



Voice User Interface Interaction Design Research Based on User Mental Model in Autonomous Vehicle

Yuemeng Du, Jingyan Qin^(✉), Shujing Zhang, Sha Cao,
and Jinhua Dou

School of Mechanical Engineering, University of Science & Technology,
Beijing, People's Republic of China
qinjinyanking@foxmail.com

Abstract. With the development of artificial intelligence, autonomous vehicle as a new form of driving comes into public view. In this new Human-Machine Interaction Model, relationship between people-vehicles and user mental model have met a great change which include the information process, sequence, sorting, information architecture, navigation, visual thinking, data visualization, interface and media design. Autonomous vehicle uses the artificial intelligence and driving big data in the Intelligent Transportation System to deal with the cognition and perception, builds new paradigm for the human-machine interaction. The mental model based on human perception and cognition is not suitable for the mental model based on artificial intelligence. To enhance the user experience, the paper presents the hybrid mental model which combine the human mental model and AI mental model, consequently analyzes the user experience and interaction design influenced by mental model which is driven by human intelligence and artificial intelligence in the traditional mode and the autonomous vehicle mode. Based on Induction and Comparative Study on different mental models, the paper concludes the advantages and disadvantages of visual and auditory channels in autonomous vehicle. Finally, the paper presents the principles of voice user interface design in the view of adaption to the new environment, meeting the gap of the needs of users' satisfaction of control and privacy protection, defines the usability targets and user experience targets for the autonomous vehicle voice user interface design.

Keywords: Autonomous vehicle · Mental model · Voice user interface design
User experience

1 Introduction

Mental model theory is an attempt to model and explain human understanding of objects and phenomenon [1, 2]. In pass, people have this models of themselves, others, the environment, and the things with which they interact and form it through experience, training, and instruction [3]. In the interaction design, mental model is the crucial element for users' perception and interaction behavior logic. Kenneth Craik defined the concept of mental model in the book *The Nature of Explanation* as the small-scale

model to try out various alternatives, conclude the best, react to future, utilize the past knowledge and experience to deal with the present and the future [4].

Jay Wright Forrester defined general mental models as the representation of the real system world concepts and relationships [5]. Susan Carey presented in the paper *Cognitive Science and Science Education* that mental model based on incomplete reality, past experience and even intuition refers to the process of thinking about the way a person works. Mental models help users to form action and behaviors, influence people's concerns in complex situations, and determine the way to solve the problem [6]. Donald Norman showed the relevance of mental models for the production and comprehension of discourse from the user human computer interaction and usability perspective. Norman pointed out the two types of mental models during the design process which are the designer's product concept and user's mental model formed during the user and the product system mapping mechanical [3]. Mental model is the dynamic construction process with time and system comprehension [7]. For the human computer interaction, mental model from users, engineers and designers bridge the gap of perception and interface, and plays the important role in HCI and design refinement [8].

In these views, mental models can be constructed from perception, imagination, or the comprehension of discourse [1], also the mapping media or translation bridge between the reality and the abstract virtual reality or the metaphysical meanings. Consequently, Artificial Intelligence rooted in big data, algorithm, computing and context uses crowds' wisdom to construct the association, prediction and imagination of the objective world, to map the subjective noosphere. The deconstruction and reconstruction between the objective and subjective leads to human-machine new interaction relationship and new mental model, simultaneously influence interaction design. In the traditional driving environment, Human Intelligence (HI) is the dominant factor for HMI (Human-Machine Interaction) which control the driver with the help of Mechatronics and mechanical installations to drive vehicles. However, autonomous vehicles which are driven by Artificial Intelligence (AI) change the dominant function in a disruptive way. AI embedded in autonomous vehicles carries out the main tasks and duties replace of Human Intelligence which include driving plan, workflow and driving command, identity traffic, navigation and wayfinding. Autonomous vehicles driven by AI liberate humans' productive forces from the core tasks of driving behaviors, transform the drivers' status into duality functions. Cognitive redundancy from the drivers and riders pour into entertainment, relax and working through quantified self which is collected driving scenarios big data or user generated contents, profession generated contents and organization generated contents.

As the new HMI and mental models come into being, the paper explores the users mental model transformation in the AI scenarios, comparative research on the main tasks and adjunct tasks with the traditional vehicles, human-machine interaction relationships and mental models originated from users' cognition, perception workflow, interaction model, information architecture, information dimensions, information types processing and media forms. Finally, the paper chooses the voice user interface for autonomous vehicle design as the case study, analyzes the user experience and interaction design based on the new mental model.

The main contributions include:

- System analyze the mental model transformation with the form of AI autonomous vehicles, summarize the cognitive process in the autonomous vehicle driving and riding scenarios.
- Comparative research on the different advantage information types between Human Intelligence and Artificial Intelligence, develop the suitable interaction design dimensions.
- Present Voice User Interface design suggestions and strategies for the autonomous vehicles with the AI mental model.
- Propose the hybrid intelligences mental models which merge the humans' perception, cognition, languages and the diversified animals' sensors or humans social network crowds' wisdom.

2 Related Work

2.1 Autonomous Vehicle

Autonomous vehicle, or self-driving car is the automatic perception of surrounding environment by vehicles equipped with artificial intelligent software and various induction devices, including vehicle sensors, radar, GPS and 2D or 3D cameras. Artificial Intelligence algorithm is used to make a correct driving decision. Hardware which include mechatronics devices and information communication equipment are used to realize the autonomous and safe driving of vehicles, reach the destination safely and efficiently, and achieve the OD goal of completely eliminating traffic accidents. SAE International's new standard J3016 [9] showed the six levels to define the degrees of vehicles automation (see Table 1). The paper researches on Level 4 and 5 autonomous vehicle.

Table 1. Definition of automobile automation levels by SAE (J3016) Autonomy Levels

Autonomous level	Details
Level 0 (no automation)	The full-time performance by the human driver of all aspects of the dynamic driving task, even when enhanced by warning or intervention systems
Level 1 (drive assistance)	The driving mode-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task
Level 2 (partial automation)	The driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task

(continued)

Table 1. (continued)

Autonomous level	Details
Level 3 (conditional automation)	The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene
Level 4 (high automation)	The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task, even if a human driver does not respond appropriately to a request to intervene
Level 5 (full automation)	The full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver

As an intelligent system, autonomous vehicle automatically senses the environment and make driving decisions by the artificial intelligent system which include perception, decision and control, vehicle platform manipulation [10]. Perception sensors and actuators, 3S (GPS, GIS, RS) positioning devices, intelligent transportation system access data of the driving environment and perform the users, vehicles and traffic environment interaction. The automated driving system which include the dynamic driving task, the parking needs, the roadway and the commuter information receives and collects the big data from the transportation elements and scenarios, the system uses the AI algorithm to analyze the travel route and control vehicles mobility. The feedforward and feedback of the human-vehicle, vehicle-vehicle and vehicle-system combine the Human-Vehicle Interaction. In the whole dynamic driving procedure, AI hastens the birth of AI proactive or active driving mode. The paper focuses on the change of the driving mode guided by Human Intelligence and Artificial Intelligence, explores the advantages and disadvantages of the different mental models and information interaction workflows supported by different intelligence.

2.2 Voice User Interface

As the Artificial Intelligence based on natural language interaction influences a variety of computing architectures as multi-cores CPU system, heterogeneous system, distributed system, vehicles composed of AI devices have the ability of hearing and understanding of the users' languages in the state-of-the-art autonomous vehicles. Human-Machine Interaction adopts not only the planned programming language, but also the natural languages to support from the "hands on", "hands off", "eyes off" to "mind off" and "steering wheel optional" autonomous driving concept. HMI two-way communication and interaction support the machine to proactive or active understand the needs of the user and feedback reply. The key technologies of speech interaction include speech recognition, speech synthesis, and semantic understanding. The input speech is converted to text or command respectively, and the text is converted to machine synthesized speech, and the natural language text is transformed into user intention, so that the machine can understand user's needs [11].

In the traditional driving scenarios, voice user interface drastically reduces the distraction and security of the driver for carrying out the driving main task, and distributes the cognition on the secondary task which is related with the non-driving task [1], such like receiving call and music media entertainment. Voice assistant is the most common way of voice interaction. It can lighten the driver's visual burden, reduce the distraction of energy and ensure driving safety.

In today's traditional car market, voice user interaction is widely applied to vehicle navigation. It relies on satellite positioning and real-time traffic monitoring technology to provide drivers with early planning routes, providing speed limit, traffic jam, illegal driving taking evidence photos and other prompts to help drivers successfully drive to their destinations. As a new form of intelligent voice interaction, voice assistant is also gradually applied to the design of vehicle voice system.

Voice User Interface (VUI) provides the new HMI interaction mode and contributes several characters for autonomous vehicles interaction design.

- Hearing faster wakes up the user's attention [2], just like the athletes starting shots. Compared to the visual image, voice as the low dimension information, become the users' cognitive focus much more simple than graphic user interface and highlight the mental model theme, however, visual information as the higher dimension information, carries non-focus and diversified information storyline according to different perspectives which include first-person perspective, second-person perspective and third-person or God's perspective. Different visual perspective causes misunderstanding of the driving tasks and distracts the users' attention.
- Voice information sets up personification mental model and eliminates user interaction psychological barriers. To compare with visual information, VUI is better than Graphic User Interface (GUI) for personification association between the strange or unfamiliar field and the past user experience knowledge. Nevertheless, excessively personification results in high user psychological expectations. When driverless vehicle is executing tasks, the usability targets for the interaction design, such as performance, accuracy and stability will damage user experience in turn.
- VUI supports space free multi direction interaction for the directionless longitudinal wave, Different from GUI multi-users interaction through multi-touch visual information, voice communication in the context of AI from multi-users simultaneously interaction produces cocktail party effect [3]. Voice interaction obstacles come from the multi-users identity and tasks in time non-interruption implementation. However, voice communication with Human Intelligence can skillfully divert attention with the selective auditory attention.
- Voice information belongs to opening structure media from the aspect of cold media and hot media, and seldom influenced by visual information metaphor. For this point of view, VUI is less dominated by the interaction designer's mental model.
- VUI decreases the information architecture and hierarchy, supports task skip and switch. the visual information in GUI is restricted by space limitations, requires a series of visual information to transact a complex task in the information hierarchy, voice interaction is not dissipative from information space, and can be directly formed the dialogue or speech task, jump directly from information sequence by overlooking the time sequence constraint.

The paper mainly discusses voice user interface interaction design in the autonomous vehicles. The user in the self-driving cars requires no active control of the dynamic driving interaction behaviors as the change of artificial intelligence guidance of the primary and secondary tasks, HMI interaction pattern with artificial intelligence engenders the design method of voice user interface with the information processing advantage.

3 Comparative Study Between Traditional Driving and Autonomous Driving

3.1 Main and Secondary Tasks Changes from the Traditional Vehicles and Autonomous Vehicles

In traditional driving context, the relationship between the user and vehicle is manipulated unilaterally by the drivers and riders. The main task of the driver is to comprehensively analyze the information of vehicle’s movement state and traffic condition, and make the correct driving strategy and make corresponding driving action [1]. However, in the context of autonomous vehicle dynamic driving and human-vehicle interaction scenarios, sensors and artificial intelligence algorithm take the place of the drivers functions to do the driving context awareness computing and riders consciousness awareness computing. Most driving main and secondary tasks transform into the AI and HI hybrid intelligence interaction scenarios which have no tasks focus and user experience storytelling. Figure 1 lists the change of the main task and the secondary task in details.

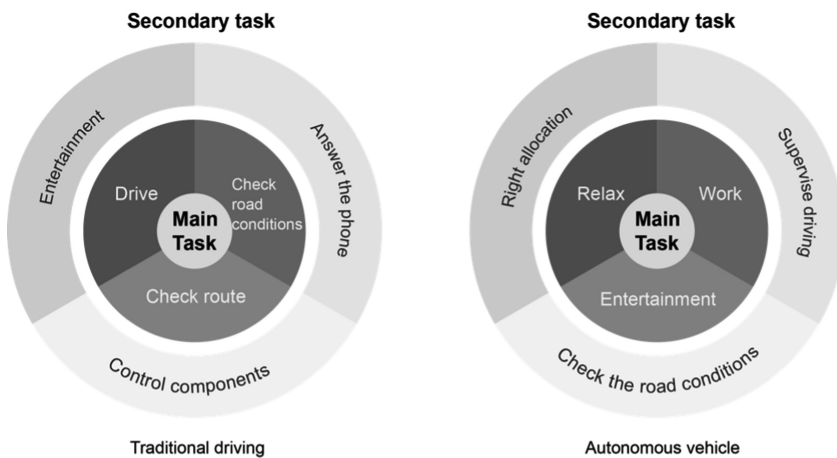


Fig. 1. Comparison of the main or primary and secondary tasks in traditional driving environment and autonomous vehicle driving environment.

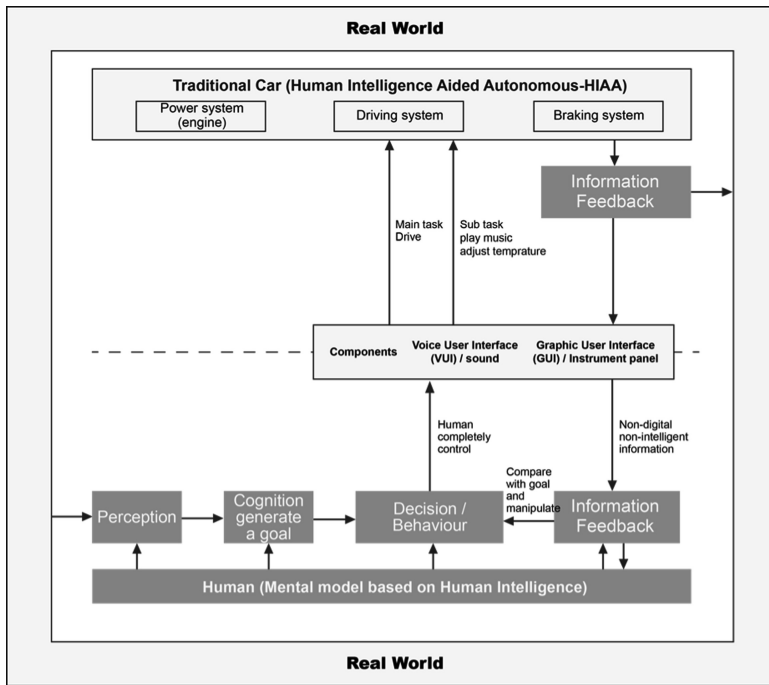
3.2 Mental Model in the Mode of Autonomous Vehicles Driving Context

Based on the comparison of main and secondary tasks, the HMI interaction logic between the user and vehicles based on traditional driving cannot fully understand and meet the new form of HMI interaction in the driverless environment. When users interact with unmanned vehicles, new mental models which combine human intelligence mental models (users, engineers, designers, etc.) and artificial intelligence (pattern, algorithm, sensors principles from animals or crowds' wisdom) will be generated to adapt and learn new systems. For designers, only by fully understanding the user's mental models, can we build an accurate interaction system model and interaction paradigm, proper information architecture, workflow, navigation, information hierarchy, interface and media visual design, make the design method fit with the user's mental models, eliminate users' doubts and face with new products, and provide a hybrid seamless user experience.

As to the mental model construction methods, some scholars presented different ways and perspectives to analyze. Indi Young thought mental model is consisted of several parts, each part is divided into groups, and the whole model can be used a series of behavior affinity diagrams [4]. Waern suggested that there are two approaches to construct mental models which depends on whether or not the learners have prior knowledge about the system. The bottom-up approach is used by learners who react to incoming bits and pieces of information, interact with the system, and gradually build a more consistent and complete mental model upward. Most users choose the top-down approach to evoke the learners existing knowledge, modify and adjust the mapping relationship and reconstruct it into a new mental model according to the information they perceive as they interact with the system. Expert users and novice programmers or learners prefer the different mental model types [5]. The construction of mental models is mostly through psychological experiments to find out some general mental models in a statistical way. As a new research direction, autonomous vehicle is a new application for the research direction, and the experienced subjects or experiments are relatively small data, which fails to form user research with large sample size. This is also the defect of this paper. Therefore, the paper studies from the past experience based on the analysis of user perception of the traditional car and the system interaction model. According to the conceptual model of unmanned vehicles, the paper analyzed the cognitive process to infer the user interaction with the drones, and then summed up the mental model under the new situation, focusing on a variety of mental models and external environmental information process.

In the traditional driving environment, the concept model of traditional automobile is made up of power system, driving system and braking system. The relationship between man and vehicle is driven by the users' single side, and the cognition process of human interaction on vehicles is based on the users' perception of the external world, and the perception information acquired by cognition forms the goal and then makes dynamic driving or riders decision-making behavior. Through the voice and graphical user interface for human vehicle interaction, instruction is conveyed by the present humans. Traditional cars are then informed by information feedback, such as lighting and voice, to compare the results of users' actions with goals, and adjust the behavior again, forming a cycle process. The whole process is supervised by the human

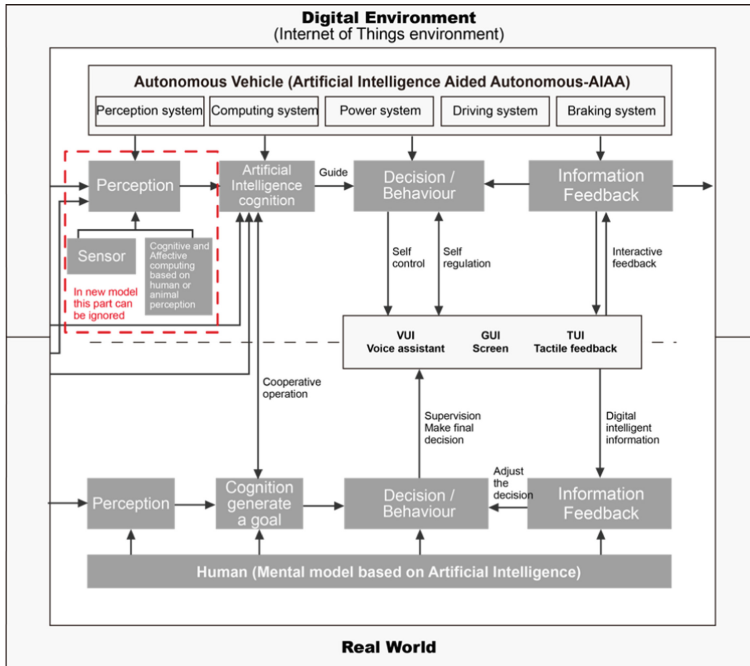
intelligence driven by the human’s mental model, which is consistent with human cognition, and then reconstructs the mental model based on information feedback, and finally generates the correct model. Figure 2 shows the cognitive process of driver’s interaction to human and car in the traditional driving environment. Taken the cognitive cycle of decision making model as the reference mentioned by Connolly and Wagner [6], the paper constructs the HMI cognition process. Through the decision cycle, decision-makers construct the cognitive perceptions of reality, complete the knowledge accumulation, adopt reasonable execution behavior, after its feedback effect on the real world’s environment, the external environment being reformed or changes in turn is recognized or understood by the decision makers again. With the establishment of new knowledge, the new mental model guides the users’ behaviors again and constitutes a cycle of cognitive decision making.



The cognitive process of driver's interaction between human and car in the traditional driving environment

Fig. 2. The cognitive process of driver’s interaction between human and car in the traditional driving environment.

The paper points out the change of the mental model based on the traditional driving environment cognition process. As shown in Fig. 3, in the autonomous vehicle dynamic driving environment and scenarios, self-driving cars are controlled by the Artificial Intelligence, Human-Machine Interaction relationship transformed into human vehicle symbiosis co manipulation. AI is different from HI and the crowds’



The cognitive process of HMI interaction in autonomous vehicle dynamic driving environment

Fig. 3. The cognitive process of HMI interaction in autonomous vehicle dynamic driving environment

wisdom result from the swarm intelligence algorithm. With the function of big data, deep learning from visual and voice computing, AI in HMI leads to new cognition process and mental model. In the new cognition process, the relationship between the perception based on sensors and the cognition based on mental models is changed dramatically. Cognition can be set up without senses perception. Sensors from the animals' senses principles which include the visual, the sound or the touch can be translated into different waves which are not sensed by humans. The new perception produces new mental model for the cognition, consequently change the logic sequences and mental model mode between the perception and cognition. Diversified mental models lead to rich types of Human-Machine Interaction modes and models. In the new cognition process, autonomous vehicles are empowered by the AI algorithm from the high-performance sensors which receive the intelligent transportation system and the Internet of Things environment. Merging with the human six senses or animals living things senses information, the driving environment which combines the driving context awareness, consciousness awareness and emotion awareness comprehensively form the Artificial Intelligence cognition. Based on big data and new AI HI hybrid cognition information, autonomous vehicles can actively or proactively make the driving decision making, self- control the driving behavior and make the route strategy. Human

intelligence supervises the driving scenarios and partially participates in the decision making. The HMI interacts with the voice, images, visual and tangible user interfaces to proceed the information exchange, feedforward, feedback and adjust the decision made by AI, finally fulfills the safety stable driving behaviors.

In the autonomous vehicle driving environment, there are two different mental models which include Human Intelligence and Artificial Intelligence mental model. The two mental models interact with each other and impact on the external environment simultaneously. Taken the double-loop learning process diagram as the reference, the paper presents the learning process new mode in the mental model in Fig. 4 and represents the impaction of the different two types of mental models on the strategy and principles after the feedback of information reaction respectively. As AI deals with the data from every possible channel from human environment and animals' senses simulation, mental model made of AI cognition goes straight forward the external environment and cognition results without human perception. The human prior knowledge, information-processing styles and universal common senses intelligence in some situations plays no role on the mental model which is dominated by AI. Some scholars research on the AI simulation on the animal intelligence. For some aspects, humans' brains structure is not as good as the brain structure of other animals. And artificial intelligence not only simulates human brain intelligence, but also simulates the intelligence of other animals [7]. In the autonomous vehicles image guidance, the compound eye structure of the fly vision system is simulated, which enables the seeker to realize the 360° search field that the human eye vision system does not possess [8].

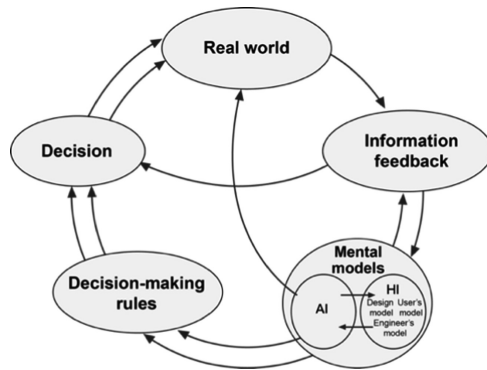


Fig. 4. double-loop learning process and double intelligence mental models.

When AI develops the new mental models on the advanced stage, intelligence learns not only from the knowledge and increases the existing mode, but also has the original creation ability to build new mental model with direct cognition without perception. In the single human intelligence dominant mental model, the learning process consists of two translation procedure which includes translation of the external environment signal data into visual or sound information and GUI or VUI translated into command interaction behaviors manipulated by humans. However, in the new

mental model which is developed by Human Intelligence and Artificial Intelligence, autonomous vehicles can use other creatures such as Hawkeye, perception of the compound eyes of flies, Pipi shrimp eyes from the second person perspective to perceive the world from the multi-dimensional perspectives while building the cognitive style, the hybrid intelligence mental model will also change the perception which omits the middle part to translate among the meta data, humans, vehicles and the environment. Nature can directly use meta information and metadata to the problem solving and driving context and do not need to translate for humans' cognition or perception. However, the mental model for the driving and learning process not only depends on the AI and other information communication technology, but also takes the non-technical problems into account, such as the SNS behaviors, social responsibility, profit distribution, in particular people still need informed consent and information feedback during the dynamic driving environment when riding in the autonomous vehicles. Taking the unmanned vehicle assessment regulations into the consideration, there is no need to intervene in the cognitive traffic law, the evaluation of executive ability, and the disposal of emergency channels, while in the joint decision-making of comprehensive driving ability, it is necessary to participate in the decision making together with humans and AI vehicles.

3.3 Comparison of Human Intelligence and Artificial Intelligence

In traditional driving and self-driving scenarios, there are two mental models that are guided by Human Intelligence and Artificial Intelligence. The gradual gain of human intelligence in the struggle between human and nature is the result of labor. Through continuous practice and the accumulation of millions of years of evolution and experience, human knowledge has been accumulated [9]. Human intelligence is an ability to solve the problem by using of the knowledge and experience to learn new knowledge, concept and ideas [10]. People can perceive the external world through their eyes, ears, nose, tongue and mind consciousness. And Artificial Intelligence is the science and technology that uses machines to simulate human thinking in order to expand and extend human brain intelligence [11]. The development of AI based on brain cognitive science, from machine recognition to pattern recognition, from natural language processing and understanding to knowledge engineering expert systems, from knowledge patterns to a new cognitive logic, from unstructured data to structured smart contents [12]. According to the comprehensive literature, human intelligence is better at dealing with fields that need intuition, inspiration, insight and creative thinking, while AI is better at carrying out reasoning and computing as the main way of thinking and decision-making. Different intelligence needs the corresponding mental model.

In Human Vehicle Interaction, because of the inherent advantages of human intelligence, we can do any dimension switching in the 11 information dimensions, and achieve cross channel perception. But artificial intelligence has no rich information dimension and limited with the three-dimensional spaces, can only switch in finite dimension. Therefore, in contrast to the difference between human intelligence and artificial intelligence in Human-Vehicle Interaction, the difference is concentrated in the one dimension (Voice User Interface), two dimensions (Graphic User Interface) and three dimensions (Tangible User Interface). In these three dimensions, there are

corresponding channels of perception, namely auditory, visual and olfactory, which are represented by one dimension speech user interface, two-dimensional graphical user interface and three-dimensional entity interface. Table 2 shows the advantages and disadvantages of human intelligence and artificial intelligence in dealing with these three dimensions' information with the mental models.

Table 2. Information dimension comparison between AI and HI.

Dimension	Sense channel	Interface	Information process characters with Human Intelligence	Information process characters with Artificial Intelligence
1 Dimension	Auditory	VUI	1. It relies heavily on time, single thread, and irreversible	1. Neglecting time and processing multiple sounds at the same time
			2. Support single focus without cocktail effect	2. Support multi focus with cocktail effect
			3. Uncertainty	3. Certainty
			4. Beyond information mapping and transformation in the 11 information dimensions with imagination	4. Need to information map one by one
2 Dimensions	Vision	GUI	1. The human eye has the persistence of vision	1. Support not only individual human eyes but also the three person perspectives and crowds' wisdom eyes, even the animal perception which is different from humans sight
			2. Perceptible temperature, that is, cross channel perception and feel the subtle space information	2. The algorithm can achieve fast comparison and processing of the past information in time
3 Dimensions	Tactile sensation	Tangible User Interface	1. Affordance	1. Born data
			2. DNA evolution	2. Feeding data
			3. Hybrid reality and virtual reality in the three spaces (cyberspace, information space and noosphere)	3. Artificial intelligence generally deals with virtual bodies and is not good at dealing with entities

Based on the above comparison, it is found that AI has advantages over human intelligence in dealing with voice information, and there is no difference in the processing of graphic information between AI mental model and human intelligence mental model, but the processing of entity tangible interface information is far from human intelligence. In the design of the Human-Vehicle Interface interaction design, the interaction design of voice user interface can be emphasized in the “eyes off, hands off, and mind off” autonomous vehicle HMI interaction design. And the special talents of human intelligence and artificial intelligence can be played together.

4 Suggestions for the Autonomous Vehicle Voice User Interface Interaction Design

4.1 Scenarios for the Autonomous Vehicle VUI

According to the third section, the change of main task and secondary task leads to mental model transformation. The driver’s main task has changed from driving to entertainment, resulting in a lot of cognitive redundancy, which liberates people’s attention and is suitable for increasing various forms of entertainment and improving user experience, especially the driver’s function change from the driving behavior to riders’ function, the user experience and interaction design focus on the vehicle data and system service. In human vehicle interaction, the way of voice interaction is simple to wake up, the level of interaction can be simplified, and it can directly reach the decision, alleviate the visual load and user perception easily. The following will reflect the design recommendations of the voice user interface from the state of emergency and safe driving.

In a safe running state, the user can liberate the hands for entertainment by taking the active control of an autonomous vehicle. Research has shown that the highly automated system reduces the participation of the driver, the driver into the role and function of passive monitoring vehicle auxiliary system from the active control of the role of the driver in low load state, easily out of control loop not to respond immediately to emergency situations [1]. The application of the voice user interface can ensure that the user can also understand the state of the autonomous vehicle without special attention. This is because compared with the visual channel, hearing is easy to attract people’s attention at any time, and has the advantages of fast reaction speed and unrestricted lighting conditions. VUI is suitable for use in an emergency dynamic driving context [2]. The information processing of the autonomous vehicle based on cloud computing and AI algorithm. It is not necessary for the VUI interaction design to present all the complex digital language to the end user, only need to run the state by visual and auditory feedback to the user, do an “informed but not overly burdensome” information communication in case of emergency, artificial intelligence can make the most rapid and accurate judgment, but due to the distribution of rights and responsibilities of the autonomous vehicles driving problem, the design needs to inform the current state of the user. At this time, the voice user interface should arouse the attention of users in the entertainment state by launching noise and lighting changes, and switch to voice and visual form to inform users of the current driving state.

4.2 The Future Trend of VUI

The development of artificial intelligence is becoming more and more powerful, and it is believed that in the near future it will be able to cross the current restrictions and become more intelligent. At present, artificial intelligence imitates human intelligence, which is mentioned in the previous article. In practice, artificial intelligence can also simulate the intelligence of other animals. In this new mental model, artificial intelligence can also be judged by the combination of multiple senses. Abe Davis researched on the visual microphone to passive recovery of sound from video [13], some blinders even use sound to enjoy the pictures by translation of image into sound. Autonomous vehicles can directly obtain the metadata from the environment to make decisions based on the detection of auditory and visual channels. There will be a great change in the voice interaction process of the human vehicle. The former pattern can be concluded that when AI is learning the thinking mode of animal intelligence, it can skip the recognition channel and perception process, directly from concept to recognition decision and cognition. In the old mental model, the information is translated to the user by machine recognition and then translated to the user secondly. In the new mental model form, the interaction process can be simplified. The autonomous vehicle, by identifying different driving concepts, does not need to inform the user of all process information directly. However, taking into account the responsibility and rights of the dangerous situation, the issue of responsibility and rights mentioned in the previous section needs a visual auditory feedback state for the Human-Machine Interaction.

Under the guidance of artificial intelligence, information can be extracted in a variety of sound confounding. Under the guidance of human intelligence, the dimension of sound and vision can be switched. GUI and VUI transforms according to different driving context with the support of HI and AI mental models. The use of artificial intelligence to hear the sound judgment of danger, the propagation of sound than can be used in artificial intelligence as the end, determine the object state by detecting sound or visual. Visual and auditory switching, which way should we take to acquire multiple voices, can also separate voice, and also can analyze the content of current nodes with visual wiretapping. The highest level of speech interaction can translate the language of the deaf and dumb people, VUI interaction design don't need to collect and compute the users' behavior pattern, finally understand the users' intentions which are told or shown by users, but only set up the OD information to go straight forward to the goal.

5 Conclusion

The mental model is influenced by perception and cognition. Artificial Intelligence and Human Intelligence develop the different mental models separately. Mental model decides the strategy and principle planning. In the autonomous vehicle driving context, AI mental model and HI mental model make the hybrid decision making with the function of perception and cognition. The interaction design takes the HI and AI interaction mode into account based on the different mental models.

Autonomous vehicles liberate humans' hands, eyes and mind off the dynamic driving context and support the Voice User Interface to deal with the users' cognition redundancy. The tradition mental model roots from the sequence of algorithm to data to mode and to behaviors, which is based on Human Intelligence. The new mental model based on Artificial Intelligence encourages the second-person perspective to simulate not only the human like but also the other creatures' perception and cognition modes, finally creates new thinking mode and mental model to do the interaction design. Mental model leads to VUI interaction design method beyond the tradition one and provides all kinds of possibilities for the interaction design.

References

1. Johnson-Laird, P.N.: *Mental Models: Towards a Cognitive Science of Language, Interface, and Consciousness*. Cambridge University Press, Cambridge (1986)
2. Kieras, D.E., Bovair, S.: The role of a mental model in learning to operate a device. *Cogn. Sci.* **8**(3), 255–273 (1984)
3. Norman, D.A.: *The Design of Everyday Things*. Basic Books, New York (2002)
4. Johnson-Laird, P.N.: The history of mental models. In: Manktelow, K., Chung, M.C. (eds.) *Psychology of Reasoning: Theoretical and Historical Perspectives*, pp. 179–212. Psychology Press, New York (2014)
5. Wiki, 7 February 2017. https://en.wikipedia.org/wiki/Mental_model
6. Carey, S.: Cognitive science and science education. *Am. Psychol.* **41**(10), 1123 (1986)
7. Nakatsu, R.T.: *Diagrammatic Reasoning in AI*. Wiley, Hoboken (2009)
8. Jih, H.J.: Mental models: a research focus for interactive learning systems. *Educ. Technol. Res. Dev.* **40**(3), 39–53 (1992)
9. Standard SAE: J3016: SAE international taxonomy and definitions for terms related to on-road motor vehicle automated driving systems, levels of driving automation (2014)
10. Shi, W., Alawieh, M.B., Li, X., et al.: Algorithm and hardware implementation for visual perception system in autonomous vehicle: a survey. *Integr. VLSI J.* **59**, 148–156 (2017)
11. Yuan, B., Xiao, B., Hou, Y., et al.: State-of-art and trend of speech interaction for mobile intelligent terminal. *Inf. Commun. Technol.* **2**, 39–43 (2014)
12. Duan, L.: *Driver mental workload evaluation and application in driver assistance system*. Ph. D., Jilin University, China (2013)
13. Lv, Z.: *Human Engineering*. China Machine Press, Beijing (2016)
14. Arons, B.: A review of the cocktail party effect. *J. Am. Voice I/O Soc.* **12**(7), 35–50 (1992)
15. Young, I.: *Mental Models: Aligning Design Strategy with Human Behavior*. Rosenfeld Media, New York (2008)
16. Waern, Y.: On the dynamics of mental models. In: *6th Interdisciplinary Workshop on Informatics and Psychology: Mental Models and Human-Computer Interaction*, vol. 1, pp. 73–93. North-Holland Publishing Co., Amsterdam (1987)
17. Wang, Y.X., Guenther, R.: The cognitive process of decision making. *J. Cogn. Inform. Nat. Intell.* **1**(2), 7–85 (2007)
18. Liu, J., Liu, X.: The relationship between artificial Intelligence and human brain Intelligence. *Research on Noetic Science in China 2011 Album*, pp. 850–853(2012)
19. Wang, G.: *Research on IR imaging guidance using fly's visual system*. Ph.D., Department of College of Astronautics Northwestern Polytechnical University (2003)
20. Liu, Q.: Human intelligence and artificial intelligence. *Comput. Sci.* **19**(2), 55–59 (1992)

21. Wang, L.: The relationship between artificial intelligence and human intelligence. *Technol. Innov. Appl.* **31**, 76 (2016)
22. Tian, Y.: *Dictionary of Thinking Science*. Zhejiang Education Publishing House, Zhejiang (1996)
23. Qin, J.: Impaction of artificial intelligence on interaction design. *Packaging Eng.* **38**(20), 27–31 (2017)
24. Davis, A., Rubinstein, M., Mysore, G.J., et al.: The visual microphone: passive recovery of sound from video. *ACM Trans. Graph.* **33**(4), 79 (2014)