Design and Implementation of Service Robots for Special Groups Based on a Multimodal Interaction System

Lan Li, Fabin Yang, Zhuanhong Qi, Jiaxue Kang, Yingying Ma, Jundi Wang*
School of Computer and Artificial Intelligence, Lanzhou Institute of Technology, Lanzhou, Gansu, China
**Corresponding Author*

Abstract: With the high-density coverage of China's railway network and improved operational efficiency, high-speed rail (HSR) has become a core travel choice for citizens, and the elderly's demand for this convenient transport is rising markedly. However, traditional HSR services, relying on manual support, have not able shortcomings in coverage, response efficiency and adaptability to special groups—failing to meet their needs for information access, easy operation and fast travel, while also having unreasonable traditional meal distribution. Against this backdrop, the "Zhixiang Lvtu Jingling" (Intelligent Travel Elf) HSR service robot, a fully automatic smart device for HSR scenarios, is key. Centered on a "multimodal interaction system," it integrates voice, vision, touch and payment interactions to build a full-process service loop of perception, integration, decision-making and execution. It addresses travel pain points of special groups (the elderly, visually impaired, etc.) and optimizes efficiency in meal distribution and pet supervision. Its design, focused on elderly-friendly principles rather than simply replacing labor, unifies service universality and precision. In short, the system, linked by elderly-friendly design, boosts HSR service efficiency and breaks demographic/scenario barriers of traditional services. This concept provides a practical solution for smart HSR services and a reference for elderly-friendly upgrades in airports, subways and nursing homes, aiding an inclusive elderly social service system.

Keywords: Multimodal Interaction; Artificial Intelligence Robot; Speech Recognition; Diverse Interaction

1. Introduction

With the continuous development and advancement of science and technology in China, the travel needs the elderly and people with mobility impairments have increasingly become a focus of attention. According to data from the National Bureau of Statistics, by the end of 2024, the population aged 60 and above in China had reached 280 million, accounting for 19.8% of the total population; the total number of people with disabilities was approximately 85 million, among whom about 38 million held disability certificates. Additionally, the number of elderly passengers traveling by high-speed rail (HSR) has increased by more than 15% annually. As an efficient, safe, and convenient public transportation method, HSR has become an important choice for the elderly to travel across regions. However, the current HSR service system is still dominated by the traditional manual service model, which has obvious shortcomings in service coverage, response efficiency, and adaptability to special groups. It is difficult to meet the personalized needs of the elderly in terms of information access and operational convenience, and also cannot effectively solve other service pain points in HSR scenarios—this aligns with the general demand for bridging the digital divide in the public service sector [1]. The level of service intelligence affects the convenient travel experience of elderly passengers and people with disabilities. Therefore, in response to the service needs of special groups during HSR travel, we have designed an artificial intelligence robot system based on a multimodal interaction system. By integrating multimodal interaction technologies such as voice, vision, and touch, and adhering to the principles of convenient travel and care for vulnerable groups, the system not only improves service efficiency but also significantly enhances the travel experience of elderly passengers and people with disabilities.

This multimodal interaction system takes multi-dimensional perception and scenario-based response as its core design concepts, integrating multimodal interaction methods including voice, vision, touch, and payment. The voice interaction integrates an offline dialect voice database, combined with speech speed adjustment and semantic fault tolerance functions to adapt to the language habits of elderly users. The visual interaction adopts an adjustable large-font touch screen and 3D real-scene navigation rendered by CesiumJS [2] to meet the needs of users with visual impairment. The tactile interaction ensures the operation of visually impaired users through vibration feedback from guide canes and raised Braille buttons. The payment interaction integrates cash recognition, QR code payment, and payment on behalf of relatives to make up for the lack of smartphone operation capabilities.

The system adopts a hierarchical progressive and collaborative interaction architecture. The perception layer integrates devices such as microphone arrays, depth cameras, and LiDAR to collect multi-source data in real time—among which the application of LiDAR and positioning devices can draw on the map construction and positioning technology of outdoor inspection robots [3], and the research on mobile robot navigation systems also provides a reference for the overall positioning solution [4]. The fusion layer realizes data feature extraction and accurate fusion through improved models, visual detection, and Kalman filtering algorithms. The decision-making layer generates service strategies based on structured information. The execution layer outputs service results in multiple dimensions through iFlytek UNI-TTS Pro speech synthesis and robotic arm movements, with iFlytek's intelligent speech basic resource database providing support for the speech synthesis function. The hardware design follows the elderly-friendly principle: the robot height is adjustable (adjustable appropriately according to actual scenarios), the body is made of lightweight ABS and PC materials with rounded corner treatment, the touch screen brightness is automatically adjusted, and the triggering force of the emergency call button is ≤5N. The software has a built-in elderly mode, which shields complex functions and prioritizes the display of high-frequency needs; at the same time, it realizes low-latency communication through the MQTT protocol, ensures security through AES-256 encryption and Hyperledger Fabric blockchain evidence storage, and guarantees service continuity when the network is interrupted with the offline cache function. This system has the advantages of high reliability, low power consumption, strong anti-interference ability, and high flexibility. It not only solves problems such as inconvenient HSR travel for special groups and low efficiency of traditional services but also significantly improves the universality of human-computer interaction, providing a full-scenario solution for smart HSR services. Application cases such as the "All-Round Patrol Robot" at Wuxi Railway Station also verify the feasibility of such robots in transportation hub scenarios.

2. Interaction Design of the Multimodal Interaction System

In the R&D system of the "Zhixiang Lvtu Jingling (Intelligent Travel Elf)" high-speed rail (HSR) service robot, multimodal interaction design is the core technical module. With the core goals of adapting to HSR scenarios, focusing on elderly-friendly services, and serving all groups, it integrates four interaction dimensions voice, vision, touch, and payment behavior to build an interaction framework featuring multi-dimensional perception and scenario-based response. This framework aims to break down usage barriers for special users such as the elderly and visually impaired, realize natural, convenient, and inclusive human-computer interaction, while also adapting to the service experience of all groups. The design is not a mere stack of single technologies, but a in-depth customization based on the characteristics of HSR scenarios and the physical and psychological needs of special groups; the implementation of each interaction function centers on lowering operation thresholds and improving the accuracy of information transmission [5].

2.1 Voice Interaction Design

As one of the most acceptable interaction methods for the elderly, voice interaction design focuses on covering dialect differences, adapting to hearing impairment, and optimizing environmental anti-interference capabilities. To address the issue that some elderly users have strong accents and are accustomed to communicating in dialects, the system integrates an offline dialect voice database, which not only covers mainstream southern dialects such as

Cantonese and Sichuan-Chongqing dialects, but also specifically includes 8 northwestern dialects like Shaanxi dialect and Gansu dialect. This ensures that elderly users traveling in these regions can access services in familiar languages, avoiding interaction barriers caused by inability to understand Mandarin. The optimization of this dialect recognition technology can refer to Haier TV's elderly-friendly AI large model, and the elderly-friendly design ideas for multimodal information interaction of smart TVs also have reference value [6]. Meanwhile, the system supports speech speed adjustment. When the microphone array detects the user's unclear accents, hesitant responses, or repeated inquiries, it will automatically switch to slow speech speed and keyword repetition mode. For example, if a user asks "How to get to Carriage 3?", the system will first identify the demand through the semantic fault tolerance function, then broadcast repeatedly in the form of "To reach Carriage 3, you need to go straight 50 meters and then turn left. Please go straight 50 meters, then turn left, and you will arrive at Carriage 3." This not only adapts to the characteristic of the elderly's slower auditory response speed, but also ensures accurate information transmission through keyword emphasis. Vosk voice-driven model technology can provide support for the underlying implementation of voice interaction. In terms of function coverage, voice interaction runs through the entire service process and provides dialect broadcasts during navigation, allowing elderly users to complete core operations without visual participation.

2.2 Visual Interaction Design

Visual interaction design centers on large fonts, high contrast, and realism, specifically addressing the issues of visual degradation (e.g., blurred vision caused by presbyopia and cataracts) and low acceptance of complex interfaces among the elderly. For the hardware carrier, the system adopts a touch screen, whose size is not only convenient for the elderly to operate with one hand but also avoids information congestion caused by an overly small screen. The screen font size is adjustable manually; combined with the high-contrast color scheme of black background and white text, it effectively reduces the elderly's visual fatigue. Even in brightly lit platforms or dimly lit carriages, users can clearly see the interface content (such as train schedules, payment amounts, and navigation routes). Meanwhile, to solve the problem of the elderly getting lost due to factors like the large space and complex routes in HSR stations, visual interaction introduces a 3D real-scene navigation function. It renders the 3D model of the HSR station's building information model through the Cesium JS engine, converting abstract transfer routes into intuitive real-scene images. On this basis, the system is equipped with the Dijkstra algorithm, which automatically plans the shortest transfer path based on the user's destination, marks directions with dynamic arrows on the 3D model, and forms dual visual and auditory guidance together with voice prompts—preventing the elderly from getting lost due to their inability to understand 2D maps. In addition, visual interaction also includes a dynamic visual feedback design: it conveys service status via LED indicators and enlarges key information through on-screen pop-ups, further reducing the elderly's information recognition cost.

2.3 Touch Interaction Design

Touch interaction design focuses on visually impaired users and elderly users with limited mobility. Through the combination of vibration feedback and physical buttons, it builds an operation path that requires no visual participation, filling the service gap of traditional visual interaction. For the guide function, the system connects to guide canes via Bluetooth and outputs standardized vibration codes based on path planning: a single vibration indicates a left turn, double vibrations indicate a right turn, and a long vibration indicates obstacles ahead (such as steps and luggage). This simple and clear directional design allows visually impaired users to independently follow the robot to complete operations like platform transfers and carriage positioning without relying on others' guidance. For elderly users with limited mobility and reduced finger dexterity, the robot body is designed with a physical button area. In addition to the core emergency call button, two shortcut buttons (for repeating commands and increasing volume) are added. The emergency call button adopts a raised design with Braille marks on its surface, and its triggering force is $\leq$5N (equivalent to the weight of a 500g object), allowing elderly users to trigger it without pressing hard. After triggering, the button immediately links to the train staff's terminal and synchronously pushes the user's location and demand type, realizing one-click help and real-time response, and solving the pain point that users with limited mobility find it hard to seek help in emergencies.

2.4 Payment Interaction Design

The core goal of payment interaction design is to cover different payment habits and solve the digital divide problem for those who cannot use smartphone payments—which is consistent with the goal of eliminating the digital divide in the information accessibility optimization of airline retail websites. Therefore, a three-tier payment system (cash payment, QR code payment, and payment on behalf of relatives) is established to ensure every elderly user can find a suitable payment method. In the cash payment module, the system is equipped with a GRGBanking NV200SA banknote recognizer and a GLORY CR coin module, supporting the recognition of banknotes (denominations of 1-100 RMB) and common coins. The payment process can be simplified as follows: the user inserts banknotes or coins → the system confirms the denomination via voice (e.g., "50 RMB banknote recognized") the touch screen displays the consumption amount and change amount in large fonts → voice prompt "Please take your change". The entire process requires no complex operations from the user, fully catering to the elderly’s fear of operating or making mistakes. For elderly users who can use smartphones, the QR code payment module connects to three mainstream platforms: WeChat Pay, Alipay, and UnionPay Cloud Quick Pass. The generated QR code has a validity period of 120 seconds and refreshes dynamically, which not only ensures payment security but also avoids QR code expiration caused by the elderly’s slow operation. For elderly users who cannot use smartphones or operate QR code payment, the "payment on behalf of relatives" function provides a convenient solution. Based on FaceNet liveness detection technology, users only need to complete a combination of blinking and nodding actions; the system then matches pre-set relative information via the cloud, and sends a 5-digit SMS verification code (valid for 60 seconds) to the relative’s mobile phone. After the relative passes verification, the payment can be completed remotely. The entire process requires no mobile phone contact from the user, completely breaking the restriction that consumption is impossible without a mobile phone. In addition, the payment interaction is designed with a comprehensive exception handling mechanism: situations such as 3 consecutive payment failures and network interruptions will automatically trigger manual intervention, avoiding the elderly’s worry about financial losses due to payment failures or network issues.

2.5 Elderly-Friendly Overall Design and Technical Support of the Multimodal Interaction System

The design of the multimodal interaction system always takes the Guidelines for Elderly-Friendly Design of Household Products as its core principle, strictly following the requirements of simplifying operation processes, enhancing comfort, and optimizing intelligent functions. At the same time, it draws on industry practical experience: it refers to the dialect recognition technology of Haier TV’s elderly-friendly AI large model to improve the adaptability of voice interaction; and learns from the interesting interaction module of the "Xiaotie" robot at Xi’an Railway Station, adding emotional design during idle periods to reduce the elderly’s sense of unfamiliarity with intelligent devices and improve their acceptance. The application of AIGC technology in elderly-friendly design in the digital era also provides a new direction for emotional design [7]. From a technical perspective, the stable operation of the multimodal interaction relies on a threefold guarantee: low-latency communication, data security, and offline adaptation. Low latency is achieved via the MQTT protocol to realize low-latency communication between the robot, the train system, and the cloud platform, ensuring the real-time performance of operations such as voice command recognition and navigation path update. For data security, the AES-256 encryption algorithm and Hyperledger Fabric blockchain evidence storage technology are adopted to encrypt and store payment information and user identity data; transaction hashes and timestamps cannot be tampered with after being uploaded to the chain. Meanwhile, WORM (Write Once Read Many) storage is used to retain 5-year audit logs, ensuring the information security of the elderly. Offline adaptation supports offline modes such as Bluetooth navigation and local transaction caching; even in areas with weak HSR signals, core interaction functions (such as voice broadcasting and cash payment) can still operate normally, avoiding service interruptions caused by no signal or network.

3. Architectural Design of the Multimodal Interaction System

The multimodal interaction system takes hierarchical progression and collaborative interaction as its core architectural design concept. It simulates the collaboration logic of an intelligent service team, and constructs a full-process service covering data collection, processing and fusion, strategy generation, and instruction implementation through the close linkage of the perception layer, fusion layer, decision-making layer, and execution layer. This architecture not only adapts to the characteristics of HSR scenarios—such as dense crowds, strong noise interference, and complex spatial movement lines—but also highly aligns with the physical and operational needs of special users like the elderly and visually impaired. Meanwhile, it relies on technologies such as low-latency communication and data encryption to ensure service stability and security, ultimately realizing high-precision, highly inclusive, and reliable smart HSR services, just like the application case of the greeting robots in HSR halls.

3.1 Perception Layer

As the neural endings of the system, it integrates diverse devices to collect data such as voice commands, visual images, spatial positions, and touch signals in real time, providing complete, efficient, and high-fidelity raw data support for subsequent layers. Core devices and functions focus on anti-interference and adaptation to the habits of special groups, as shown in Table 1.

Table 1. Core Components and Functions

Device Type	Core Configuration	Core Aging-friendly Functions
Voice Perception Device	Knowles microphone array	Far-field wake-up and noise-reducing directional sound pickup
Visual Perception Device	4K camera, depth camera	Pet monitoring, assisting in obstacle avoidance and collision prevention
Spatial Positioning Device	Laser radar, UWB beacon	Accurate 3D navigation, resistance to metal interference
Tactile and Payment Perception Device	Pressure sensor, banknote recognizer, QR code module	Capturing touch, adapting to cash and QR code payment habits

3.2 Fusion Layer

The Fusion Layer receives raw data from the Perception Layer. Through deep learning algorithms and signal processing technologies, it completes the transformation of data cleaning, feature extraction, and multi-source fusion—converting fragmented, multi-format raw data into structured, high-reliability information, and providing usable, accurate input for the Decision-Making Layer. Its core processing logics include voice data processing, visual data processing, and spatial position processing.

First, for voice data processing: ambient noise in voice signals is filtered out via noise suppression algorithms, and the signals are then converted into text commands through an improved ASR-Lite end-to-end model. Meanwhile, the system automatically matches offline dialect voice databases to convert dialects into standard text, avoiding recognition deviations caused by differences in language habits.

Second, in terms of visual data processing: images captured by 4K cameras are processed using the YOLOv5 object detection model to achieve rapid recognition of human faces, gestures, abnormal pet behaviors, and obstacles. Depth images generated by the Intel RealSense D435i are fused with the point cloud data from the Velodyne VLP-16 LiDAR [8]. The accuracy of the 3D coordinates of obstacles is optimized through point cloud registration algorithms, which avoids visual misjudgments caused by light changes and ensures the obstacle avoidance safety of navigation for elderly users. Finally, spatial position data processing involves high-precision positioning correction [9]. Spatial position data collected by UWB Beacons and LiDAR is used to correct errors in real time via the Kalman filter algorithm, controlling the robot's positioning accuracy within a certain range. At the same time, combined with the BIM (Building Information Modeling) models of high-speed railway (HSR) stations, absolute coordinates are converted into relative scene coordinates, providing intuitive spatial references for path planning in the Decision-Making Layer. The core value of the Fusion Layer lies in eliminating data ambiguity, enhancing information credibility, and addressing the limitations of single devices in the complex HSR environment. It ensures the accuracy of information transmitted to the Decision-Making Layer, laying a foundation for the accurate generation of subsequent service strategies.

3.3 Decision-Making Layer

As the brain center of the system, it generates service strategies adapted to user needs and scenario characteristics based on the structured information output by the Fusion Layer, combined with the preset rule base and dynamic scenario analysis. It serves as the core for realizing personalized and precise services, with its decision-making logic centered on user identification and scenario demand matching. In the dynamic identification of user identity and needs, the system automatically determines the user type based on facial data, voice features, and operational behaviors from the Fusion Layer. For instance, users who use dialects, speak at a slow pace, and prefer touching physical buttons are identified as elderly users; while users detected carrying a white cane or triggering tactile interaction signals are identified as visually impaired users. In the intelligent generation of scenario-specific strategies, the system invokes corresponding algorithms and rules for different users and scenarios. When a visually impaired user initiates a navigation request, the system automatically activates the guide mode, invokes the TEB path planning algorithm and Social LSTM pedestrian prediction model, plans the optimal route with a reserved 1.2-meter safety distance, and triggers the vibration feedback command for the white cane. When an elderly user selects cash payment, the system prioritizes activating the GRGBanking NV200SA Banknote Recognizer Module, simultaneously triggers voice guidance processes such as “Please insert banknotes; the system will automatically recognize the denomination”, and disables the complex QR code payment interface. In the face of abnormal scenarios (e.g., three consecutive payment failures caused by banknote recognition errors or expiration of the verification code for payment on behalf of relatives), the system automatically triggers manual intervention: it provides a voice prompt stating “Staff have been notified to assist; please wait a moment” and pushes the user’s location to the staff terminal. In case of network interruption, it activates offline service rules to prioritize ensuring core functions such as cash payment and local navigation. The core advantage of the Decision-Making Layer lies in balancing flexibility and security. It not only ensures the standardization of services (e.g., elderly-friendly operation processes) through preset rules but also realizes on-demand adjustments via dynamic scenario analysis, avoiding cumbersome service modes.

3.4 Execution Layer

As the final implementation link of elderly-friendly design, the Execution Layer undertakes the service strategies from the Decision-Making Layer. It converts abstract strategies into user-perceivable service outcomes through multi-dimensional output methods (including voice, vision, touch, and mechanical actions), while ensuring the output forms are adapted to the perception capabilities of special groups. For voice output, the FLYTEK UNI-TTS Pro speech synthesis engine is adopted, supporting broadcast in dialects such as Cantonese and Sichuan-Chongqing dialects as well as speech rate adjustment [10]. For elderly users, it uses a loud volume and warm titles (e.g., “Uncle Zhang, your change of 70 yuan is ready; please collect it”) and avoids professional terms like “offline voucher”, replacing them with easy-to-understand expressions such as “paper receipt” to ensure information comprehensibility. Visual output is realized via a 10.1-inch IP5-rated touchscreen, with a font size ≥ 24pt and a high-contrast black-background white-text color scheme. Core information (e.g., payment amount, navigation distance) is displayed via pop-up magnification for 5 seconds, preventing elderly users from missing key content due to impaired vision. In 3D navigation scenarios, dynamic arrows are used to intuitively present routes (e.g., green for smooth traffic, red for congestion), reducing reliance on text. For touch output, the system connects to white canes via Bluetooth and uses standardized vibration codes: a single left vibration indicates a left turn, double right vibrations indicate a right turn, and a long vibration indicates an obstacle ahead, providing direction guidance for visually impaired users.

The emergency call and repeat command buttons in the physical key area adopt a raised design with an actuation force $\leq$ 5N, facilitating pressing by elderly users with limited mobility. After a button is triggered, a red LED flashes simultaneously to indicate that the command has been received. In terms of mechanical and scenario-based output, the robotic arm grabs and delivers items in a low-speed stable mode to avoid colliding with elderly users. In pet supervision scenarios, the on-board spray module reduces odors, and the LED screen displays the prompt “Pet under supervision, please do not approach” to balance the needs of pet owners and other passengers. The core design principle of the Execution Layer is multi-dimensional redundant output [11]. Considering the dual degradation of vision and hearing in elderly users, dual voice-visual feedback is adopted (e.g., payment results are both broadcast and displayed). For visually impaired users, dual voice-touch guidance is used to ensure users with different perception capabilities can accurately receive service information.

In conclusion, the architecture of the multi-modal interaction system breaks the limitations of traditional high-speed railway (HSR) services, such as heavy reliance on manual work and poor adaptability. Through the Perception Layer’s full-dimensional data collection to cover user needs, the Fusion Layer’s high-precision processing to ensure information credibility, the Decision-Making Layer’s scenario-specific strategies to deliver precise services, and the Execution Layer’s multi-dimensional output to adapt to special groups, it eventually forms a seamless closed loop from user needs to service implementation. Particularly for the elderly group, the architecture thoroughly addresses pain points like the digital divide and operational barriers through designs including anti-interference perception, simplified decision-making, and redundant output. This provides an implementable technical framework for enhancing the inclusiveness and efficiency of HSR smart services.

4. Conclusion

The design and implementation of the "Intelligent Journey Elf" HSR Service Robot and its Multi-Modal Interaction System take elderly-friendliness as the core orientation, with multi-modal interaction technology as the carrier. It integrates four interaction dimensions—voice, visual, tactile, and payment—and relies on a hierarchical, progressive, and collaborative architecture consisting of the Perception Layer, Fusion Layer, Decision-Making Layer, and Execution Layer, thus constructing a full-process closed loop from demand perception to service implementation. Its design is not a simple replacement for manual work; instead, through anti-interference data collection, accurate information fusion, personalized decision-making output, and redundant service feedback, it has not only effectively addressed the pain points of special groups (such as the elderly and visually impaired individuals) in HSR travel but also significantly improved service efficiency in scenarios like meal delivery and pet supervision, realizing the unification of service universality and precision. Meanwhile, the elderly-friendly design principles, multi-modal interaction technology integration plan, and hierarchical architecture logic adopted by the system not only provide an implementable solution for HSR smart services but also offer technical references and practical paradigms for the elderly-friendly transformation of other public service scenarios. This is of great significance for promoting the construction of a more inclusive service system for an aging society.

Acknowledgments

This paper is supported by Science and Technology Planning Project of Lanzhou Science and Technology Bureau (No. 2025-2-49).

References

[1] Zhonghua Wang, Liwei Cai, Xiongqing Li, et al. Bridging the Digital Divide: Current Situation Assessment and Optimization Strategies for Information Accessibility of Airlines' Retail Websites. Journal of Civil Aviation, 2024(5): 20-23.

[2] Xiaoyan Wang, Xiandong Chen. 3D Model Visualization Based on Cesium. Geospatial Information, 2022, 20(1): 4-12.

[3] Hongmeng Zhang. Design of an Intelligent Inspection Robot. China Machinery, 2024, (8): 40-43.

[4] Yuanshao Hou. Research on SLAM Technology of Inspection Robots with Multi-Source Information Fusion Based on Deep Learning. Information Technology and Informatization, 2023(8): 217-220.

[5] Li Wei. Intelligent Elderly Care Leads the

Future. China Venture Capital, 2015, 000(027): 34-39.
[6] Lei Chen. Design of Elderly-Friendly Apps in the Context of Digital Media. Wireless Internet Technology, 2023, 20(16): 57-61.
[7] Yang Xi. Elderly Care in the Digital Age: Application of AIGC Technology in Elderly-Friendly Design. Shanghai Packaging, 2025(04): 56-58.
[8] Xueqin Xiang, Zhigeng Pan, Jing Tong. Application Research of Depth Cameras in Computer Vision and Computer Graphics. Journal of Frontiers of Computer Science and Technology, 2011, 5(6): 481-492.
[9] Wenling Xia, Zhaopeng Gu, Tangsheng Yang. Implementation of Real-Time 3D Reconstruction Algorithm — 3D Reconstruction Based on Kinect and Monocular Visual SLAM. Computer Engineering and Applications, 2014, 50(24): 199-203.
[10] Chengwei Huang, Yan Zhao, Yun Jin, et al. Research on Feature Analysis and Recognition of Practical Speech Emotions. Journal of Electronics & Information Technology, 2011, 33(1): 112-116.
[11] Zhiyu Sun, Wei Liu. Research on the Design of Intelligent Products for Pet Travel. Art Science and Technology, 2019, 32(5): 215-216.