen, during the tasks there is a significant increase of heart rate and SPR signal.

Heart Rate Variability Analysis

To validate real-time assessment obtained with the VI-grade experimental set-up in the Driving Simulator , the same sensor system was tested in the electrophysiology laboratory in the BACPIC, where the quantitative HRV analysis was also performed with the Kubios software (version 3.0.2), with time-varying algorithms and, according to the European Society of Cardiology guidelines [7], in the time domain (TD), in the frequency domain (FD), and with non-linear (NL) methods. In the FD, the LF/HF ratio, calculated from LF and HF in normalized units, was accepted as an index of sympathovagal interaction adequate to explore autonomic modulation [8]. For comparison, 12-lead ECG was continuously recorded also with Mortara Surveyor/X-Scribe

The validation protocol consisted of five phases: implying: 1) 10-minutes baseline supine; 2) 10-minutes Head-up tilting 70^0 (HUTT); 3) 10-minutes supine recovery); 4) 20-minutes mental stress with "Mensa" preliminary Tests; 5) exercise test, at bicycle-ergometer, until muscle exhaustion.

Experimental results

SPR Activity

As a first analysis, we want to investigate if there is a significant difference in SPR activity from a task to another among all the subjects. To do this, since every person has different electrodermal activity, we evaluated the RMS value of the SPR over the tasks normalized with respect to the RMS evaluated on the entire trace. Figure 6 shows the graph with the normalized SPR for manual (red line) and autonomous (blue line) driving.

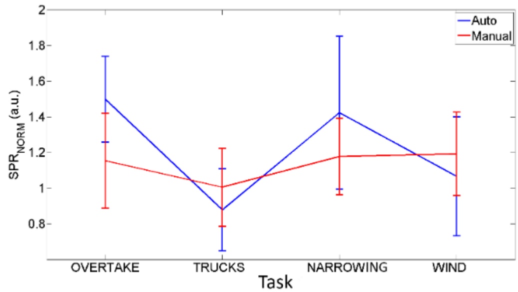

Figure 6: Normalized SPR over the tasks for autonomous (blue line) and manual (red line). Error bars represent standard deviation among the subjects

Referring to manual driving (red line), we observe that “wind” is slightly more stressful than the other tasks and that “trucks” is slightly less stressful than the others, although SPR activity is comparable in all tasks when the subject is driving manually ($p>5\%$). Instead during autonomous driving (blue line), “overtake” is perceived as more dangerous compared with “trucks” and with “wind” ($p<5\%$). “Trucks” are perceived as less dangerous than lane narrowing, which is more dangerous than “wind”.

As a second analysis, we wanted to evaluate if the tasks are perceived more dangerous in autonomous or manual drive. Performing the t-test for each obstacle, we obtain that “overtake” is perceived significantly more dangerous ($p=0.2\%$) in autonomous. Although non-significant the lane narrowing is perceived slightly more stressful in autonomous drive ($p=8\%$), whereas “trucks” and “wind” are perceived in the same way ($p=18\%$ and 28%, respectively).

As a third analysis, we want to investigate if the four tasks (in general) are perceived more stressful in autonomous or manual drive by each participant. Figure 7 shows the results.

Before providing any comment, we must point out that two different autonomous algorithms were tested: an aggressive autonomous drive algorithm (subjects 1 – 6) versus a smooth autonomous algorithm (subjects 7 – 13).

As it can be seen in Figure 7, there is a huge variability among the subjects. However, increasing the confidence of t-test to 20%, subjects 1 and 3 perceived autonomous drive as more dangerous than manual, subjects 7, 9, 10, 12 and 13 had the opposite reaction (autonomous drive was perceived less stressful than manual), whereas no difference was appreciable in the remaining ones.

Particularly interesting is the fact that, for most participants the autonomous drive was significantly less stressful than manual when the smooth algorithm was used. Instead, with the aggressive autonomous algorithm, two people over six perceived autonomous drive as significantly more dangerous than manual.

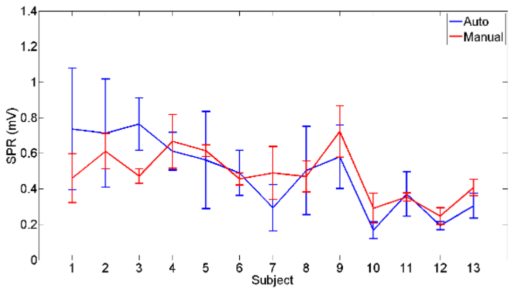

Figure 7: SPR averaged on the tasks for each subject in autonomous (blue line) and manual (red line). Error bars represent standard deviation among the tasks.

Finally, we evaluated if, over the entire distance, the autonomous drive is perceived as less stressful than manual. The autonomous drive resulted significantly ($p<5\%$) less stressful than manual for subjects 7 – 13 (i.e. when smooth autonomous algorithm was used. (Figure 8).

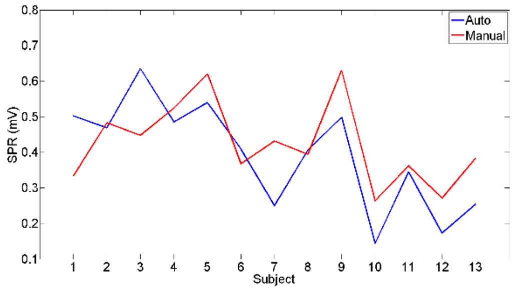

Figure 8: SPR averaged on the entire distance for each subject in autonomous (blue line) and manual (red line).

Real- time HRV-Sensor validation in the physiology laboratory.

The signals quality obtained with the VI-grade sensor system was optimal for reproducible real-time calculation and monitoring of time-variant HRV parameters (Figure 9).

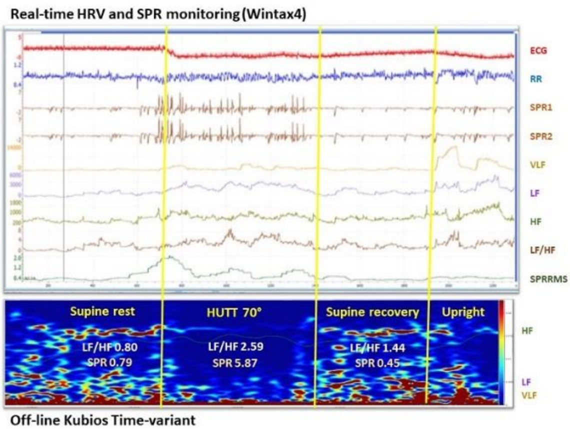

Figure 9: example of comparison between real-time monitoring of HRV and SPR parameters with WINTAX 4 and and off-line time-variant HRV analysis performed with Kubios from ECG simultaneously recorded with the BACPIC Mortara Surveyor/X-Scribe system.

Baseline SPR RMS activity ranged between 0.01 and 0.8 mV, with wide inter-individual variability. A significant ($p<0.05$) increment (up to 2.6 mV) was induced by HUTT, but not by physical exercise. SPRRMS and LF/HF had similar trend along test session. Good agreement was found between HRV parameters calculated in real-time, their off-line recalculation with Kubios software and those obtained from Mortara telemetry.

Questionnaire method and results

We have considered several subjective metrics to fully understand the overall users' experience in described context [9 – 13]. The sense of presence, defined as the sense of being in a virtual environment, it is a crucial variable to assess whenever considering a virtual environment (VE). The better a VE is designed the higher sense of presence is experienced. In different studies various parameters were manipulated in order to comprehend their influence on the perceive sense of realism of the VEs like field of view increment or presence of autonomous traffic. A second relevant self-reported measure is the User Experience (UX), namely, "the perception and reactions of a user that derive from the use, or from the prediction of use of a product, system or service" (ISO, 2009). In the automotive research field, UX questionnaires are used to assess the interacting experience regarding the in-Vehicle Infotainment Systems (IVIS) or the driving simulator itself. A third aspect important to evaluate is the systems usability insofar as it definitely affects the drivers' experience. It is defined by Shackel (2009) as "the capability to be used by humans easily and effectively" and, like UX, its applicable both to IVIS and to simulators. Several questionnaires were considered. A demographic questionnaire was used to gather background information (e.g., age, gender, experience with virtual reality devices, driving videogames, large screens). The NASA Task Load Index (NASA-TLX) is a multidimensional scale index employed to assess subjective workload. It comprises six subscales: mental, physical demands, and temporal demands, frustration, effort, and performance. The response scale ranged from low (0) to high (100). The performance is the only dimension that present different labels good (0) and poor (100). The System Usability Scale (SUS) allows evaluating the perceived usability of the system. The responses can be provided on a 5-point Likert scale. Sum scores can range between 0 and 100. A Presence questionnaire was also administered. Finally, three ad hoc questionnaires were administered to assess respectively: perceived difficulty and danger, as well as UX. The Perceived Danger and Difficulty questionnaires were considered to assess the perception of drivers, in terms of and danger and difficulty, considering the obstacles faced during the driving tasks. Both questionnaires comprised 4 items on a 5-point scale (from 1-low to 5-high). The UX questionnaire aimed at evaluating three dimensions: pleasantness, engagement, and utilization and time flow. It comprises a total of 12 items. The response could be provided utilizing a

5-point Likert scale. The Presence questionnaire shows a statistic difference in the possibility to act dimension, higher for the manual condition (p-value 0.023), due to the control exercised by participants over the vehicle. In both conditions, the realism and the quality of interface dimensions were close to the maximum scale. This information shows the high fidelity of the driving simulator used. System Usability Scale doesn't show statistical differences among conditions but the results show a general very high usability of the driving simulator. Moreover, also the User Experience questionnaire shows the high quality of the virtual environment used, with the three scales, Pleasantness, Engagement and use&Time flow, with results near the maximum for all of them. One of the main limitations of this experiment was the restricted sample size. Nevertheless, on the first half of the participants, we wanted to verify the effect of an aggressive behavior of the robot driver: sudden brake actions (instead of smooth) during the execution of some tasks. With the second half of the participants, the robot driver acted much more smoothly for the same tasks. The analysis of the questionnaires on the total population shows some important trends that are indicated in the form of percentual increment between the two conditions. For a preliminary comparison with bio-signal data, we selected perceived Difficulty from questionnaires. Figure 10 shows the results of questionnaires relative to the perceived difficulty.

Finally, NASA-TLX shows, as expected, that manual driving is more demanding in terms of physical demand (20%), mental demand (5%) and effort (20%). On the contrary, the frustration and performance dimensions show the impact of the AI aggressiveness on participants. In fact, in the manual condition, the performance was evaluated lower by 10%. This can be explained by the perceived difficulty results in the 7-13 pool that shows that the Wind obstacle is the most challenging in manual condition. This information gives us an explanation of the performance higher value for autonomous driving. The frustration dimensions, higher in the autonomous condition, is due to the errors in the behaviors of AI vehicles, that cannot be avoided without control over the vehicle, producing a higher value in the participants.

Conclusions

Our experimental set-up has proven reliable for real-time monitoring and quantitative estimate of driver's instantaneous stress reaction induced by driving demands. The research performed has shown some promising results for the methodology used to define a subjective/objective metrics to classify the effect of various robotic driving styles in highway realistic driving on a professional Driving Simulator, which offers a a seamless real/virtual human reaction. This means that, when a statistically meaningful number of participant data would be available, the methodology applied on the Driving Simulator could be extended to a real-life case. While with the former the external traffic environment is provided as deterministic, with the latter a random variation is unavoidable, thus the approach is very useful to isolate the most influencing parameters. The research team is planning to repeat the experiment to reach at least a total of 30-40 participant in the upcoming months.

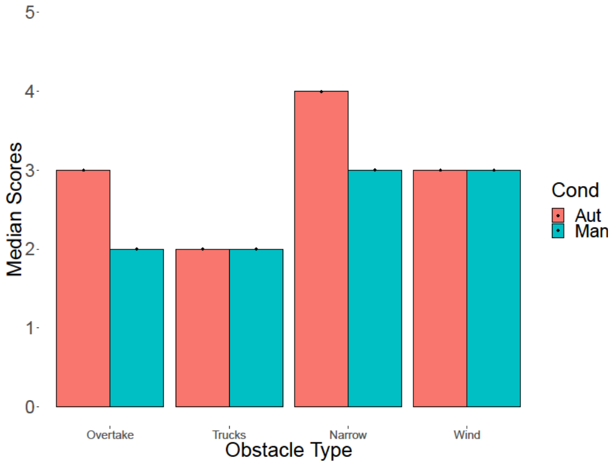

Figure 10: Median scores of the questionnaires on the perceived difficulty.

References

[1] A. Affanni, “Wearable instrument to measure simultaneously cardiac and electrodermal activities”, IEEE International Symposium on Medical Measurements and Applications, MeMeA 2016 - Proceedings, 2016.

[2] A. Affanni and G. Chiorboli, “Wearable instrument for skin potential response analysis in AAL applications”, 20th IMEKO TC4 Symposiu-m on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014, 2014.

[3] A. Affanni and G. Chiorboli, “Design and characterization of a real-time, wearable, endosomatic electrodermal system”, Measurement, vol. 75, pp.111-121, 2015.

[4] A. Affanni, G. Chiorboli, and D. Minen, “Motion artifact removal in stress sensors used in driver in motion simulators ”, Medical Measurements and Applications (MeMeA), 2016 IEEE International Symposium on, 2016.

[5] A. Affanni, R. Bernardini, A. Piras, R. Rinaldo, and P. Zontone, "Drivers stress detection using skin potential response signals", Measurement, vol. 122, pp. 264-274, 2018.

[6] http://docs.gestionaweb.cat/1662/catalogo-magneti-marelli-2017.pdf

[7] Task force of the European Society of Cardiology and the North American Society of pacing and Electrophysiology.Heart rate variability. Standards measurement, physiological interpretation, and clinical use. Eur Heart J. 1996;17:354-81.

[8]. Lombardi, F., & Stein, P. K. (2011). Origin of heart rate variability and turbulence: an appraisal of autonomic modulation of cardiovascular function.Front Physiol, 2, 95.

[9] Wright, S., Ward, N. J., & Cohn, A. G. (2002). Enhanced presence in driving simulators using autonomous traffic with virtual personalities. Presence: Teleoperators & Virtual Environments, 11(6), 578-590.

[10] Lin, J. W., Duh, H. B. L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In Proceedings ieee virtual reality 2002 (pp. 164-171). IEEE.

[11] Hess, A., Jung, J., Maier, A., Taib, R., Yu, K., & Itzstein, B. (2013, June). Elicitation of mental states and user experience factors in a driving simulator. In 2013 IEEE Intelligent Vehicles Symposium (IV) (pp. 43-48). IEEE.

[12] Schultheis, M. T., Rebimbas, J., Mourant, R., & Millis, S. R. (2007). Examining the usability of a virtual reality driving simulator. Assistive Technology, 19(1), 1-10.

[13] Robillard, G., Bouchard, S., Fournier, T., & Renaud, P. (2003). Anxiety and presence during VR immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. CyberPsychology & Behavior, 6(5), 467-476.