How Human-Computer Interaction Work in An Intelligent Cockpit

Junting Hao *

Nanjing Foreign Language School, Nanjing, China

* Corresponding Author Email: Alan.hao27@nflsicc.com

Abstract. As technology quickly develops, cockpit in automobile has developed from traditional mechanical seating into modern cockpit. This means more physical buttons are replaced by new car central control display. Human-Computer Interaction (HCI) plays an important role in intelligent cockpits to make these designs useable. This paper is going to introduce some of the latest design of HCI in intelligent cockpit. Start from the definition of HCI, this paper will introduce some application and design about HCI in intelligent cockpit. The first design is a language model that can record drivers' driving habits so it can predict drivers' behaviors and give out personal services. In the second paper, the author proposes a model based on YOLOv5s to capture and recognize gestures. A playing system is also designed, so together, drivers can control music playing by making different gestures. The last application is about AR-HUD. This display system has considered several factors of driver and different warning situation. It turns out that it can shorten drivers' reaction time. In the discussion part, this paper points out that the future job should be based on optimizing computing models and ensure model performance in complex and real-life situations.

Keywords: Human-Computer Interaction, intelligent cockpit, application.

1. Introduction

As technology quickly develops, automobile interiors have developed from mechanical seating into the intelligent cockpit [1]. The traditional physical buttons are replaced by the new car central control display with a touch screen. The intelligent cockpit refers to an advanced interior system that includes Artificial Intelligence (AI), Internet of Things (loT), big data and Human-Computer Interaction (HCI). It breaks the barrier between drivers, cars and the outer environment. The main feature of the intelligent cockpit holds recognizing voice instructions, gesture control, studying drivers' preferences, predicting future needs based on data analysis and integrating with smart devices in automobiles. HCI plays an especially important role in these services. HCI helps to make the drivers and the computers on automobiles connect efficiently. Through HCI, humans and computers can understand the order or instructions transmitted by each other. In this way, the driving experience can be improved, and automobiles become more convenient and safer.

The idea of HCI was first proposed by American psychologist called Licklider JCR, he produced the viewpoint of 'Human-Computer Close Symbiosis' [2]. This is considered as the foundation of HCI, based on this principle, after decades, the idea HCI formed. Nowadays, HCI is defined as the interaction between humans and computers to send information. It focusses on how to make humans and computers communicate efficiently, naturally, and comfortably [3]. HCI can be used in many fields. Take Apple as an example, before 2007, the mobile phone market was dominated by Nokia and Blackberry. Till Jobs published the iPhone with touchscreen, this improved the interaction experience of mobile phones. Subsequently, based on the idea of HCI, iOS systems were gradually formed, this makes Apple became competitive in market. Also, HCI can be used in the medical area since Brain-Computer Interaction (BCI) is also a kind of HCI. It is widely used to help paralytic recovery [4]. Moreover, HCI helps to realize some functions in intelligent cockpit. Cameras and microphones are used to make recordings. So that computers in automobiles can analyze and detect the state of the driver. So, cars can warn the driver to focus on driving, this can avoid many accidents caused by road rage and fatigue driving. Since this domain is promising and quite important in the current era, it is required to provide a comprehensive overview of this field.

The rest of the paper is going to be organized as follows. First, in Section 2, this paper aims to introduce some different methods of HCI proposed by other researchers. Then, in Section 3, this paper will analyze the technologies and point out their shortcomings. Based on this knowledge, the prospects for the future will be given out. Finally, in Section 4, this paper will summarize the entire article.

2. Method

2.1. Introduction of HCI

HCI is defined as the study of the technology of humans and computers to influence each other [5]. But as technology develops quickly, the role of computers in the process of interaction has been replaced by robots. Also, the ways of interacting have changed over time. Based on different ways of interacting, they can be separated into text interaction, voice interaction, image interaction and cross-modal interaction [6]. HCI can also been distinguished by the medium of interaction between humans and computers. Interaction using keyboard, mouse or screen is called contact interaction. Non-contact indirect interaction includes eye tracking and facial recognition. The last one is called non-contact direct interaction; this usually represents Brain-Computer Interface (BCI). HCI usually works in four steps. Firstly, users use input devices to ask for advice and instructions. Then the system uses sensors or input interfaces to receive the information given by users. The system will also convert the information into recognizable signals for further processing. After the processing part, the system will provide feedback through output devices in a form that users will understand. In the field of intelligent cockpit, HCI is widely used in many ways since it has powerful abilities in dealing with instructions.

2.2. The Application of HCI in Intelligent Cockpit

2.2.1. Cockpit-Llama: Driver Intent Prediction in Intelligent Cockpit via Large Language Model

In this paper, the researchers designed a language model to predict drivers' behaviors to give out more proactive and personal services [7]. The decision of this model is based on many different factors. It will analyze drivers' operations like the control of steering wheels, pedals and control buttons. The historic driving will also be considered, like past routes, music preferences and the location usually visited. Moreover, the state of drivers will also be considered. The sensors can determine the emotion and the state of fatigue of the driver. The driving environment is also considered. The research team also built a dataset including the aspect above to improve the accuracy of predictions. In the adjusting state, the Low-Rank Adaptation (LoRa) method is adopted. LoRa can efficiently optimize the parameters of the model. Moreover, it can maintain good performance at the same time reducing money cost and training time. This helps the Cockpit-Llama work more efficiently. After a test, it turns out that Cockpit-Llama's prediction performance surpasses other advanced methods. Compared to ChatGPT-4, its accuracy improves relatively of 92.34%, 183.61%, 95.54% and 201.27% in four different tests. This means it can better understand drivers' potential requirements and greatly improve the driving experience. In the paper, it is assumed that the corresponding authors grant the copyright to use the paper for the book or journal in question.

2.2.2. Research on Gesture Recognition Technology in Human-Machine Interaction for Intelligent Cockpit

This paper proposes a model called Bi-YOLOv5s based on YOLOv5s to deal with gesture recognition with tiny movement or in uneven illumination environment [8]. Firstly, the author uses adaptive Gamma image enhancement pre-processing method based on Retinex theory to improve the robustness of model in uneven illumination. Then, to reduce the loss of information and improve the accuracy of recognition of small-scale gestures, the author constructed a novel bi-directional feature pyramid structure in the feature fusion network. It turns out that the accuracy of recognition of Bi-YOLOv5s reached 98.3%. Also, to achieve a balance between accuracy of recognition and complexity of the model, the author compressed and accelerated the constructed Bi-YOLOv5s. The author also introduces ShuffleNetv2, a modified lightweight network which has been incorporated into SKNet to Bi-YOLOv5s. This process not only improves recognition ability in complicated environments but also reduces the recognition time by 3 seconds. Moreover, the author also designed a music playing system controlled by gesture base on PyQt5. OpenCV was used to deal with pictures, then, through HD-SKBNet, a lightweight model, the gesture can be recognized. In this way, simple instructions like switching songs or stopping playing can be realized by custom gestures.

2.2.3. Design and Evaluation of Ecological Interface of Driving Warning System Based on AR-HUD

AR-HUD can directly project information and quickly catch driver's attention; it plays an important role in HCI in intelligent cockpit [9]. However, bad design may influence road safety. In this paper, based on cognitive load theory and environmental interface design theory, the researchers proposed an ecological interface cognitive balance design strategy for AR-HUD DWS. This ensures drivers can understand the information quickly and accurately. This model considers several warning situations including crashing, lane departure in different driving conditions. Researchers obtain data by using an eye tracker to follow the driver's line of sight movement trajectory, they also collect the driver's heartbeat and blood pressure to assess driver's mental stress. Moreover, they record driver's operation to analyze their reacting time accuracy to information. When dealing with the data, they extract key information and rank it based on importance. The team also optimizes the information present part by choosing color, changing sizes and locating reminders. At the same time, they improve Multimodal Interaction by balancing visual sense, auditory sense and tactile sense. After testing, it turns out that the new-designed AR-HUD ecological system can greatly shorten driver's interacting time, making driving experience more comfortable and safer. In addition, this system has better performance. It can work efficiently in bad weather, providing accurate and timely warnings to drivers.

3. Discussion

3.1. Challenges and Limitations

3.1.1. Data size and complexity

During the model development, researchers only used about 60,000 samples to adjust Cockpit-Llama. This may not be enough to include every case. So, the respond given out may not be hundred percent precise. Also, because of the size of the parameter, it takes relatively longer time to calculate. So, the model cannot give out feedback timely. This means Cockpit-Llama cannot meet the need of real-time responding of commercial driving. Additionally, the model cannot completely capture complex behaviors, its judgement may also be influenced by external influencing factors during driving states [7].

3.1.2. External influencing factors

In this article, researchers focus on static gesture recognition, however, during a car ride, jounces may happen, this will make the picture not clear, so at last it will influence the accuracy of gesture recognition. Moreover, when gesture interactions happen, usually there is more than one hand in the cockpit. The model may not be able to choose the right hand to interact with. This will lead to misunderstanding or giving wrong feedback. Lastly, this system can only be used to play or switch music. It cannot realize common instructions like answering phone calls or using navigation [8].

3.1.3. Personal services

This experiment is based on the drive simulator, however, in real life, more extreme and sudden accidents may happen, the system may not be able to deal with these situations. Also, since the participants are mainly young people who lack driving experience, in another word, different people have individual differences, the system should consider this. Additionally, this new system only focusses on warnings during the driving process. It didn't consider the synergy with other driving systems like navigation and entertainment; there may be conflict between different information. For example, the information may overlap with each other.

3.2. Future Prospects

3.2.1. Increase sample and optimize model

To get a more precise result, future research should build a bigger database by using public data like data from Kaggle or UCI. More importantly, to make the model more usable, different methods should be used to optimize inference time. The easiest one is to improve hardware. By using GPU, it can work efficiently to deal with matrix models, in this way, interference time can be shortened. Moreover, the model can be quantified by expressing the parameter and data in the model using numerical value with lower precision. For example, converting 32-bit floating-pointing numbers to 16-bit floating number can reduce memory usage and increase computing speed.

3.2.2. Simplify environment and ensure aims

To decrease the influence of complex environments, the method introduced by the article can be used in dynamic gesture tracking with complex background changes. In this way, the system can still judge the driver's gesture no matter how complicated the environment is. Additionally, simply using camera with Wide Dynamic Range (WDR) can help to form clear pictures even when the automobile is moving fast in a vibrant environment [10]. To ensure the judging aim which means the hand that the system will judge to give feedback, researchers can set a standard distance from the camera to drivers' hand, this can make sure the system will not be misunderstood.

3.2.3. Adopt complex situations and systems

To make the system able to deal with real-life accidents and give out information, researchers should extend training and database by using real accident data. Real-time traffic information can also be used to help adjust priority of warning. To give personal warning experience, the system can build separate databases for different drivers and adjust the warning ways for different people. In addition, researchers should build an information framework, for example, when navigation and lane keeping appear at the same time. AR-HUD system should adjust their location automatically. In this way, information overlaps can be avoided.

4. Conclusion

This article conducted a comprehensive review in the field of using HCI in the intelligent cockpit. This article introduces several designs including how to predict drivers' behavior and offer personal and proactive services by analyzing their daily driving behaviors. Another design mentioned is about recognizing drivers' gestures. Through modified model, intelligent cockpit can understand driver's requirement proposed by different gestures. The last design is about AR-HUD, this design did a couple of research about drivers' physical and mental preferences. Based on this data, they design an AR-HUD system which can shorten interacting time. In the discussion part, this article points out that the model now may lack data or its computing time is too long. So, in the future job will focus on increasing data collection and simplifying the model. To deal with information problems, the priority of information should be clear.

References

[1] Feng B. Security design of passenger-vehicle intelligent cockpit systems based on deep learning. Internet Weekly. 2025; (1): 38 – 41.

[2] Licklider JC. Man-computer symbiosis. IRE Transactions on Human Factors in Electronics. 2008 May; 7 (1): 4 – 11.

[3] Zuo Y. Research on the application of human–computer interaction design in mobile phone products [Master's thesis]. Qingdao University of Technology; 2013.

[4] Li Y, Zheng B. Design and development strategies of rehabilitation medical products under brain–computer interface technology. Footwear Technology and Design. 2025; 5 (7): 168 – 170.

[5] Dong S. Progress and challenges in human–computer interaction. Journal of Computer-Aided Design & Computer Graphics. 2004; (1): 1 – 13.

[6] Guo Q, Zhang J. The historical evolution, core drivers, and future trends of human–computer interaction (HCI).

[7] Chen Y, Li C, Yuan Q, Li J, Fan Y, Ge X, et al. Cockpit-Llama: Driver intent prediction in intelligent cockpit via large language model. Sensors (Basel). 2024; 25 (1): 64.

[8] Yan H. Gesture recognition technology for human–computer interaction in intelligent cockpits [Master's thesis]. Xi'an University of Engineering; 2023.

[9] Ma J, Li Y, Zuo Y. Design and evaluation of ecological interface of driving warning system based on AR-HUD. Sensors. 2024; 24 (24): 8010.

[10] Jia Z, Sui Y, Qian L, Ren X, Zhao Y, Yao R, et al. Electrochromic windows with fast response and wide dynamic range for visible-light modulation without traditional electrodes. Nature Communications. 2024; 15 (1): 6110.