

Optimizing Cabin Comfort Based on Artificial Intelligence

Minhang Wang

College of Computer and Artificial Intelligence, Beijing Wuzi University, Beijing, China
wmh20051231@outlook.com

Abstract. The comfort of the intelligent cockpit has surpassed the traditional standards and evolved into an active, personalized, and holistic intelligent service that is related to health and safety. The comfort of the intelligent cockpit is the core of contemporary automotive competition and has shifted from being an optional feature to a requirement. Under the trend of electrification and intelligence, the comfort experience directly influences users' choices and has become the key for automakers to create differentiation and enhance brand value. This article addresses the issue of insufficient regulation and control over temperature, sound, and lighting environments in the intelligentization of automobile cockpits. Analyze factors that enhance the personalized comfort experience for drivers. The adaptive linkage adjustment of systems such as seat heating, audio tuning, and road warning response significantly enhances the cockpit experience. This system can enhance the overall satisfaction with comfort. It can also reduce energy consumption. This article provides an effective solution for achieving active intelligent control of the cockpit environment. This article also holds significant importance in enhancing product competitiveness and driving the development of intelligent cockpit technology.

Keywords: artificial intelligence, intelligent cockpit, deep learning

1. Introduction

With the development of intelligent cockpit technology. The car cockpit is evolving from a traditional driving space to an intelligent scenario. Against this backdrop, user experience and comfort have become the key to differentiated competition among automobile brands. It surpasses traditional temperature and humidity control. It encompasses a personalized perception of environmental temperature, sound environment, visual transformation, air quality, and even olfactory atmosphere. Improving cabin comfort is of crucial significance for driving the value upgrade of the automotive industry.

Currently, the technology of intelligent cabins has made certain progress. Chen researched how large-scale AI models can enhance the multimodal interaction of the intelligent cabin of new energy vehicles. By analyzing the multi-modal interaction technology architecture, the multi-modal interaction hierarchical architecture, and other related concepts are introduced. Point out the challenges posed by AI large-scale models in enabling multimodal interaction in the intelligent cabin of new energy vehicles, and propose strategies for optimizing the interaction system and its application [1]. Liu, based on the theory of context awareness, restructured the driving context. The

application mechanism of context awareness in the interactive design of intelligent cabins was explored, and a user-centered design strategy was extracted, enhancing the human-machine compatibility and emotional warmth of the intelligent cockpit [2]. However, the existing technical system still has significant limitations. Firstly, most of the comfort subsystems operate independently and lack multimodal interaction integration, making it difficult to meet the comprehensive comfort demands of users in complex and dynamic environments.

Secondly, the perception dimension is single and passive. Existing solutions rely heavily on limited environmental physical parameters, and there is a lack of real-time and precise capture of key information, such as the physiology and subjective feelings of drivers and passengers. This results in lagging regulation and insufficient personalization.

In response to the aforementioned issues, this article proposes the design of a sensor network deployment scheme and constructs a prediction and decision-making model based on deep learning and reinforcement learning to support dynamic analysis of multi-dimensional comfort. In addition, the collaborative logic of integrated adaptive closed-loop regulation for the interior environment of vehicles has been defined. This article aims to establish a theoretical model for active collaboration in intelligent cockpit comfort.

2. Current status of the traditional intelligent cockpit

Currently, traditional intelligent cockpits have entered a stage of widespread adoption marked by large screens, voice recognition, and basic connectivity. However, the interactive voice recognition system is not accurate enough in understanding user intent, and the replacement of physical buttons with touchscreens sacrifices some flexibility. The ecosystem lacks in-depth applications tailored for driving scenarios. For example, Yin tested three traditional cockpits: pure hardware interaction, software-hardware combined interaction, and pure software interaction, at speeds of 30 km/h and 60 km/h. He concluded that the usability of pure hardware interaction is limited. Pure software interaction requires more visual attention from the driver, which may have a negative impact on safety, especially during high-speed driving. The integration of software and hardware interaction attempts to balance pure software interaction and pure hardware interaction and provides a compromise solution. It not only enhances usability but also reduces the driver's scanning time, making it safer than pure software [3].

3. Key technologies for optimizing cockpit comfort through artificial intelligence

3.1. Machine learning and adaptive learning system

3.1.1. User behavior habit construction and preference prediction

The core of intelligent control of cabin comfort is the adaptation of the system to user habits. The traditional solution is fixed, which makes it difficult to deal with the various behavioral habits of users. Kong proposed an intelligent service robot control system based on Reinforcement Learning. Based on Raspberry Pi 4b+ as the core controller, sensor cooperation, and voice interaction, this system realizes the interaction between the agent and environment through reinforcement learning. The system is also integrated with the autonomous decision-making module, which enables the robot to independently plan the walking path when guiding customers, and to independently decide whether to continue working or return to the energy storage area when completing the killing work, so as to further reduce the dependence on human intervention. This system can effectively reduce

the risk of infection among medical staff and further improve the efficiency of epidemic prevention in China [4]. Cabin comfort can also use reinforcement learning and sensor fusion technology to build a dynamic learning mechanism for user habits. Through a variety of sensor feedback from seats, steering wheels, air conditioners, and other equipment, the system collects the habits of drivers and passengers in real time, forming a multi-dimensional user feature vector. On this basis, the reinforcement learning model takes user feedback as the reward signal, and constantly iterates and optimizes the realization of functions such as temperature change, air conditioning adjustment, and seat heating. To achieve accurate control of user habits. In addition, an online learning method based on a dynamic neural field is introduced into travel scenario prediction. This method actively provides services at the corresponding time points by learning the periodic behaviors of users, such as commuting time, common routes, parking habits, etc., such as early seat heating, autonomous hot car, etc., so as to push the comfort experience from passive to active.

3.1.2. Adaptive adjustment algorithm for a dynamic environment

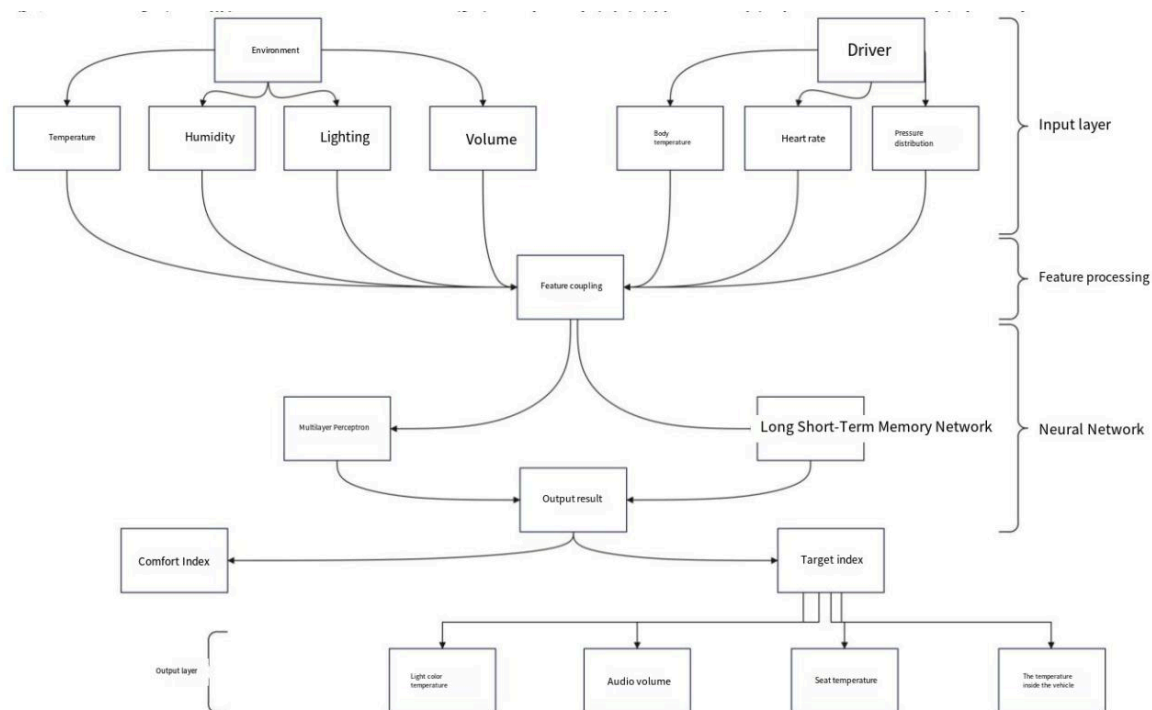


Figure 1. Cabin comprehensive comfort prediction model (picture credit: original)

Cabin comfort is a nonlinear result of the interaction between the environment and the pilot. The biphasic prediction model investigated by Zhou uses facial thermal imaging and environmental factors to assess the thermal comfort of passengers. He combined the advantages of the SAM-ResNet34 architecture with machine learning algorithms and conducted a comparative analysis of the standalone SAM-ResNet34 model and various hybrid deep learning-machine learning (DL-ML) models. The results show that the DL-KNN hybrid model has superior predictive performance compared to other models. It achieved a prediction accuracy of 89.45% and 94.44% on the summer and winter test sets, respectively, as well as an F1 score of 88.78% and 94.49% [5]. Intelligent driving can form a comprehensive cabin comfort prediction model based on deep neural networks. The input layer of the model integrates two types of features: environmental features (temperature, humidity, light, volume) and driver features (body temperature, heart rate, pressure distribution).

Through a hybrid architecture combining multi-layer perceptrons and long short-term memory networks, the model is capable of outputting the overall cabin comfort index in real time and decomposing it to generate the target values for each subsystem. As shown in Figure 1.

3.2. Multimodal fusion interaction technology

3.2.1. Speech recognition and natural language understanding

Voice interaction serves as the entry point for human-machine communication in current intelligent vehicle cabins. The level of its intelligence directly determines the satisfaction of users with the comfort control of the cabin.

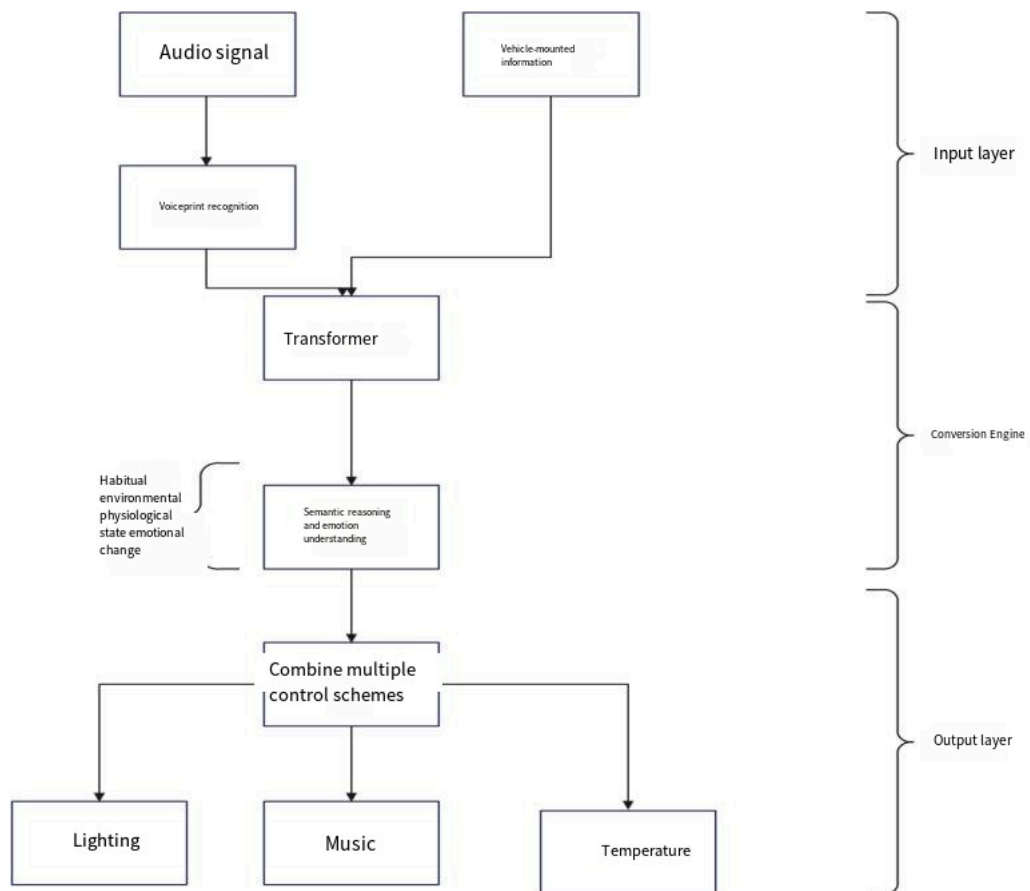


Figure 2. Speech and meaning understanding engine (picture credit: original)

Peter's voice-gesture fusion based on Transformer enables stable human-computer communication. By merging multimodal data into a unified sentence and then processing it by a language model, probability embedding is used to handle uncertainties, and context scene understanding is integrated to address ambiguous references. Transformer is evaluated in simulation and real-world experiments to prove its robustness to noise, dislocation, and missing information. The results show that the performance of the transformer is better than the deterministic baseline. Especially in scenes that need more context knowledge, so as to realize more powerful and flexible man-machine communication [6]. The intelligent cockpit can build an end-to-end speech semantic understanding engine based on the transformer architecture, break through the traditional delay bottleneck, and use the joint encoder to realize the parallel processing of speech features and

vehicle information. Introduction of intention reasoning mechanism enhanced by knowledge map—When the user gives a fuzzy command such as a little cold, the system will no longer mechanically perform the temperature rise operation, but will comprehensively conduct joint inference and decide to output a composite control scheme of multi-system regulation. In addition, the introduction of emotional computing power has become another focus -- through voiceprint feature extraction and intonation analysis. The system can identify users' emotional states such as fatigue, anxiety, and pleasure. When it detects that the driver is in a high-pressure working state for a long time, it can actively switch to the soothing mode, jointly dim the lights, reduce the volume, turn on the fragrance, and suggest stopping in the rest area to change from a tool to a partner. As shown in Figure 2.

3.2.2. Visual inspection and gesture recognition

Wang conducted in-depth research on gesture recognition. The intelligent cockpit can also utilize the visual detection subsystem of the on-board infrared camera to achieve the recognition of drivers' and passengers' gestures (such as sliding, tapping, clenching fists, palm orientation, etc.) and the positioning of the gaze area [7]. Introducing the fusion of speech and gestures using a Transformer. For example, when both the temperature increase command and the pointing gesture are present, the system will, based on the historical startup records of each subsystem, preferentially adopt the modal output that is more clearly pointed and has a lower probability of accidental touch. An important application of visual perception is driver fatigue monitoring. By using features such as eye-opening/closing degree and blinking frequency, a fatigue grading mechanism is established. When fatigue accumulation exceeds the threshold, the system automatically activates intervention measures such as light adjustment and music rhythm, achieving a dynamic balance between comfort and safety.

3.2.3. Touch sensation and force feedback enhance the realism of interaction

Tactile interaction, as the underlying channel of multimodal fusion, holds irreplaceable value for aspects such as the safety of blind operation and the sense of operation confirmation. Kern achieves this by embedding vibration tactile actuators on both sides of the steering wheel, thereby transmitting navigation steering instructions to the driver. The experimental results show that drivers can accurately identify directional vibrations, and tactile navigation requires less distraction of attention compared to visual navigation. The intelligent cockpit can embed linear resonant actuators and piezoelectric tactile feedback units in the steering wheel, center console, and seat side panels. When users trigger control instructions through gestures or voice, the system generates directional vibration feedback at the corresponding positions [8]. Traditional tactile feedback is mostly based on single-frequency vibrations and lacks spatial directionality. The steering wheel, as the most continuous and direct contact interface between the driver and the vehicle, is equipped with an interactive sensing function - when the lane departure warning is triggered, the left or right steering wheel spokes will produce directional vibrations, conveying the direction of deviation to the driver; when the seat massage mode is switched, the progressive inflation and deflation of the pressure distribution array creates a wave-like push-back sensation, significantly enhancing the authenticity and pleasure of the operation feedback.

4. Discussion and analysis

Currently, intelligent cabins have problems such as insufficient accuracy in multimodal interaction integration, a single perception dimension, the absence of a user habit prediction mechanism, and insufficient realism in interaction feedback. For instance, the system has difficulty in quickly conducting joint judgment and intent reasoning for multimodal data such as voice, gestures, and physiological characteristics, resulting in a low accuracy rate for ambiguous instruction interpretation. The perception dimension overly relies on environmental physical parameters and lacks real-time monitoring of the crew's physiological state and emotional characteristics. It is unable to form an "environment - physiology - behavior" adaptive regulation/user habit learning that is dependent on solidification, rigidity, and lacks dynamic preference prediction driven by reinforcement learning. It is difficult to actively trigger comfortable personalized services in some changeable and complex scenes. Tactile feedback is still dominated by single-frequency vibration, lacking real interactive experience, such as steering wheel directional vibration and pressure array wave pushing back feeling. In the future, it is necessary to introduce the joint coding architecture based on a transformer, deploy multimodal sensor networks to build a deep learning comfort prediction model, and apply reinforcement learning to realize the dynamic prediction of user habits. Spatial directional tactile feedback is embedded in the steering wheel and other interfaces to promote the evolution of cockpit interaction from passive response to active service, and comprehensively improve comfort and personalized experience [9, 10].

5. Conclusion

This paper investigates the research progress and technical path in the field of comfort optimization of the intelligent cockpit. Although the intelligent cockpit has currently entered the stage of widespread adoption, it still encounters several bottlenecks, including inadequate precision in multimodal fusion, a singular and passive perception dimension, a lack of user habit prediction capabilities, and insufficient realism in interactive feedback. In view of the above problems, the existing research shows a trend of evolution from independent subsystems to integrated collaborative regulation. By deploying multimodal sensor networks to collect real-time environmental physical parameters and crew physiological behavior data, a comprehensive comfort prediction model based on in-depth learning is constructed to realize the dynamic evaluation of temperature, sound, light, and other dimensions. Reinforcement learning and IOT sensor fusion technology are introduced to continuously optimize the strategy through user feedback. In the interaction layer, a multimodal fusion engine is built based on the transformer architecture, combined with the knowledge map to enhance intention reasoning, and the tactile feedback with spatial directivity is embedded. Research shows that the above technology path can effectively drive the seat, air conditioning, audio, and other systems for adaptive linkage and optimize energy consumption while improving comprehensive comfort satisfaction. Future research can further explore the application of generative artificial intelligence in emotional interaction and scene generation, and provide key technical support for the transformation of automotive intelligence.

References

- [1] Chen, Q. (2026). Research on multi-modal interaction of intelligent cockpit in new energy vehicles empowered by AI large models. *Automotive Knowledge*, 26(2), 8–10.
- [2] Liu, W., & Li, J. (2026). Application of screen printing in user experience design of context-aware intelligent cockpits. *Screen Printing*, (2), 42–44.

- [3] Yin, H., Li, R., & Chen, Y. V. (2024). From hardware to software integration: A comparative study of usability and safety in vehicle interaction modes. *Displays*, 85, 102869.
- [4] Kong, F. (2023). Intelligent service robot control system based on reinforcement learning. *Internet of Things Technology*, 13(5), 77–79.
- [5] Zhou, X., Miao, Z., Yuan, M., Li, K., Guo, H., Lin, X., Zeng, Y., Tu, R., & Zhong, J. (2024). Dual-phase prediction model of passenger thermal sensation using facial thermal imaging and environmental factors. *Case Studies in Thermal Engineering*, 58, 104439.
- [6] Vanc, P., & Stepanova, K. (2025). Transformerger: Transformer-based voice-gesture fusion for robust human-robot communication. *arXiv preprint arXiv:2504.01708*.
- [7] Wang, B., Yu, L., & Zhang, B. (2024). AI-mobilenet: A novel model for 2D gesture recognition in an intelligent cockpit based on multi-modal data. *Artificial Intelligence Review*, 57, 282.
- [8] Kern, D., Marshall, P., Hornecker, E., Rogers, Y., & Schmidt, A. (2009). Enhancing navigation information with tactile output embedded into the steering wheel. In H. Tokuda, M. Beigl, A. Friday, A. J. B. Brush, & Y. Tobe (Eds.), *Pervasive Computing. Pervasive 2009. Lecture Notes in Computer Science (Vol. 5538)* (pp. 302–318). Springer.
- [9] Liu, B. (2025). Analysis of the current development status and trends of multi-modal interaction technology in intelligent cockpit [Paper presentation]. Academic Symposium on the Development of Artificial Intelligence and Economic Engineering, Jiangling Automobile Co., Ltd., Chongqing, China (pp. 421–424).
- [10] Li, L., Zhang, B., & Duan, T. (2025). Design strategies for intelligent automotive cockpit based on the integration of artificial intelligence technology and user experience. *Popular Automobile*, (03), 4–6.