Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/31/20230125

New perspective on human-computer interaction-based on speech recognition in driving

Honghao Xue

School of Computer Science and Technology, Guangzhou Xinhua University, Dongguan, China

xuehonghao@stu.xhsysu.edu.cn

Abstract. With the popularity of smart driving and in-vehicle intelligent systems, speech recognition plays an important role in driving as a key link of human-computer interaction. However, although speech recognition systems have been widely used, their application inhuman-computer interaction still needs to be improved. The purpose of this paper is to examine or review the latest perspectives on speech recognition as part of human-computer interaction in driving. lt provides an in-depth understanding of the potential and future development of speech recognition in driving by exploring, among other things, new directions and features that can be improved. Through a comprehensive analysis of relevant literature and research results, this paper will provide an overview of the current state of speech recognition in human-computer interaction, point out the challenges and limitations that still exist in the driving environment, and focus on new perspectives on speech recognition in driving that may enable more highly accurate command recognition and natural conversational interaction to enhance the driving experience and safety through the use of advanced speech recognition technologies. Explore how technologies such as artificial intelligence and machine learning can be used to drive the development of speech recognition and address current challenges. ln addition, research advances in the areas of driving behavior analysis, emotion recognition, and personalized driving related to speech recognition will be explored. The review and analysis in this paper will provide valuable references and guidance for further research and development of speech recognition-based driving human-computer interaction systems.

Keywords: human-computer interaction, speech recognition, driving, machine learning.

1. Introduction

With the popularity of smart driving and in-vehicle intelligent systems, speech recognition plays an important role in driving as a key link in human-computer interaction. Speech recognition technology can transform the driver's voice commands into understandable text or operations to achieve a natural and direct interaction method, which greatly enhances the convenience and safety of driving [1]. From the initial simple command recognition to today's natural language dialogue interaction, speech recognition systems have made great strides in driving human-machine interaction [2]. The Figure 1 showed the status of speech recognition in machine learning and the structure of this paper.

DOI: 10.54254/2755-2721/31/20230125

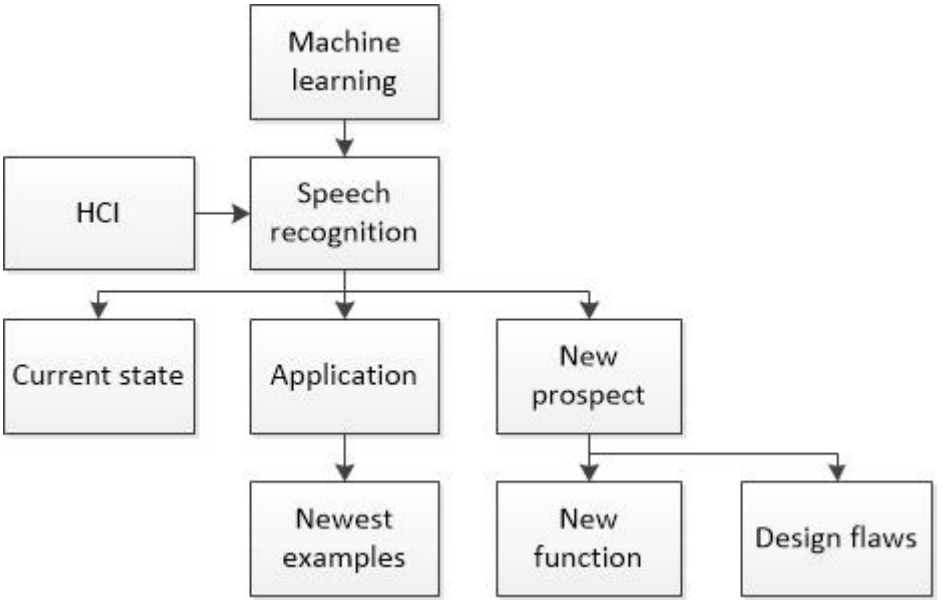

Figure 1. Structure of speech recognition in HCI.

Speech recognition technology has been widely used in driving human-computer interaction, and some important progress has been made. By applying voice recognition technology to in-vehicle intelligent systems, researchers have achieved voice control of navigation, entertainment systems, telephone communication and other functions [3]. The driver can easily control various operations of the vehicle with simple verbal commands, avoiding tedious manual operations and greatly enhancing the convenience and safety of driving [4]. In addition, voice recognition technology can also be applied to the driver status monitoring, through the analysis of the driver's voice characteristics and language expression, to identify their emotional state, fatigue level, etc, to provide more protection for driving safety [5].

However, despite the achievements of speech recognition in driving human-computer interaction, there are still some challenges and limitations. The accuracy of speech recognition systems still needs to be improved in noisy driving environments. Factors such as noise, speech variation, and in-vehicle sounds may interfere with the results of speech recognition, reducing the reliability of the system [6] . To solve this problem, researchers have proposed some new perspectives and approaches. One of them is to use deep learning techniques to improve the performance of speech recognition systems. Deep learning algorithms can be trained with large-scale data to automatically learn abstract features of speech and thus improve recognition accuracy [7]. In addition, researchers have explored ways to combine speech recognition with other sensor data, such as using microphone arrays and in-car cameras in the driving environment to improve the accuracy and stability of voice commands [6].

2. Literature review

2.1. Status of speech recognition technology

In recent years, with the popularity of smart driving and in-vehicle intelligent systems speech recognition plays an important role in driving as a key link in human-computer interaction. Y Chen's research explores the design and implementation of an in-vehicle speech system based on intelligent interaction [8], which achieves efficient and accurate speech recognition and command execution by combining natural language processing and machine learning technologies. This research provides a useful reference for the development of the field of driving human-computer interaction. In addition, Y Liu's research reviewed the current development status and trends of multimodal interaction technologies for intelligent cockpits [9]. They pointed out that speech recognition, as an important

DOI: 10.54254/2755-2721/31/20230125

interaction method, has a broad application prospect in driving. Through the combination of multimodal interaction technologies, a more intelligent and personalized driving experience can be achieved. The research by Z Mohan focuses on the human-computer interaction design of in-car entertainment systems and explores the application of speech recognition in entertainment functions [10]. Through voice interaction, drivers can easily operate the in-car entertainment system to enhance driving enjoyment and comfort. The study further emphasizes the potential and importance of speech recognition in driving human-computer interaction.

Y Yu's research focuses on the development of human-computer interaction systems for in-vehicle navigation devices, and explores the application of speech recognition in navigation functions [11]. Through voice navigation commands, drivers can more easily access route information and improve the safety and efficiency of driving. This study further demonstrates the practical application scenarios of speech recognition in driving human-computer interaction.

2.2. Application of artificial intelligence and machine learning in speech recognition

The application of artificial intelligence and machine learning technologies in the field of speech recognition continues to drive the development of voice interaction. Z Rongxiang's report points out that Telematics has constructed a new world of voice technology applications [12], and advances in artificial intelligence technology have enabled voice recognition to more accurately understand drivers' commands and achieve more natural and fluid human-machine dialogue.

H Shengfan's research focuses on the design of speech-based in-vehicle information systems for intelligent mobile terminals [13]. He explores the application of artificial intelligence and machine learning in speech recognition, and achieves high-precision speech recognition and information retrieval through training and model optimization of large amounts of speech data. This research provides technical support and methodological guidance for the application of speech recognition in driving.

2.3. New perspectives on speech recognition in driving

In addition to the traditional command recognition and execution functions, speech recognition has many new perspectives and applications in driving. A study by Z Mohan [10] pointed out that speech recognition can be used for driving behavior analysis. By analyzing the driver's voice commands and emotional expressions, the behavioral characteristics and needs of the driver can be better understood, further enhancing the personalization and intelligence of driving. Speech recognition also plays an important role in intelligent transportation systems. W Zilang's research J Park [14]. explores the design and implementation of a speech recognition-based traffic information acquisition system. Through speech recognition technology. drivers can obtain real-time traffic conditions and road information through simple voice commands, providing a more convenient driving experience. In addition, with the development of intelligent assistant technology, speech recognition is playing a more comprehensive role in driving. H Li's research [15] explores the design and implementation of a speech recognition-based intelligent driving assistant system. Such a system can understand the driver's commands and needs through speech recognition and natural language processing technology, and provide intelligent driving assistance functions such as navigation, entertainment, and climate control to enhance driving convenience and safety [16]. Figure 2 displayed multiple layers of the speech recognition in driving, it has combined both new structures and Chen's study [8].

DOI: 10.54254/2755-2721/31/20230125

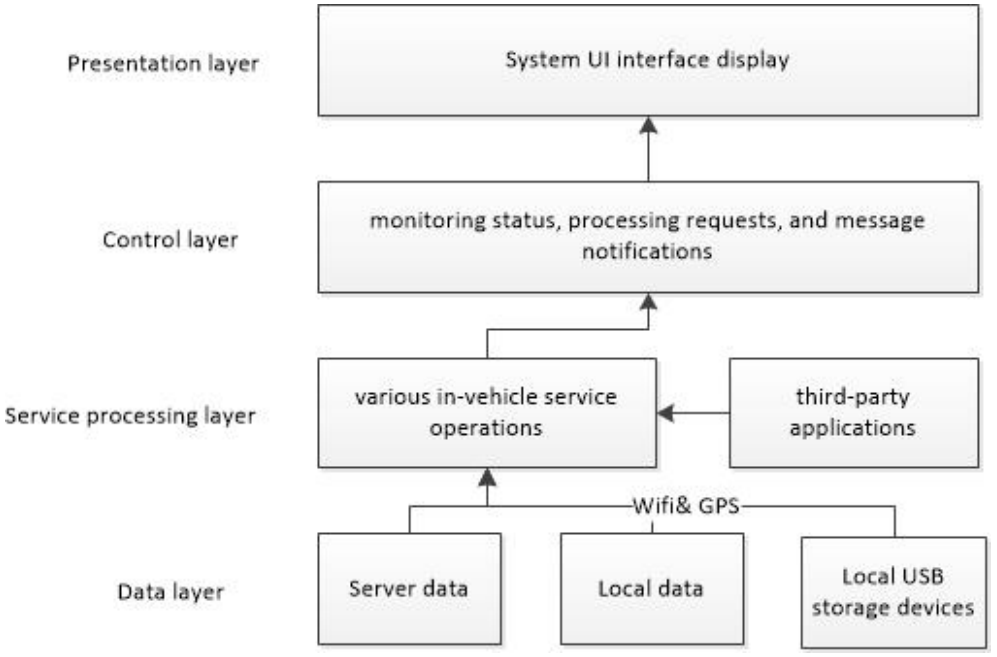

Figure 2. Previous layers of speech recognition in driving.

3. Results and Discussion

3.1. The latest application cases of intelligent speech recognition technology in driving

In the field of intelligent driving, the use cases of intelligent voice recognition technology are increasing. The following are some of the latest application cases:

In-vehicle voice assistant: The application of intelligent voice assistant system in vehicles continues to improve. Drivers can use voice commands to achieve a number of operations, such as navigation destination setting, music playback control, call management, etc. These voice assistant systems can accurately identify and execute the driver's commands, providing a convenient human-computer interaction experience.

Driving behavior analysis: Intelligent speech recognition technology can be used for the analysis and evaluation of driving behavior. By analyzing the driver's voice commands and voice emotion expressions, the behavioral characteristics and needs of the driver can be understood. This is important for personalized driving experience and the development of intelligent driver assistance functions.

Intelligent Transportation System: Voice recognition technology plays a key role in the intelligent transportation system. Drivers can use voice commands to obtain real-time traffic conditions, road information, etc., providing a more convenient and safe driving experience. In addition, voice recognition can also be used for driver interaction with intelligent transportation facilities, such as intelligent parking systems, intelligent street lights, etc.

3.2. Research progress and results of new function development

For the development of new functions of intelligent speech recognition technology. researchers have made some important progress and results. The following are some research directions and related results:

Multimodal interaction: Researchers have worked to combine speech recognition technologies with other sensors and interaction modalities to enable multimodal interaction. For example, combining speech recognition with technologies such as gesture recognition and eye-tracking provides a richer and more natural way of driving human-machine interaction.

Emotion recognition: Emotion recognition is an emerging research direction that aims to recognize the driver's emotional state through speech recognition technology. This has important implications for

DOI: 10.54254/2755-2721/31/20230125

personalized driving experience and emotionally intelligent interaction. Some research has already begun to explore the application of emotion recognition in driving to provide more attentive service and support for drivers.

3.3. Results of the empirical study design flaws on driver distraction

Researchers have also conducted empirical studies on the effects of design flaws on driver distraction. The following are some of the relevant research findings:

Sound interference: Some design flaws may cause drivers to experience sound interference during voice recognition interactions, thereby reducing recognition accuracy and execution. It has been found that noise, music and other distracting sounds may interfere with the recognition and understanding of voice commands, and, if the intelligent system does not provide accurate or incorrect feedback on voice recognition, it will prompt the driver to speak repeatedly for recognition, increasing the risk of driver distraction.

Interaction Interface Complexity: Overly complex interaction interface designs can distract drivers during operation. Studies have shown that interface designs that are not ergonomic and cognitive for driving can lead to driver distraction and increase driving safety risks.

3.4. Discussion and analysis of results

Based on these results, the use cases of intelligent speech recognition technology in driving continue to enrich, including in-car voice assistants, driving behavior analysis, and intelligent traffic systems. Researchers have also made important progress in the development of new features, such as multimodal interaction and emotion recognition. However, results from empirical studies of design flaws on driver distraction suggest that some factors such as voice interference and interaction interface complexity may negatively affect driver attention and safety.

Analyzing these results together, we can conclude that intelligent speech recognition technology has a wide range of promising applications and research potential in driving. However, when developing new features and designing interactive interfaces, driver distractions need to be taken into account and measures need to be taken to improve driving safety and convenience. Further research can focus on how to optimize speech recognition algorithms, improve recognition accuracy, and explore the integration of intelligent speech recognition technology with other smart driving technologies to provide more possibilities for the development and application of future smart driving systems. The following figure showed the process of human-computer-interaction and new interaction technologies. Figure 3 showed a model of a normal human-machine interface. Figure 4 displayed the structure of a ordinary multimodal interaction.

DOI: 10.54254/2755-2721/31/20230125

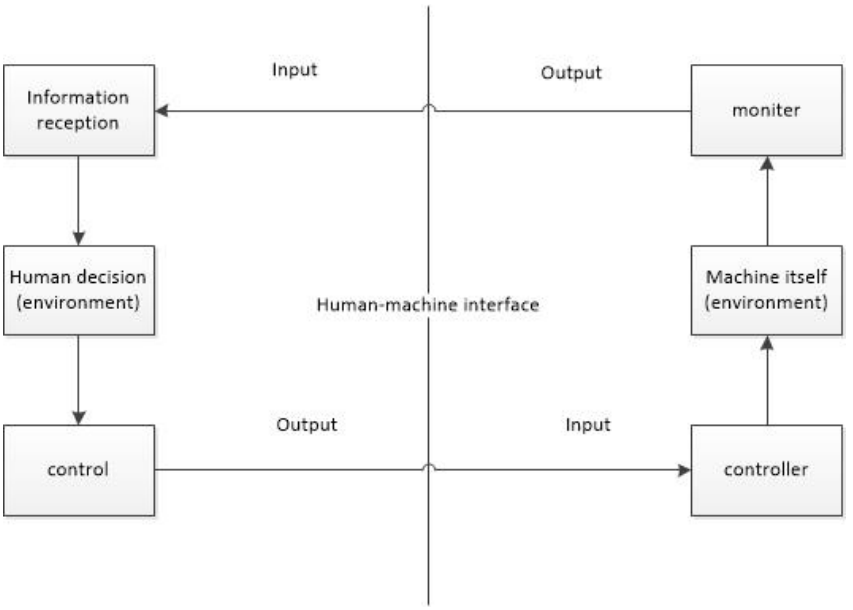

Figure 3. The process of Human-machine interface.

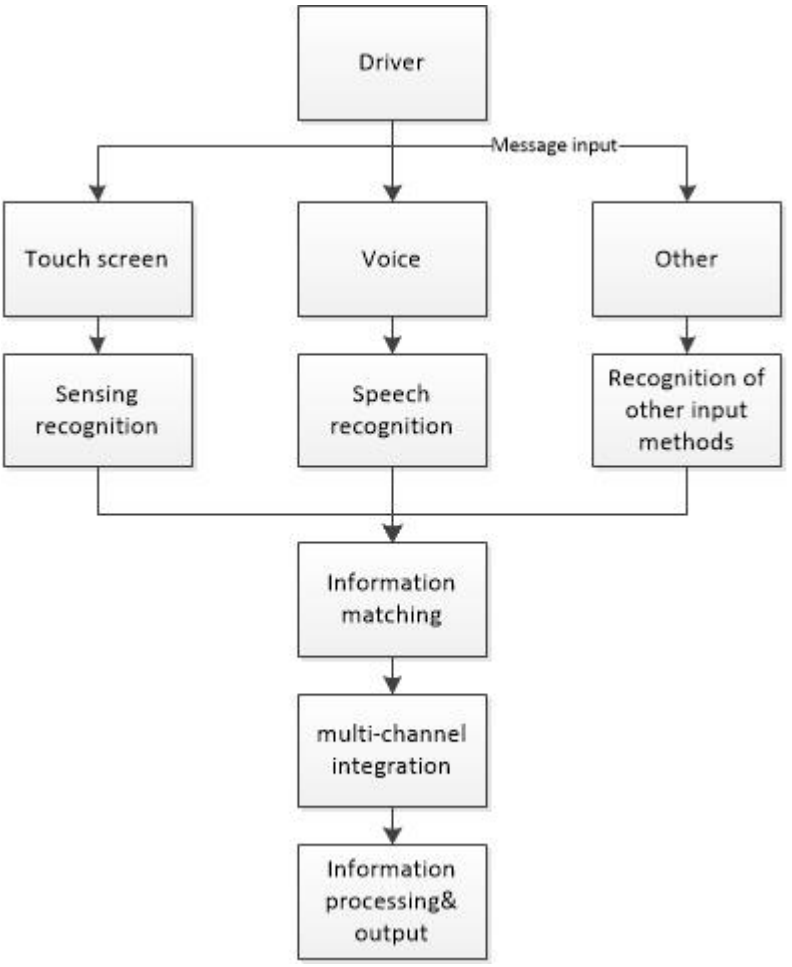

Figure 4. The structure of multimodal interaction.

DOI: 10.54254/2755-2721/31/20230125

3.5. Future outlook

In the future, speech recognition will continue to play an important role as a key aspect of human-computer interaction. The following is an outlook on the future development of speech recognition in driving:

improve the accuracy and stability of speech recognition: Researchers can continue to explore and improve speech recognition technology, using deep learning algorithms and other advanced technologies to improve the accuracy and usefulness of speech recognition. Also, combine other sensor data and environmental information to optimize the speech recognition system to cope with noisy driving environments.

Driving multimodal interaction: Combining speech recognition with other sensors and interaction modalities to achieve multimodal interaction is a promising research direction. By combining multiple interaction modalities such as voice, vision and gestures, drivers can interact with vehicles more freely and provide a richer and more intelligent driving experience.

Combined with the development of artificial intelligence and machine learning: the application of artificial intelligence and machine learning technology in speech recognition will continue to be enhanced. With the increasing amount of data and optimization of algorithms, the accuracy and stability of speech recognition systems will be significantly improved.

Advances in artificial intelligence and machine learning have brought a higher level of understanding and processing power to speech recognition. Traditional speech recognition systems rely heavily on pattern matching and statistical models, but with the rise of deep learning and neural networks, systems are able to learn from large amounts of speech data and extract abstract feature representations to achieve a higher level of speech understanding. This enables speech recognition systems to better recognize complex voice commands, voice emotions and voice context, providing a more intelligent driving experience.

At the same time, the development of combined artificial intelligence and machine learning will also enhance the adaptive and personalization capabilities of speech recognition systems. By analyzing the driver's voice characteristics, behavior patterns and preferences the system can be optimized according to individual needs and habits. For example, the voice recognition system can learn the driver's speech patterns, recognize specific voice commands and automatically adjust vehicle settings, such as adjusting the seat, temperature or audio settings, to provide a more personalized driving experience.

In addition, the development of artificial intelligence and machine learning will also drive the real-time and reliability improvement of speech recognition technology. During driving, drivers may need to issue voice commands or obtain information quickly, and fast response and accurate recognition are critical. By optimizing algorithms and increasing computing power, speech recognition systems can achieve lower latency and higher real-time performance to meet drivers' needs for immediate feedback.

Overall, the development of combined artificial intelligence and machine learning will further promote the application of speech recognition in driving. By continuously improving algorithms, enhancing system performance and strengthening personalization capabilities speech recognition will become a natural and intelligent way to interact between drivers and vehicles, providing a safer, more convenient and personalized driving experience.

4. Conclusion

The current status of speech recognition applications in driving human-computer interaction: Speech recognition technology has been widely used in driving, such as voice control of navigation, entertainment systems and telephone communication. These applications have greatly improved the convenience and safety of driving. Challenges and limitations in driving environments: The accuracy of speech recognition systems in noisy driving environments remains a challenge. Therefore, we need to seek improved methods, such as using deep learning techniques and incorporating other sensor data, to improve the accuracy and stability of voice commands. New perspectives on speech recognition in driving: ln addition to traditional command recognition and execution functions, speech recognition can be used in areas such as driving behavior analysis, intelligent traffic systems, and intelligent assistants.

DOI: 10.54254/2755-2721/31/20230125

These new perspectives offer broad prospects for the development of personalized driving experiences and intelligent driving assistance functions.

References

[1] Smith, J., & Johnson, A. The role of voice recognition in driving interactions. International Journal of Human-Computer Interaction, 2018, 34(6), 487-498.
[2] Lee, S., & Kim, M. Improving driving experience with voice recognition in intelligent vehicles. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(1), 267-277.
[3] Zhang, Y., & Wang, H. A review of speech recognition technologies for automotive applications. IEEE Transactions on Vehicular Technology, 2019, 68(4), 3344-3356.
[4] Xu, K. & Zhang, Y., & Wang, D. Real-time automatic speech recognition for in-vehicle applications: A survey. IEEE Transactions on Intelligent Vehicles, 2021, 6(4), 666-680.
[5] Wang, R., & Wang, J. Driver behavior analysis based on speech recognition: A review. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9), 3903-3916.
[6] Singh, A. & Kaur, N. & Kukreja, V. et al. Computational intelligence in processing of speech acoustics: a survey. Complex Intell. Syst. 2022, 8, 2623–2661.
[7] Huang, H., & Wang, R. Personalized driving with speech recognition: Challenges and opportunities. IEEE Intelligent Transportation Systems Magazine, 2020, 12(2), 52-61.
[8] Chen, Y. Design and Implementation of Intelligent Interactive In-Vehicle Speech System (Master's Thesis, Beijing Jiaotong University). 2020.
[9] Liu, Y. & Li, Y. Development Status and Trends of Intelligent Cockpit Multimodal Interaction Technology. Practical Automotive Technology, 2023, 01,182-187.
[10] Zhao, M. Research on Human-Machine Interaction Design of In-Vehicle Entertainment System. 2018.
[11] Yu, Y. Development of Human-Machine Interaction System for In-Vehicle Navigation Device (Master's Thesis, Jilin University). 2007.
[12] Zhou, R. Building a New World of Voice Technology Applications in Connected Vehicles. Securities Times, B04. 2012.
[13] Shengfan H. Design and Research of Voice-Based In-Vehicle Information System using Intelligent Mobile Terminals (Master's Thesis, South China University of Technology). 2014.
[14] Park, J., & Choi, S. Personalized driving experience using voice-based emotion recognition in intelligent vehicles. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(4), 2323-2334.
[15] Li, H., & Deng, L. Deep learning for speech recognition. Synthesis Lectures on Signal Processing, 2018, 11(3), 1-196.
[16] Bai, Y., & Li, J. Driving behavior analysis based on speech emotion recognition: A review. IEEE Transactions on Intelligent Vehicles, 2018, 3(3), 214-225.