Passenger Communication System for Next-Generation Self-Driving Cars: *A Buddy*

M Talha Bin Ahmed Lodhi[1]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Faisal Riaz[2]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Yasir Mehmood[3]
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Muhammad Farrukh Farid[4]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Abdul Ghafoor Dar[5]
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Muhammad Atif Butt[6]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Samia Abid[7]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Hasan Ali Asghar[8]
Control, Automotive and Robotics Lab
aff. with NCRA
Department of Computer Science and Information Technology
Mirpur University of Science and Technology (MUST)
Mirpur-10250, AJK Pakistan

Abstract—**With the rapid emergence of autonomous vehicles, there is a need to build such communication systems which help the passengers to communicate with autonomous vehicles (AVs) robustly. In this regard, this research work presents a multimodal passenger communication system. The communication system is known as "buddy" for AVs. Buddy is an all in one control system for AVs which incorporates touch, speech, text, and emotion recognition methods of interaction. Buddy makes it easy for passengers to interact with AVs. It enables the communication between the passengers and the AV which eventually provides a safe driving experience. Moreover, we have proposed and developed our own simulator two evaluate the performance of our proposed passenger communication system. We have also conducted extensive infield-tests to test the effectiveness of the proposed system. The extensive rigor analysis validates the results and hence the significance of the proposed passenger communication system.**

Keywords—*Autonomous Vehicles (AVs); passenger communication system; the simulation engine*

I. Introduction

Communication is an important part of human beings to convey their messages to each other. Popularizing communication in the vehicles is key in the proliferation of this rapidly emerging technology. Autonomous vehicles (AVs) are the latest type of robots and there is a need to build a communication system that helps the passengers to communicate with AVs. This is signified by the work of Wang et al. [1] which presents methods of preparation for the mass emergence of AVs. They show how AV capabilities can be extended to underwater army submarines thus signifying the impact of AVs in military applications. Hence, the need arises to engineer autonomous systems with Human-Computer Interaction (HCI) capabilities.

HCI is a field that helps humans to interact with machines. Extensive research has been done in HCI to make systems autonomous. Valeria et al. [2] have evaluated the factors played by psychology in designing systems. The role of HCI is vast and applies to many unique types of systems.

Moreover, in the HCI-based research, different fields have been explored at large. For example, Venuto et al. [3] have presented the state of the art HCI system for the remote control of a mechatronic actuator, such as a wheelchair or a car. Leminen et al. [4] have proposed a project detailing an HCI for a smart home geared for research purposes. Pons et al. [5] worked on interactive software geared for animals, referred to as ACI (animal-computer interaction). However, the role of HCI has not been explored to tailor novel interaction solutions for a passenger to AVs.

A lot of research has been done to make HCI systems for controlling AVs. Woo et al. [6] have worked on remote driving tools for AVs which are a collaboration between remote driving tools and sensor fusion displays. Xiao et al. [7] have proposed sensor fusion methodologies while Poveda et al. [8] have developed hybrid source seeking controllers specifically for AVs. However, these HCI systems cannot compute the effective state of the passenger sitting in AV.

Considering the above-mentioned limitations, we have tailored affective computing inspired passenger to AV (P-AV) interaction system. Since there is a need for building a communication system that helps the passengers to communicate with AVs, we propose a passenger communication system (PCS) which consists of 4 modules as shown in Fig. 1. For the testing and evaluation of the proposed system in different environments, we have proposed and developed our simulator. Moreover, the extensive testing of the proposed system has been done on EMO (AV) to test the performance of the proposed PCS.

The rest of this paper is organized as follows. Section II presents the literature review. The system methodology has been formally defined in Section III. The simulation results are described in Section IV. Finally, the conclusions are presented in Sections V.

II. LITERATURE REVIEW

Communication in autonomous vehicles' is an important emerging technology. We discuss its importance in different threads as follows. Hernandez et al. [9] describes how different types of interactions can help to measure and manage the stress of a driver. For this purpose, they join both wearable technology and business into the steering wheel of a vehicle which allows no upsetting spotting. Ragot et al. [10] describe how human emotions can recognize by physiological signals. They trained models using SVM classifier and compare emotion recognition accuracy by using laboratory sensor "Biopac MP150" and wearable sensor "Empatica E4". P. Karthikeyan et al. [11] focused on identifying the stress analysis with the advanced processing of physiological signals which is based on the mental arithmetic task. Chang et al. [12] have proposed physiological signals for an emotion recognition system based on the support vector recognition.

A driver's driving may affect due to emotions and non-driving associated reasoning tasks which may cause traffic accidents. Boril et al. [13] focused on the task of classification of emotion and cognitive load of real driving scenarios to estimate "speech production-based" and "Cepstral-based acoustic" features. They applied SVM and GMM classifiers that gave 79% and 95.2% classification performance in the task of classification of emotion as neutral vs. negative and

cognitive classification respectively. Tarnowski et al. [14] presented a 3D face model to calculate the features. The results of these features are based on facial expressions. With 2D images, light condition, and movement of head positions play a sensitive role to recognize emotions using cameras. For features classification, the "K-NN" classifier and "MLP neural network" are performed. The experiments showed the best results of these seven emotional states. However, we can also recognize the emotions of passengers in AV via facial expressions. In another work, Hickson et al. [15] presented a novel way for data collection and also provide a novel approach to increase the accuracy of CNN deep learning through personalization. Currently, the emotion recognition system is suffering in facial image conversion performance. To overcome this problem, Wang et al. [16] proposed a novel approach for emotion recognition system using "Jaya algorithm". This algorithm has been used to train the dataset and guaranteed that it won't stick with the training set to local optimum point and has fewer necessities over hyper-parameters. In this article, Eyben et al. [17] have presented a novel approach to improve the intelligent measures for driving safety in automatic driving systems. Gordon McIntyre and Roland Gocke [18] discussed particular problems faced in emotion recognition system and try to deal with natural way. They presented a novel approach which incorporates semantic descriptions and feature sets of computer vision. Computer vision process is used to detect facial expression emotions with combination muscle movements' codebook. Automatic speech recognition and computer vision are a dealer in machine learning and pattern matching. Their framework, effective communication consists of the generic model with the ontology domain. They recognize emotions by facial expressions, but they didn't consider a passenger's emotions of an AV.

III. SYSTEM METHODOLOGY

This section provides details about the proposed agent architecture of the PCS as follows:

A. Proposed Agent Architecture of Passenger Communication System

In this section, we present the proposed agent architecture of the PCS as shown in Fig. 1. Our agent consists of six modules which includes the sensory module, microsoft cognitive services, artificial amygdala, dialog generation, strategy selection and actuator module. Initially, our agent receives the perceptual data from the sensory module. From the obtained data, it will analyze the facial expressions of the passenger/client. The data is then forwarded to the microsoft cognitive services which reads the emotions from the facial expressions of the passenger/client. The results generated by this module are then forwarded to the artificial amygdala module which processes the fear level experienced by the passenger/client. After examining the fear evaluated by the passenger/client two processes will be executed:

- Strategy Selection Module
- Dialog Generation Module

First, the strategy selection module creates a direct link from the artificial amygdala module. It picks a strategy upon
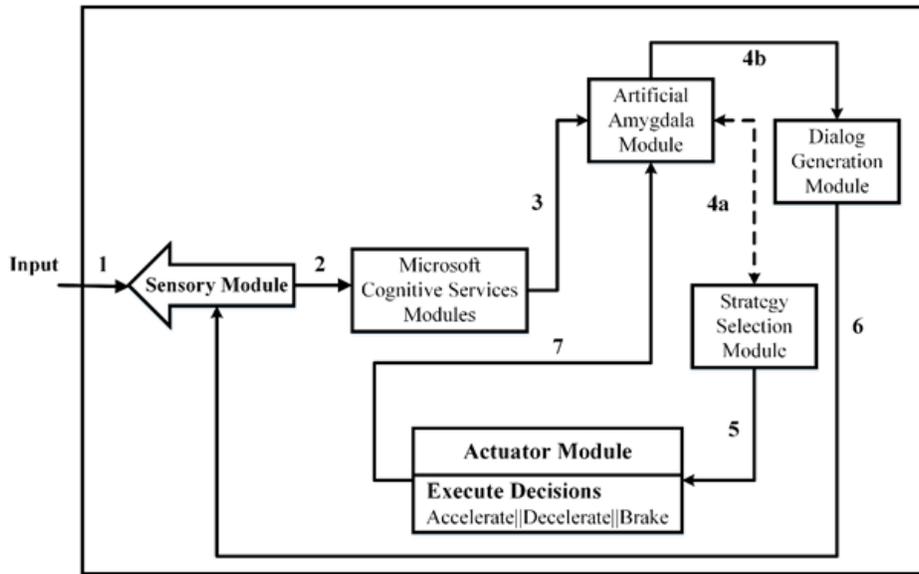
Fig. 1. The Proposed Agent Architecture of the Passenger Communication System

---

Algorithm 1 Proposed algorithm of PCS

1: **procedure** PROCESSINPUT(TEXT)
2: $\quad action \leftarrow LoadActionFile()$
3: $\quad command \leftarrow Action.Parse(text)$
4: $\quad$ **if** $Command! = Empty$ **then**
5: $\quad\quad$ **return** Command
6: $\quad$ **if** $Command == Empty$ **then**
7: $\quad\quad$ Throw Exception
8: **procedure** PROCESSSPEECH(VOICE_ INPUT)
9: $\quad InitializeSpeechModule$
10: $\quad Text \leftarrow SpeechModule.Recognize(voice\_input)$
11: $\quad Command \leftarrow CallProcessInput(text)$
12: $\quad$ **return** Command
13: **procedure** RECOGNIZEEMOTION()
14: $\quad$ **Initialize** CameraModule
15: $\quad Photo \leftarrow CameraModule.GetFrame()$
16: $\quad Result \leftarrow CognitiveServices(Photo)$
17: $\quad$ **if** $Result.Contains(Face)$ **then**
18: $\quad\quad$ **return** Result.Face.DominantEmotion
19: $\quad$ **if** $Command == Empty$ **then**
20: $\quad\quad$ Throw Exception

---

which the agent drives itself and then it sends the strategy to the actuator module which in turn accelerates/decelerates or applies the Brake. It continuously checks the fear level from the artificial amygdala and selects the strategy after checking what sort of emotion is being experienced by the passenger/client. Secondly, after checking the emotion of the passenger/client the agent tries to soothe the passenger/client with its pattern matching language (PML) in the dialog generation module and assess the expression of the passenger/client through the input obtained from the sensory module. The sensory module consists of multiple sensors that are commonly deployed with AVs like sonar, camera, Lidar. The step-wise procedure of the proposed PCS is provided in the Algorithm 1.
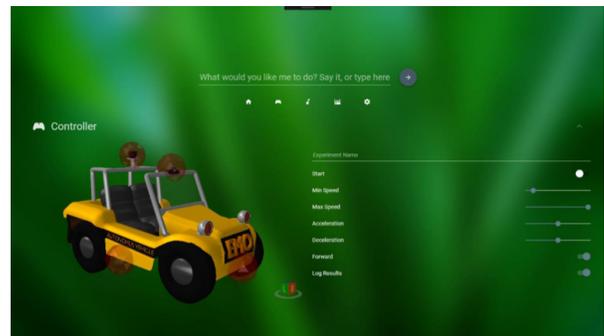
Fig. 2. Simulation Screen of the Controller Integrated into "Buddy" App

IV. EXPERIMENTAL SETUP AND RESULTS

In this section, we discuss the results of our proposed PCS (known as "Buddy"). For our PCS "Buddy" can be used on different platforms we have first targeted desktop application. Our desktop application is fully functional on any distribution of windows. Here, we briefly discuss the functionalities of PCS as follows:

A. PCS

Our proposed PCS is very user-friendly. It communicates with users to provide a comfortable environment in which a user can say whatever they want as shown in Fig. 2. We have implemented our language processor called PML which uses defined Regular Expression (REG-EX). REG-EX understands what the user currently desires for at time 't'. The user can type or say it and our language processor shows results accordingly. The language processor is capable enough to process what the user says and gives results e.g. "I want to go to Lahore". It checks in its predefined REG-EX and tells whether you have entered a valid statement or not.

After exploring the Home Section, next comes the con-
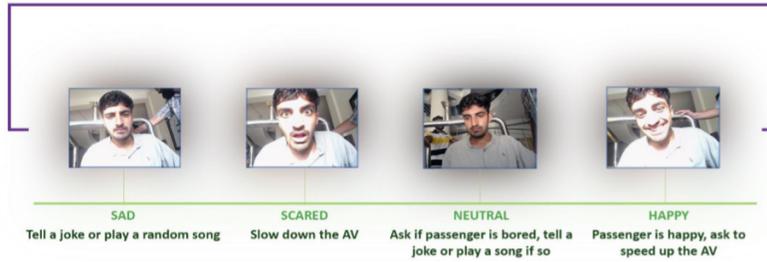
Fig. 3. Various Considered Facial Expressions of the Passenger
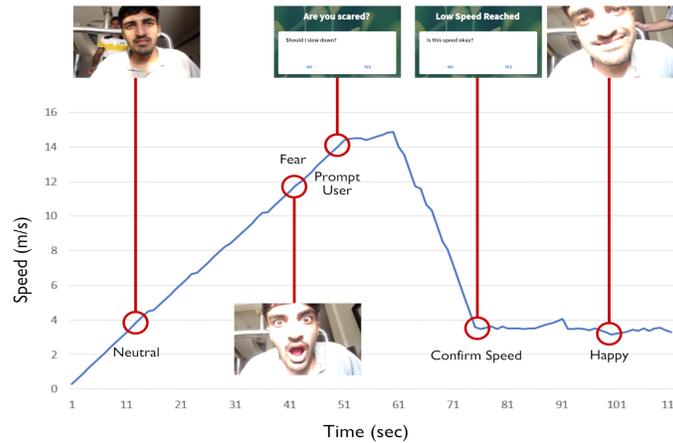
Fig. 4. Effect of 'Fear' on Driving Strategy
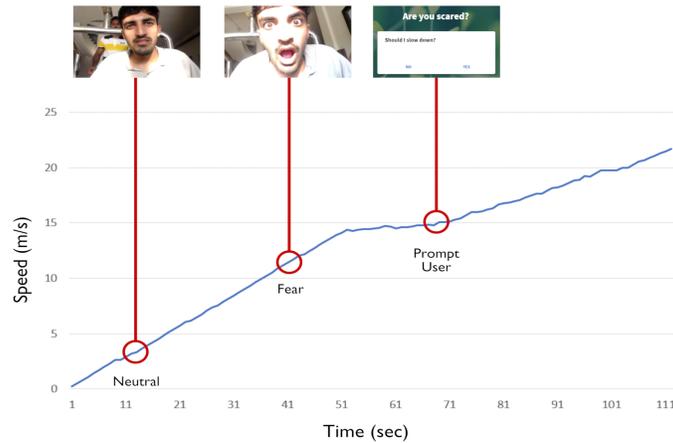
Fig. 5. Effect of Feeling 'Fear' on Driving Strategy when a Passenger wants the AV to Accelerate

troller module. This module is directly related to any sort of autonomous agent mainly "vehicles". This module provides the 3D display of the desired autonomous agent. In Fig. 2 on the left side of the screen we have implemented the 3D model of EMO the first autonomous vehicle of Pakistan. As we can see in the above figure there are red circles on certain parts of EMO. These red circles indicate that there is no lidar, camera or sonar yet attached to the system. At the bottom of

the screen is a cube that is used for 360°-degree rotation of EMO.

In the rightmost corner, we have set an experiment logger. We simply have to write the name of the experiment that we want to log e.g. "BIKE_AV_COLLISION_EXPERIMENT1". The experiment will start logging after the log results button is turned ON. Otherwise, the experiments will not be logged. The Start Button is used to turn ON the AV. We can also

adjust the max and min speed as well as acceleration and deceleration of the AV. The next feature is about moving the AV either Forward or Backwards as the user desires.

We have performed the emotion recognition test on the bases of the facial expression of the passenger and examined the driving strategy being employed by the AV to sustain a safe driving experience. In this regard, several images are provided to the PCS and the adopted strategy is noted against each recognized emotion. Fig. 3 represents various facial expressions of the passenger and the messages passed to the AV in return. Fig. 4 represents the result of the "fear" emotion on the driving strategy being employed by the AV. Here, the x-axis represents the recording time of the simulation (in seconds) while the speed of the vehicle is depicted in the y-axis. Initially, when the system recognized that the passenger is feeling "neutral" the AV started accelerating its speed. At 41 seconds, when AV reached its maximum acceleration the passenger started feeling fear; the PCS sensed the "fear" and generated the prompt "Are you scared? should I slow down?". Upon receiving a positive response, the AV started decelerating itself. At 75 seconds, the vehicle reached its least acceleration, the PCS generated a prompt to confirm if the passenger is happy with the speed and continues with the same speed on receiving the consent of the passenger. In the same way, Fig. 5 represents the driving strategy adopted by the AV when the passenger felt "fear". However, in this scenario, when the prompt generated "are you scared?" on sensing the "fear" from the facial expression of the passenger, the passenger withheld to decrease the speed of the vehicle and consequently, AV continued to increase its speed keeping road conditions under consideration.

Whereas, the driving strategy being adopted by the AV in response to the "happy" emotion is represented in Fig. 6. Initially, the facial expressions of the passenger were "neutral". Hence, the AV started accelerating its speed considering the road conditions. While accelerating (at 43 seconds), the PCS recognized the "happy" emotions of the passenger and generated a prompt "you seem happy! want to speed up?" to procure the assent of the passenger to accelerate the AV. In this scenario, the passenger agreed to increase the speed and consequently, the AV started accelerating while considering the road situations. However, Fig. 7 also represents the driving strategy adopted in response to the "happy" emotion. But, in this scenario, the passenger refused the PCS to increase the speed of the vehicle, and hence the AV has maintained the same speed considering the varying conditions of the road.

V. Testing and Evaluation

To assess the performance of an AV in a particular environment, many simulation engines have been developed. The existing simulators used by different driving groups are the racing simulators mainly built on the gaming engines; however, these simulators do not generate the exact results and do not provide the required feedback. Moreover, the existing simulators do not different environment scenarios like urban, desert, snowy conditions, etc.

The proposed simulation exhibits the features of full 3D support, behavior space, and enables to integrate human emotions in AVs. Besides, it allows us to integrate multiple sensors on AVs i.e. camera, SONAR, LiDAR, etc. Our developed simulator consists of its own custom road/map designer which makes it easier to generate different road scenarios to test the performance of an AV. A 3D map created using our map designer is shown in Fig. 8. Several environments have been designed which allow the testing of AVs in different scenarios based on their location. These include environments as shown in Fig. 10.

Whereas, we have assessed the perfomance of our proposed system in different scenarios as shown in Fig. 10. To set up experiments, the speed of AVs randomly set to [2-4]. Where, each test has been executed 5 times and saved in the log file

In the first scenario (Fig. 10(a)), a single path was given to the AV and it managed to reach its destination without any deviation. While in the second scenario, a location was fed to the simulator; however, two paths were available to reach the destination as shown in Fig. 10(b). When the vehicle reached the T-zone where it had to decide which path to select in order to continue its operation, it detected an obstacle on one of the route and continued to pursue its operation on the other path. While. in the 3rd scenario (Fig. 10(c)), a destination was fed to the simulator and multiple ways were available to reach the destination, however, using its optimal path selection algorithm, our proposed system optimally selected a path to continue a smooth journey without any obstacles.

A. Discussion

Our Sim-Engine is fully functional and has been deployed in the testing of first autonomous vehicle of Pakistan EMO. Our Sim-Engine is capable of carrying out multiple scenarios based on different aspects and also it is capable of choosing different driving strategies when faced with different circumstances such as what to do when a large truck is in front of you. Furthermore, our Sim-Engine also has the capability of checking out passenger's emotions that what he/she is feeling at that time and based on that it will generate a response which will help the user to calm down or get happy. Our Sim-Engine also logs all the data (results) which are being collected at that particular instance. A comparison study of the proposed simulator is given in Table I.

VI. Conclusion

Communication in AVs is a very hot research area. However, very few systems have been proposed to enable communication between passengers and the AV in order to provide a comfortable driving experience. Hence, software-based solutions are of immense significance to enable communication between passenger and AV. Though, considering the need of a reliable PCS for the AVs, this research work presents a novel PCS to test various autonomous agents. The proposed PCS is very user-friendly and can be embedded in any sort of autonomous agent, however, we have mainly targeted AVs. The performance of the proposed PCS has been verified through in-field experiments and evaluating its effects on the driving strategy being employed by the AV. In the future, we aim to investigate performance of the proposed PCS to ensure the road safety considering different constraints.

Fig. 6. Effect of Feeling 'Happy' on Driving Strategy
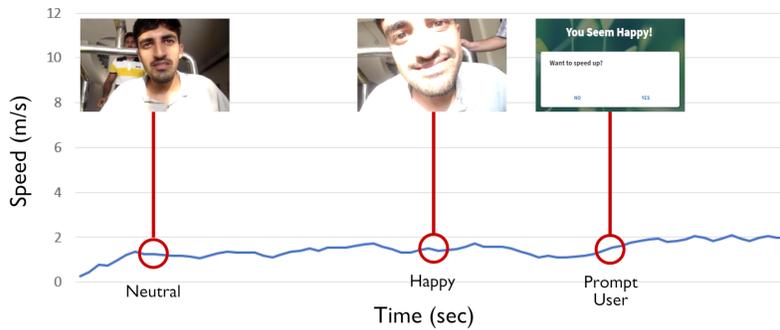
Fig. 7. Effect of Feeling "Fear" on Driving Strategy when a Passenger doesn't want the AV to Accelerate
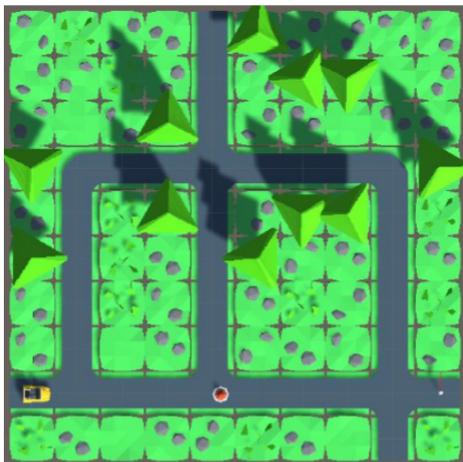
Fig. 8. Map Generated by our Custom Map Designer

REFERENCES

[1] Wang, Wen-hua, Ya-zhen Du, Lin-lin Wang, Yi Huang, Yu-xin Yao, and Hao Gao. "Development of the New Concept of Sandglass-Type Floating Production, Storage and Offloading System." In The 28th International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers, 2018.

[2] Righi, Valeria, Sergio Sayago, and Josep Blat. "When we talk about older people in HCI, who are we talking about? Towards a 'turn to community' in the design of technologies for a growing ageing population." International Journal of Human-Computer Studies 108 (2017): 15-31.

[3] De Venuto, Daniela, Valerio F. Annese, and Giovanni Mezzina. "An embedded system remotely driving mechanical devices by P300 brain activity." In Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017, pp. 1014-1019. IEEE, 2017.

[4] Leminen, Seppo, and Mika Westerlund. "Categorization of innovation tools in living labs." Technology Innovation Management Review 7, no. 1 (2017).

[5] Pons, Patricia, Alicia Carrion-Plaza, and Javier Jaen. "Remote interspecies interactions: Improving humans and animals' wellbeing through mobile playful spaces." Pervasive and Mobile Computing 52 (2019): 113-130.

[6] Woo, Jinseok, Janos Botzheim, and Naoyuki Kubota. "A socially

TABLE I. A COMPARISON OF THE PROPOSED SIM-ENGINE WITH STATE OF THE ART SIMULATORS

Features/Simulators	NETLOGO	CARLA	Orchestra	NVIDEA	Appolo	Unity	Sim Engine (Unity-Based)
Specific to AVs		✓	✓	✓	✓		✓
Full 3D Support		✓	✓	✓	✓	✓	✓
Emotion Enabled							✓
Behavior Space	✓						✓
Environment Variations		✓			✓		✓
Custom Map Designer					✓		✓
Inbuilt Sensor Emulation		✓	✓	✓			✓
Open Source	✓	✓					✓
Cloud Based				✓	✓		✓
Windows Support	✓	✓	✓	✓	✓	✓	✓
Mac Support	✓				✓	✓	✓
Linux Support	✓				✓	✓	✓

(a)   (b)

(c)   (d)

(e)   (f)
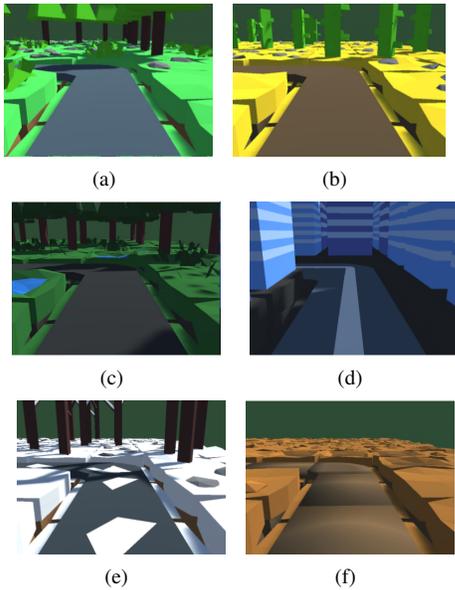
Fig. 9. Environment Variations: (a) Forest View ; (b) Desert View ; (c) Rain View, (f) City View(c) Snowy View, (f) Mud View

(a)

(b)

(c)

Fig. 10. Environment Variations: (a) Single Path for an AV ; (b) Two Paths for an AV ; (b) Multiple Paths for an AV

interactive robot partner using content-based conversation system for information support." Journal of Advanced Computational Intelligence and Intelligent Informatics 22, no. 6 (2018): 989-997.

[7] Xiao, Fuyuan. "Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy." Information Fusion 46 (2019): 23-32.

[8] Poveda, Jorge I., Mouhacine Benosman, Andrew R. Teel, and Ricardo G. Sanfelice. "A hybrid adaptive feedback law for robust obstacle avoidance and coordination in multiple vehicle systems." In 2018 Annual American Control Conference (ACC), pp. 616-621. IEEE, 2018.

[9] Hernandez, Javier, Daniel McDuff, Xavier Benavides, Judith Amores, Pattie Maes, and Rosalind Picard. "AutoEmotive: bringing empathy to the driving experience to manage stress." In Proceedings of the 2014 companion publication on Designing interactive systems, pp. 53-56. ACM, 2014.

[10] Ragot, Martin, Nicolas Martin, Sonia Em, Nico Pallamin, and Jean-Marc Diverrez. "Emotion recognition using physiological signals: laboratory vs. wearable sensors." In International Conference on Applied Human Factors and Ergonomics, pp. 15-22. Springer, Cham, 2017.

[11] Palanisamy, K., M. Murugappan, and S. Yaacob. "Multiple physiological signal-based human stress identification using non-linear classifiers." Elektronika ir elektrotechnika 19, no. 7 (2013): 80-85.

[12] Chang, Chuan-Yu, Jun-Ying Zheng, and Chi-Jane Wang. "Based on support vector regression for emotion recognition using physiological signals." In The 2010 International Joint Conference on Neural Networks (IJCNN), pp. 1-7. IEEE, 2010.

[13] Boril, Hynek, Seyed Omid Sadjadi, and John HL Hansen. "UTDrive: Emotion and cognitive load classification for in-vehicle scenarios." In The 5th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems. 2011.

[14] Tarnowski, Paweł, Marcin Kołodziej, Andrzej Majkowski, and Remigiusz J. Rak. "Emotion recognition using facial expressions."

Procedia Computer Science 108 (2017): 1175-1184.

[15] Hickson, Steven, Nick Dufour, Avneesh Sud, Vivek Kwatra, and Irfan Essa. "Eyemotion: Classifying facial expressions in VR using eye-tracking cameras." In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1626-1635. IEEE, 2019.

[16] Wang, Shui-Hua, Preetha Phillips, Zheng-Chao Dong, and Yu-Dong Zhang. "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm." Neurocomputing 272 (2018): 668-676.

[17] Eyben, Florian, Martin Wöllmer, Tony Poitschke, Björn Schuller, Christoph Blaschke, Berthold Färber, and Nhu Nguyen-Thien. "Emotion on the road—necessity, acceptance, and feasibility of affective computing in the car." Advances in human-computer interaction 2010 (2010).

[18] McIntyre, Gordon, and Roland Göcke. "The composite sensing of effect." In Affect and Emotion in Human-Computer Interaction, pp. 104-115. Springer, Berlin, Heidelberg, 2008.