SCIENCE CHINA
Information Sciences

• PERSPECTIVE •

May 2021, Vol. 64 156101:1–156101:3
https://doi.org/10.1007/s11432-020-3171-4

Future vehicles: interactive wheeled robots

Nan MA[1,2], Deyi LI[3*], Wen HE[3*], Yue DENG[4], Jiahong LI[2], Yue GAO[5], Hong BAO[1,2], Huan ZHANG[6], Xinkai XU[7], Yuansheng LIU[1,2], Zhixuan WU[1,2] & Li CHEN[1,2]

[1]*Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing* 100101*, China;*
[2]*College of Robotics, Beijing Union University, Beijing* 100101*, China;*
[3]*Institute of System Engineering, Academy of Military Science, Beijing* 100036*, China;*
[4]*School of Astronautics, Beihang University, Beijing* 100191*, China;*
[5]*School of Software, Tsinghua University, Beijing* 100084*, China;*
[6]*Smart City College, Beijing Union University, Beijing* 100101*, China;*
[7]*Demonstration Center of Experimental Teaching in Comprehensive Engineering, Beijing Union University, Beijing* 100101*, China*

Received 14 October 2020/Revised 24 November 2020/Accepted 8 December 2020/Published online 12 April 2021

Citation Ma N, Li D Y, He W, et al. Future vehicles: interactive wheeled robots. Sci China Inf Sci, 2021, 64(5): 156101, https://doi.org/10.1007/s11432-020-3171-4

Owing to the lack of variety of interactions in automatic driving, the interactive cognition and evolution growth of self-driving vehicles in uncertain and complex environments have been proposed, making future vehicles interactive wheeled robots.

Indeed, self-driving vehicles, as future vehicles, will change people's mode of traffic, transportation, and even production and life. Currently, studies on automatic vehicles mainly focus on perception, planning, decision, and control. However, self-driving vehicles need to communicate with their surroundings, including people, vehicles, and roads, to maintain a convenient and efficient operation. Consequently, interactive intelligence would be vital in future self-driving vehicles.

Lack of interactive cognition in automatic driving. Currently, when considering the implementation of self-driving technology, people are more concerned about how to display drivers' skills and techniques on the vehicle computer platform [1]. However, the rich interactive cognitive abilities that a driver must possess before being issued a license are often ignored. For example, if someone waves on the roadside, it may mean calling for a car; if someone waves while crossing a road, it may mean to go first; if a driver waves when two cars meet in opposite directions in a narrow road, it may mean to go first so as not to block the traffic. All these interactions are an important reflection of courteous driving and social civilization. Future vehicles are also expected to have a strong adaptive ability to immediately deal with an emergency at an intersection. These perceptual interactions are obtained anytime in mobile life, and future vehicles will shuttle between different nationalities, regions, even cultures, to meet the interactive cognition demands of humans, which may be developed over time.

Composition and infrastructure of self-driving interactive cognition. Considering physical space, the interaction of a self-driving vehicle can be divided into inner-vehicle, outer-vehicle, and remote cloud interactions. First, in **inner-vehicle interaction**, passengers, security staff, and other people can set different travel tasks, enjoy entertainment services, and share necessary travel information in the vehicle. Passengers can use voice, text, and other media to interact with the vehicle online using the natural-language-understanding technology. Second, **outer-vehicle interaction** refers to the interaction behavior of recognizing and predicting the posture of pedestrians and traffic policemen. Third, **remote cloud interaction** can be used for intelligent RoboTaxi services, remote command and dispatch, remote intervention requests, and execution of self-driving tasks. There are many uncertainties when dealing with emergencies in these perceptual interaction situations, such as sending real-time information to achieve diversified human intervention and making more reasonable decisions with a minimum loss [2]. Such desired ability can be used to explore the similarities among interaction uncertainties in various scenarios, such as autonomous parking, fixed-route logistics, and intercity driving.

According to different cognitive objects, interactive cognition in self-driving vehicles can be divided into vehicle-to-human, vehicle-to-vehicle, and vehicle-to-environment interactions. The **vehicle-to-human interaction** refers to the understanding of changes in the surrounding environment and human behaviors, and it is necessary to enhance the accurate and efficient recognition of multiple objects in a hybrid environment with occlusions [3] (top panel, Figure 1). For example, first, complementary multi-perspective and high-quality visual data can be obtained through vehicular sensors. Second, data features are extracted from multi-view pictures through a multi-stage, multi-branch Convolutional Neural Network (CNN) structure, and then focused on depth features encoding. Finally, employing a

* Corresponding author (email: lidy@cae.cn, colcloud@126.com)

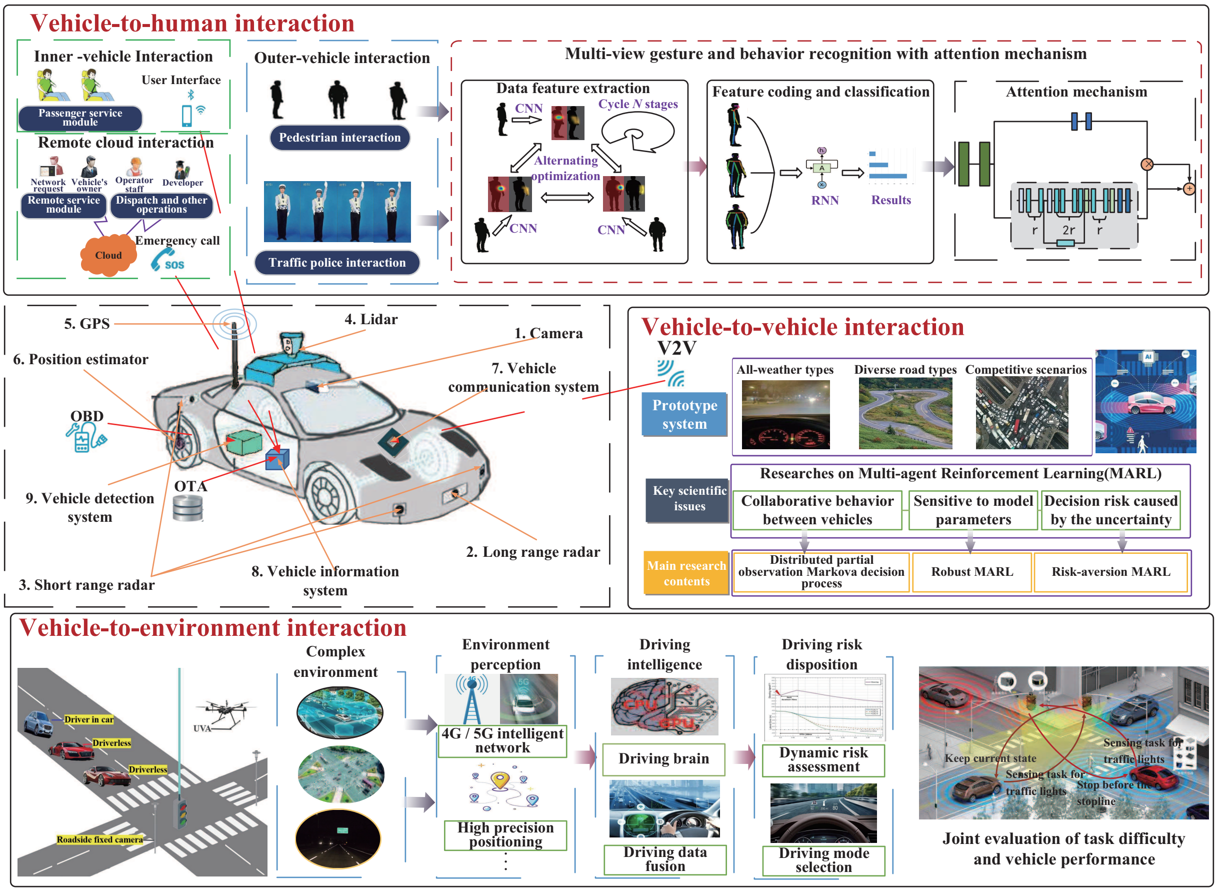

Figure 1 Interactive cognition and architectural design in automatic driving.

cyclic neural network to identify dynamic gestures in self-driving according to the timing relationship, can effectively improve the recognition efficiency by combining attention mechanisms [4]. The **vehicle-to-vehicle interaction** is employed while driving in shared public space with various road users, e.g., overtaking, narrow road meeting, and double merging (middle panel, Figure 1). A fundamental challenge is to achieve multi-agent coordination on perception, task, and motion to ensure safety, stability, comfort, and energy-saving [5]. The technical challenges and recent advances in future control methods and models were investigated in [6]. To adapt the model uncertainty and reduce potential risks in complex environments, such as all-weather types, diverse road types and competitive scenarios, Multi-agent reinforcement learning (MARL) algorithm is an effective tool, which can be combined with various methods [7], including risk-aversion [8] and variational Bayesian estimation [9] etc., to overcome the sensitivity of system parameters induced by outliers, and transfer learning to speed up the learning procedure in new tasks by reusing knowledge from past tasks [10]. The **vehicle-to-environment interaction** efficiently uses the sensing information from the 4G/5G intelligent network (bottom panel, Figure 1). To achieve fast localization and cooperation among future vehicles and guarantee intelligent driving in complex traffic environments and time-varying driving conditions, V2X sensor coverage would be improved in road networks to quickly collect more types of traffic information [11]. Self-driving vehicles should adopt the task-driven and data-centric AI testing approach [12] and have the cognition ability to drive intelligently [13], deal with as many kinds of accidental situations as possible, and become an interactive wheeled robot.

The evolution and growth of future automobiles in interaction. The basic features of human intelligence are to learn and evolve while interacting with the environment. Without intelligent interaction and evolution, future vehicles may not work well. With the advent of the intelligent era, the development of automobiles has made people realize that robot driving can coexist with human driving for a long time. Thus, it is important to discuss how wheeled robots can learn and evolve with and beyond humans. The process can be divided into three stages. In the first stage, wheeled robots learn from the model drivers to inherit their excellent driving skills by obtaining prior knowledge with supervised learning methods. The second stage is autonomous driving under human intervention. In this stage, robots attempt to finish work under human guidance through various learning methods, such as reinforcement, semi-supervised, and weekly supervised learning. The knowledge of robots gets further tested and strengthened. In the last stage, robots can drive the vehicles autonomously. Furthermore, robots can learn independently to accumulate experiences and strengthen or modify existing ones through unsupervised learning. After evolution, the evolved robots acquire driving skills, a typical situation-handling ability, and conventional accident prevention skill as prior knowledge, which enhances knowledge sharing and promotion. The evolutionary speed of interactive wheeled robots is much faster than the growth of a model driver in traditional training methods. These stages reflect the evolution and iterative development of interactive cognition in intelligent driving.

Application and demonstration of interactive cognitive technology. To discuss the necessity of interactive cognition in self-driving vehicles, we propose a new paradigm of interactive cognition, expound the interactive solution in the environment of "human–vehicle–road–cloud," and design a basic framework of interactive cognition. Since 2016, we have undertaken projects entrusted by Beijing Automotive Group Co., Ltd. (BAIC) and other enterprises. The aforementioned theories and methods of interactive cognition have been successfully employed in various autonomous vehicles, such as BAIC Foton TOANO electric vans, AUMARK light truck electric trucks, BAIC new energy LITE electric cars, and EU260 electric cars. The research findings have featured in the Tianjin World Intelligent Driving Challenge (WIDC) for three consecutive sessions (2018–2020) and won the leading prize (champion) of virtual scene competition for all the times. RoboTaxi intelligent online vehicle reservation system was operated in the 3rd World Intelligent Congress (WIC), which enriched experiences in shared transportation. The cloud intelligent interactive system for driverless buses developed by the team was also operated in the 4th WIC.

In conclusion, self-driving vehicles are challenged by realizing not only driving skills and techniques, but also intelligent interactions. Self-driving vehicles would be accepted by the public when incorporated with adequate interactive cognitive ability. The closed-loop experiment on vehicle dynamics with "the driving brain" provides a new method for vehicle testing. In the future, automobiles will be interactive, learnable, and evolvable wheeled robots.

Acknowledgements This work was partially founded by National Natural Science Foundation of China (Grant Nos. 61871038, 61931012, 61803034), Generic Preresearch Program of Equipment Development Department of Military Commission (Grant No. 41412040302), Consulting Project of Chinese Academy of Engineering (Grant No. 2019-XY-69), and Premium Funding Project for Academic Human Resources Development in Beijing Union University (Grant No. BPHR2020AZ02). We would like to thank Dianen ZHANG from Beijing Union University and Guilin PANG from Beijing Jiaotong University for the help on this work. We really thank anonymous reviewers' constructive suggestions.

Supporting information Appendixes A–C. Appendixes A is an introduction about the intelligent interaction team of Beijing Union University. Appendixes B and C cover the research achievements of the intelligent interaction team of Beijing Union University. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Ma N, Gao Y, Li J H, et al. Interactive cognition in self-driving (in Chinese). Sci Sin Inform, 2018, 48: 1083–1096
2 Li D Y, Ma N, Gao Y. Future vehicles: learnable wheeled robots. Sci China Inf Sci, 2020, 63: 193201
3 Ahmad B I, Murphy J K, Langdon P M, et al. Intent inference for hand pointing gesture-based interactions in vehicles. IEEE Trans Cybern, 2016, 46: 878–889
4 Yan C G, Tu Y B, Wang X Z, et al. STAT: spatial-temporal attention mechanism for video captioning. IEEE Trans Multimedia, 2020, 22: 229–241
5 Chen J, Zhang X, Xin B, et al. Coordination between unmanned aerial and ground vehicles: a taxonomy and optimization perspective. IEEE Trans Cybern, 2016, 46: 959–972
6 Ding Y L, Xin B, Chen J. A review of recent advances in coordination between unmanned aerial and ground vehicles. Unmann Syst, 2021, 2019: 1–21
7 Zhou M, Luo J, Villela J, et al. SMARTS: scalable multi-agent reinforcement learning training school for autonomous driving. 2020. ArXiv:2010.09776
8 García J, Fernández F. A comprehensive survey on safe reinforcement learning. J Mach Learn Res, 2015, 16: 1437–1480
9 Li J H, Deng F, Chen J. A fast distributed variational Bayesian filtering for multisensor LTV system with non-Gaussian noise. IEEE Trans Cybern, 2019, 49: 2431–2443
10 Silva F L D, Costa A H R. A survey on transfer learning for multiagent reinforcement learning systems. J Artif Intell Res, 2019, 64: 645–703
11 Molina-Masegosa R, Gozalvez J. LTE-V for sidelink 5G V2X vehicular communications: a new 5G technology for short-range vehicle-to-everything communications. IEEE Veh Technol Mag, 2017, 12: 30–39
12 Li L, Wang X, Wang K F, et al. Parallel testing of vehicle intelligence via virtual-real interaction. Sci Robot, 2019, 4: 4106
13 Chen S T, Jian Z Q, Huang Y H, et al. Autonomous driving: cognitive construction and situation understanding. Sci China Inf Sci, 2019, 62: 081101