Conceptual Design of a Shoulder-mounted Guiding System for the Visually Impaired based on Master-Slave Coordination

Bochong Chen
Bohai University, Jinzhou, Liaoning, China

Abstract: With the advancement of artificial intelligence, embedded computing, and multi-sensor fusion technologies, intelligent guide devices tailored for the visually impaired have gradually become a key focus in accessibility research. Although traditional white canes and guide dogs can assist visually impaired individuals to some extent in their daily travels, they still face challenges such as limited sensing ranges, insufficient environmental information dimensions, and low interaction efficiency in complex urban environments. To address these issues, this paper proposes a conceptual design for a multimodal shoulder-mounted guide system aimed at the visually impaired community.

This study takes the shoulder-mounted structure as the primary information carrier and constructs a system architecture featuring "environmental perception-edge computing-multimodal feedback-primary-secondary collaboration." The system design integrates various sensors, including wide-angle vision cameras, LiDAR, and millimeter-wave radar, and leverages an edge AI computing platform to perceive and analyze obstacles, dynamic targets, and road conditions in complex environments. At the same time, this paper proposes a dual-channel information mapping strategy based on "auditory spatial cues + shoulder tactile feedback" to reduce the cognitive load that might be caused by single-speech announcements. In terms of system coordination, this paper further designs a primary-secondary linkage mechanism among the shoulder-mounted main device, the smart blind cane, and the voice terminal, and conducts a theoretical analysis of the communication and data synchronization scheme based on Bluetooth Low Energy (BLE). This paper discusses the feasibility of this guide dog system from several perspectives, including system architecture, sensor layout, multimodal perception logic, and interaction methods. The study suggests that the shoulder-mounted carrier has certain potential advantages in terms of wearing stability, multi-device coordination, and spatial distribution of information feedback, which could provide new research directions for the design of future intelligent guide dog devices. Since this paper currently focuses primarily on conceptual design and theoretical architectural analysis, and has not yet completed the engineering implementation of a full-scale physical system or conducted large-scale field tests, the related functionalities still require further experimental validation in subsequent studies.

Keywords: Visual Impairment Assistance; Multimodal Perception; Scapular-Guided Navigation; Primary-Secondary Collaboration; Barrier-Free Interaction.

1. Introduction

1.1 Research Background

Visually impaired individuals have long faced challenges in daily travel, including difficulties in environmental perception, limited ability to navigate paths, and insufficient recognition of dynamic obstacles. In complex urban environments, factors such as motor vehicles, temporary construction zones, and dense pedestrian flows further increase the risks associated with independent mobility for the visually impaired. In traditional guide methods, the detection range of white canes is relatively limited, making it difficult to provide information about the environment over medium and long distances. Meanwhile, training guide dogs is costly and their widespread adoption remains limited, thus failing to meet the growing demand. In recent years, with the advancement of artificial intelligence and multi-sensor fusion technologies, smart assistive devices for the visually impaired have gradually emerged as an important research direction aimed at facilitating their mobility. [1-2] Some studies have begun exploring the integration of technologies such as visual recognition, laser ranging, and voice interaction into guide systems, thereby enhancing these devices' ability to perceive complex environments. However, most existing devices still focus primarily on optimizing single functions-such as obstacle avoidance, localization, or voice-based navigation-without establishing effective coordination mechanisms among different system components. [3,5,7]At the same time, certain devices still have limitations in terms of long-term wearing comfort, adaptability to complex scenarios, and efficiency of information feedback.

1.2 Current Research Status at Home and Abroad

Currently, research on intelligent guide dog technologies is primarily focused on areas such as multi-sensor environmental perception, path planning, and human-machine interaction. In the realm of environmental perception, some studies leverage cameras combined with deep-learning algorithms to achieve road recognition, obstacle detection, and traffic-sign recognition. [7-8] Meanwhile, technologies like LiDAR, ultrasonic sensors, and millimeter-wave radar are also gradually being integrated into guide dog systems to enhance the devices' ability to perceive distance information and detect dynamic objects. [3,7] As for human-machine interaction, in complex urban environments, continuous voice output from speech-based systems can increase the user's auditory burden and be susceptible to environmental noise interference. Consequently, some studies have begun exploring non-visual interaction methods such as haptic feedback and spatial audio cues, aiming to improve users' efficiency in understanding environmental information through multi-channel information feedback. [9-10] Regarding device form factors, existing smart assistive devices tend to focus on single functionalities-for instance, guide canes concentrate on detecting nearby physical obstacles, while wearable glasses emphasize visual recognition and path navigation, with relatively little attention paid to the collaborative mechanisms among different devices. For example, the issue of data sharing and functional complementarity between guide canes and wearable devices remains largely unexplored in a systematic manner. Overall, current research has already made significant progress in...

Some progress has been made in environmental perception and intelligent interaction, but there is still room for further optimization in areas such as multimodal fusion, wearing stability, and collaborative design across multiple devices.

1.3 Research Objectives and Content

Addressing the existing challenges faced by current smart assistive devices for the visually impaired-such as limitations in perception range, interaction methods, and device coordination-this paper proposes a conceptual design for a multimodal shoulder-mounted guidance system tailored for the visually impaired community. The key contributions include:using the shoulder-mounted structure as the core carrier to develop a multimodal sensing system that integrates wide-angle vision, LiDAR, and millimeter-wave radar; coupled with an edge AI computing platform to enable unified processing and analysis of environmental data;designing a dual-channel interaction strategy combining "auditory spatial cues + tactile feedback from the shoulder" to reduce cognitive load potentially caused by single-speech announcements; and proposing a master-slave coordination mechanism among the shoulder-mounted main device, the smart white cane, and voice terminals, along with an analysis of their communication and data synchronization logic.

2. Literature Review

2.1 Current Status of Research on Smart Guide Devices for the Blind

With the advancement of artificial intelligence and embedded technologies, smart assistive devices for the visually impaired have gradually become a key research direction in the field of visual impairment assistance. Currently, related research primarily focuses on several product forms, including smart canes, smart glasses, and wearable assistive devices. Smart canes incorporate features such as ultrasonic ranging, GPS positioning, and voice broadcasting to enhance their environmental perception capabilities. However, since these devices are designed to operate close to the ground, their environmental sensing range is somewhat limited. [3-4,6] Smart glasses typically use cameras to capture image data from the front and combine this with object detection algorithms to identify obstacles and provide navigation guidance. While this approach enables the device to acquire a broader range of visual information [7], prolonged wear of head-mounted devices can affect user comfort. Existing wearable assistive devices-such as backpacks, chest-mounted units, or clothing-integrated structures-can reduce the burden on the head and improve the devices' adaptability to various environments[2]; yet, they still face limitations in terms of sensing range and stability. In addition, the concept of multi-device collaboration has increasingly attracted attention. Researchers have begun exploring ways to integrate the functionalities of smart canes, voice terminals, and mobile navigation devices to achieve more comprehensive environmental perception and information interaction. Overall, current smart assistive devices for the visually impaired have already made significant progress in obstacle detection, path guidance, and voice-based interaction. Some progress has been made in these areas, but there is still room for further optimization in aspects such as multimodal collaborative perception in complex environments, inter-device coordination, and long-term wearing comfort.

Multimodal sensing technology has emerged as a key research direction in recent years for intelligent navigation and assistive devices. Its core idea is to enhance the system's ability to perceive complex environments by fusing data from sensors of different types. Compared to single-sensor systems, multimodal systems can mitigate, to some extent, the impact of environmental changes on device performance, making them widely used in fields such as autonomous driving, robotic navigation, and smart assistive devices. Visual perception is currently one of the most common approaches for environment recognition in research. Object detection methods based on cameras can capture rich semantic information about the scene, including road boundaries, traffic signs, and moving objects. However, visual perception is susceptible to factors such as lighting variations, occlusions, and weather conditions; thus, relying solely on cameras makes it difficult to ensure system stability in complex environments. LiDAR boasts high precision in spatial distance sensing. By leveraging point cloud data, LiDAR can accurately determine the spatial location and structural information of obstacles, playing a crucial role in 3D environment modeling and path planning. [7] Nevertheless, LiDAR typically suffers from drawbacks such as high cost and significant power consumption. Millimeter-wave radar, on the other hand, has gradually been adopted for dynamic object detection tasks due to its superior adaptability to rainy and foggy conditions. Compared to visual sensors, millimeter-wave radar demonstrates more stable detection performance under challenging weather conditions. To improve the overall perception capabilities of a system, an increasing number of studies in recent years have focused on heterogeneous sensor fusion techniques. Such approaches typically involve steps like temporal synchronization, spatial calibration, and feature fusion to integrate data from visual, radar, and inertial sensors, thereby enhancing the system's robustness in complex environments. [3,7] However, multimodal fusion systems also bring challenges such as increased computational complexity, higher hardware power consumption, and larger system size. Thus, striking a balance between performance and portability remains a critical issue in current research. Since visually impaired individuals cannot rely on vision to obtain environmental information, how to convey navigation information through non-visual means has become a key focus in the research on intelligent guide systems. Currently, voice-based announcements are a widely used method for information delivery, [6] but in complex environments, continuous voice output can place an excessive auditory burden on users. To reduce cognitive load, some studies have begun exploring the use of tactile feedback to assist navigation. [10] By employing vibration modules to provide users with directional or hazard alerts, such approaches can minimize auditory interference. In addition, spatial audio and three-dimensional auditory cues have also emerged as promising research directions. Some studies leverage virtual surround sound to map the location of obstacles into spatially perceptible sound signals, thereby helping users develop a better sense of their surroundings. [9] Although existing research on non-visual interaction has already made considerable progress, how to achieve effective collaboration among different modalities and how to further reduce users' cognitive burden remain important issues worthy of continued investigation. [11] A review of existing research reveals that although current smart guide devices have made certain progress in environmental perception and assistive navigation, they still face several rather prominent challenges. Most devices still rely on a single core device, and there is a lack of effective coordination among different systems. Some devices still have certain limitations in terms of wearing style-for example, head-mounted devices may place excessive strain on the head after prolonged use; handheld devices, on the other hand, may suffer from unstable sensor orientation due to changes in user posture, thereby compromising the accuracy of environmental recognition. Currently, some studies overly depend on voice feedback as the primary mode of interaction. In complex environments, continuous voice output can increase the user's auditory cognitive load, while ambient noise can also impair the efficiency of information transmission. [11] Finally, although multimodal sensing systems can enhance environmental recognition capabilities, their hardware complexity, power consumption, and device size all tend to rise accordingly. For wearable assistive devices, striking a balance among system performance, weight, battery life, and wearing comfort represents a critical challenge that must be addressed in practical applications.

2.2 Research Approach of This Paper

Addressing the issues existing in current research regarding device collaboration, wearing methods, and information interaction, this paper proposes a conceptual design for a multimodal shoulder-mounted guidance system tailored for the visually impaired. Unlike traditional single-device designs, this paper explores using the shoulder-mounted structure as the core information carrier and establishes a collaborative working model of "primary device + auxiliary device." Specifically, the shoulder-mounted primary device is primarily responsible for medium- and long-range environmental perception and multimodal data processing, while the smart cane serves to provide supplementary detection of nearby obstacles. The voice terminal, meanwhile, undertakes certain semantic interaction tasks. By distributing functions among different devices, this approach holds promise for enhancing information acquisition capabilities in complex environments to some extent. In terms of information feedback, this paper attempts to combine auditory spatial cues with tactile feedback from the shoulders, thereby reducing the cognitive load that might otherwise be imposed by purely verbal announcements. At the same time, this paper conducts a holistic architectural analysis-from a theoretical perspective-of multimodal perception, primary-auxiliary device coordination, and non-visual interaction methods, aiming to provide a new research direction for the future design of intelligent guidance devices.

3. Design of a Multimodal Shoulder-Strap Guiding System

3.1 System Overall Architecture Design

The multimodal scapular-based guidance system proposed in this paper is primarily designed to assist visually impaired individuals in complex urban environments. Navigation requirements. Unlike traditional single-device guidance systems, this study aims to develop a collaborative assistive system comprising a shoulder-mounted main device, an intelligent white cane, and a voice terminal. By leveraging multimodal sensing and non-visual interaction methods, the system enhances the ability to acquire environmental information. As illustrated in Figure 1,the overall system design follows a fundamental logic of "environmental perception-information processing-feedback interaction-device collaboration." The shoulder-mounted main device is primarily responsible for collecting environmental information and processing multimodal data; the intelligent white cane provides supplementary detection of nearby obstacles; and the voice terminal handles semantic interaction and navigation requests. Different modules share information via wireless communication, thereby forming a relatively complete assistive navigation framework. In contrast to head-mounted designs, the shoulder-mounted structure can reduce the burden on the head to a certain extent while also providing a more stable mounting space for multi-sensor configurations. Since the shoulder area is close to the human body's center of gravity, sensor orientation changes are relatively minor; thus, in theory, it is better suited as a platform for multimodal sensing systems. Moreover, the shoulder-mounted structure offers greater spatial flexibility for integrating tactile feedback modules, which facilitates the implementation of directional vibration cues. Currently, this study primarily analyzes relevant solutions from the perspectives of system architecture and functional logic, with a particular focus on the feasibility of multi-device collaboration and multimodal information interaction, rather than on specific hardware product implementations.

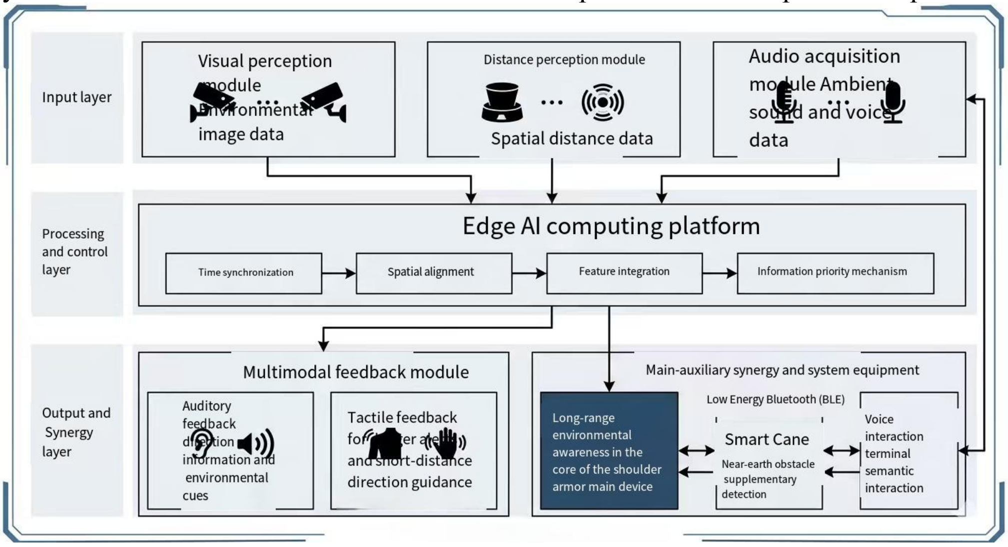

Figure 1. System Architecture Diagram

The guide system designed in this paper consists primarily of three components: the shoulder-mounted master device, the smart white cane, and a voice interaction terminal. The shoulder-mounted master device serves as the system's core component; its structure can integrate visual sensors, distance-sensing modules, and audio capture modules, enabling it to acquire information about the road ahead, obstacles, and moving objects in the user's environment. After being processed by an edge computing module, the collected data can be further converted into non-visual feedback that is easily understandable for visually impaired users. The smart white cane primarily supplements the near-field environmental detection capabilities while retaining the traditional white cane's close-range sensing function. By incorporating ultrasonic ranging and posture-sensing modules into the cane's structure, it can assist in detecting step edges, ground obstacles, and changes in walking posture. The voice terminal, on the other hand, is mainly responsible for semantic interaction between the user and the system. Users can use voice commands to request navigation instructions or inquire about location information, and the system will then generate appropriate navigational feedback based on these requests. Considering the real-time requirements in complex environments, some basic functions can be performed locally, while more sophisticated semantic analysis tasks can be handled by leveraging cloud-based models. By distributing functional responsibilities among different devices, this approach enhances the overall coverage of environmental perception to a certain extent and reduces the reliance on any single device.The hardware burden of having a single device handle all functions.

3.2 Scapular Structure Design

In wearable assistive devices, the positioning of the device carrier directly affects sensor stability, wearing comfort, and the effectiveness of information feedback. In existing studies, some devices adopt a head-mounted structure; although this design can provide a better visual field of view, prolonged use may increase the burden on the head and is prone to being affected by the user's movements.

The shoulder area serves as the primary platform for the guide system. Located near the center of the human torso, the shoulder region remains relatively stable during walking and possesses a certain load-bearing capacity, making it an ideal location for mounting various sensors and computing modules. Moreover, compared to the head region, the shoulder area offers greater spatial flexibility for distributing the device's weight, thereby reducing fatigue associated with prolonged wear. In terms of structural design, the shoulder plate is engineered using an ergonomic approach that closely conforms to the shoulder and neck regions, minimizing interference from the device's edges with the user's movements. The system's structure prioritizes lightweight construction; theoretically, composite materials or lightweight engineering materials could be used for the outer casing to further reduce overall weight and enhance the device’s durability. It should be noted that this paper currently discusses the feasibility of the scapular structure primarily from a conceptual design perspective and has not yet undergone systematic experimental verification regarding specific material properties, long-term wearing comfort, or structural strength.

To enhance the system’s perception capabilities in complex environments, as illustrated in Figure 2 ,this paper explores the integration of multiple environmental sensing modules into the shoulder structure, including a visual sensing module, a distance-sensing module, and an audio capture module. The visual module is primarily used to acquire image information from the road environment and can be positioned on the front side of the shoulder to improve coverage of the forward environment.

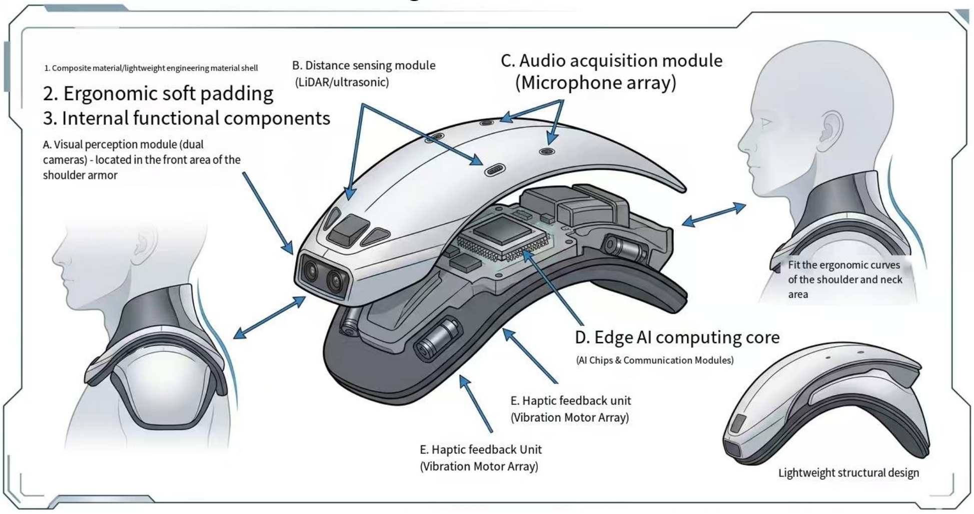

Figure 2. Schematic Diagram of Shoulder Armor Structure

The capability module. The distance-sensing module is primarily used to acquire spatial distance information from the environment, assisting in path analysis and hazard alerts. The audio-acquisition module is mainly used for voice interaction and ambient sound collection. It can be positioned on headphones or the upper area of the shoulder blades.

Since different sensors vary in terms of sampling frequency, data format, and environmental adaptability, time synchronization and information fusion remain necessary among multiple sensors. Multimodal collaborative perception can, to some extent, mitigate the limitations of a single sensor in complex environments; however, it also increases system complexity and power consumption requirements. Therefore, future research still needs to further optimize the related structural design.

3.3 Multimodal Perception Module Design

The visual perception module is primarily used to acquire image information from the road scene and assist the system in performing object recognition and environmental analysis tasks. However, visual perception is easily affected by factors such as lighting, weather, and occlusions; therefore, relying solely on visual information makes it difficult to ensure system stability. The distance perception module is primarily used to acquire spatial distance information between obstacles and the user. By incorporating distance-sensing technologies such as LiDAR, millimeter-wave radar, or ultrasonic sensors, the system can enhance its ability to understand the spatial structure of the environment. LiDAR can capture spatial information in the environment in the form of point clouds, giving it certain advantages in obstacle localization and path analysis. Millimeter-wave radar exhibits greater stability under challenging weather conditions such as rain and fog, making it a valuable complementary approach for dynamic object detection [7].

The audio module primarily handles voice interaction and ambient sound capture. Users can issue navigation requests to the system via voice, and the system will generate supportive feedback based on relevant information.

4. Multimodal Perception and Information Processing

4.1 Edge Computing Architecture Analysis

In intelligent guide systems, environmental information typically features large data volumes, rapid changes, and high real-time requirements. Therefore, relying solely on cloud servers to perform all data processing may be affected by factors such as network latency, communication stability, and privacy security. In recent years, edge computing has gradually been adopted in intelligent navigation and assistive perception systems. For vision-impaired assistance systems,as illustrated in Figure 3, local processing can, to a certain extent.Shorten the information feedback time and enhance the system's responsiveness in complex environments. A multimodal assistive guidance system can integrate an embedded AI computing platform to perform unified processing of visual, distance-sensing, and audio information. Such platforms typically feature certain parallel computing capabilities, which can be leveraged to support tasks such as object recognition, environmental analysis, and data fusion. Meanwhile, some advanced semantic interaction functions can still be implemented using cloud-based models to reduce the computational load on local devices.

Edge computing architectures also bring about challenges such as increased power consumption, heightened thermal management demands, and larger device sizes. Therefore, in wearable assistive devices, striking a balance between system performance and device lightweighting remains a key issue that future research needs to focus on.

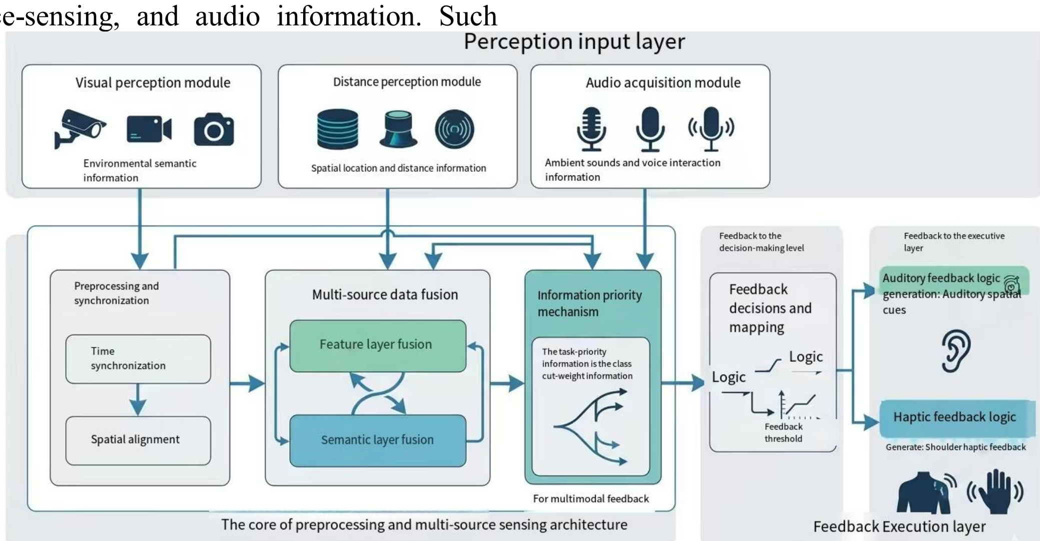

Figure 3. Information Flow and Data Processing Logic Diagram

4.2 Multimodal Information Fusion Approach

Compared to single-sensor systems, multimodal systems can, to a certain extent, mitigate the impact of environmental changes on device performance. To achieve information synergy among different sensors, the system needs to perform unified processing of multi-source data. Since different sensors vary in terms of sampling frequency, data format, and sensing range, data fusion typically involves steps such as time synchronization, spatial alignment, and feature integration. In the design presented in this paper, the visual module is primarily responsible for acquiring semantic information about the environment, such as road regions, obstacle categories, and traffic signs; the distance-sensing module mainly provides spatial location and distance information; and the audio module assists in capturing environmental sounds and speech interaction data. After acquiring different types of data, the system organizes them using a unified timeline and, combined with environmental analysis logic, generates structured environmental information.

It should be noted that this paper currently discusses the approach to information fusion primarily from a systems-logic perspective and has not yet conducted experimental validation of specific algorithmic models or fusion accuracy. Therefore, the relevant functionalities still need to be further explored and validated in subsequent research. Further optimization for real-world scenarios.

4.3 Information Synchronization and Collaboration Mechanism

Since the system consists of multiple modules-including a shoulder-mounted main device, an intelligent white cane, and a voice terminal-stable information synchronization mechanisms must be established among these devices to ensure that they can work together seamlessly. The main device and auxiliary devices can exchange data via low-power wireless communication. During data synchronization, the system needs to minimize time discrepancies between different devices as much as possible, thereby avoiding delays or conflicts in environmental information. Furthermore, given the differences in data types among various devices, the system also needs to standardize and organize environmental information uniformly. Overall, multi-device collaboration can to some extent expand the system's environmental sensing range; however, it also increases the complexity of communication management and system synchronization. Therefore, future research still needs to further analyze the efficiency and stability issues of collaboration among different devices.

The above design demonstrates theoretical feasibility in terms of wearing stability, device coordination, and information feedback.

5. Auditory and Tactile Collaborative Feedback Design

5.1 Design Approach for Auditory Feedback

Auditory feedback is primarily used to convey directional information and environmental cues to users. Compared to traditional monaural voice broadcasts, spatialized audio cues can help users develop a sense of direction to some extent. For example, when an obstacle is located in front of the user’s left side, the system can prioritize generating a cue signal through the left channel, thereby enhancing the user’s perception of the obstacle’s spatial location. Spatialized auditory cues are well-suited as a medium for providing environmental feedback over medium-to-long distances. Compared to continuous voice broadcasts, directional audio cues can reduce the need to process complex linguistic information, thus alleviating the user’s cognitive load to a certain degree. In addition to directional cues, auditory feedback can also be used to convey certain environmental information. For example, the system can distinguish between ordinary obstacles, moving targets, or hazardous areas through different types of sounds, helping users quickly grasp changes in their surroundings. Compared to lengthy voice announcements, sound-based symbolic cues can reduce information redundancy to some extent. For instance, short, sharp beeps can indicate nearby obstacles, while sustained tones can be used to signal moving objects. Dangerous targets. However, if there are too many different sound types, it could also increase the user’s learning burden; therefore, feedback design should strive to remain as simple as possible. Moreover, ambient sounds themselves can also carry certain auxiliary meanings. For example, vehicle horns, pedestrian activity, and traffic warning tones can all help users gauge the state of their surrounding environment. Consequently, during system design, we should not completely block out external ambient sounds-rather, we should do our utmost to avoid disrupting the user’s normal auditory perception.

5.2 Tactile Feedback Design Approach

5.2.1 Directional tactile cues

Haptic feedback is primarily used to assist with directional cues and hazard alerts. Compared to auditory feedback, haptic information offers greater immediacy and is less affected by ambient noise, making it well-suited as a complementary feedback modality during navigation. Vibration feedback modules are arranged in the shoulder area, and different vibration patterns at various locations provide directional cues. For example, vibration on the left side can indicate that the user should adjust their direction to the left, while vibration on the right side signals a turn to the right. Compared to handheld devices, the shoulder area has certain spatial advantages, making it more appropriate for delivering directional haptic feedback.

5.2.2 Hazard alert mechanism

In addition to directional cues, haptic feedback can also be used for high-priority hazard alerts. For example, when the system detects an obstacle in close proximity or a dynamically approaching target, it can prioritize triggering a stronger vibration signal to promptly alert the user to adjust their course of travel. Compared to voice prompts, haptic alerts generally offer greater immediacy, making them well-suited as an emergency notification method. Particularly in environments with complex background noise, haptic feedback can partially compensate for the potential interference of voice messages by ambient sounds. However, the hazard-alert mechanism must also carefully regulate the frequency of feedback. If the system triggers alerts too frequently, users may become accustomed to the vibration signals, thereby reducing the effectiveness of the alerts.

Therefore, how to establish a reasonable information-prioritization mechanism remains a key issue that future research should focus on.

5.3 Dual-Channel Feedback Collaborative Analysis

As illustrated in Figure 4,the dual-channel feedback approach proposed in this paper is primarily based on the functional division of labor principle: "Auditory feedback is responsible for environmental understanding, while tactile feedback handles immediate alerts." Specifically, auditory feedback is better suited to tasks such as providing directional information, route guidance, and descriptions of environmental conditions, because sound can convey a richer variety of information. In contrast, tactile feedback is more appropriate for short-term, high-priority alerting tasks-such as direction adjustments and danger warnings.

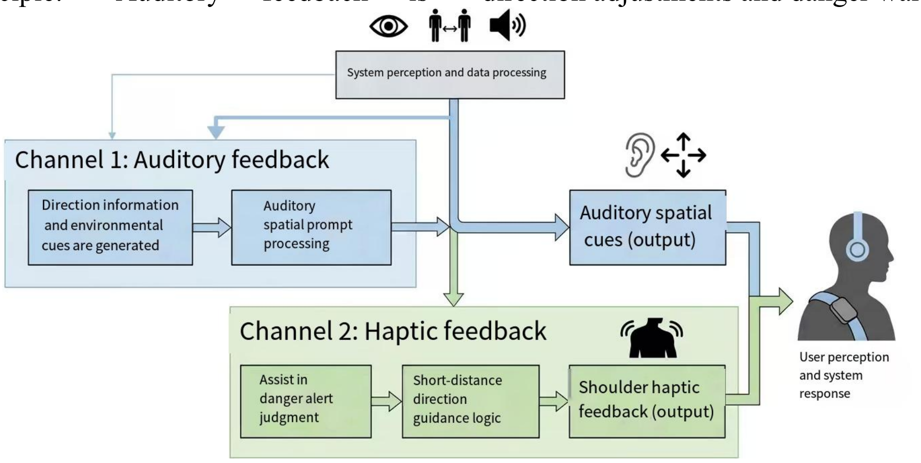

Figure 4. Logic Diagram of Dual-channel Feedback

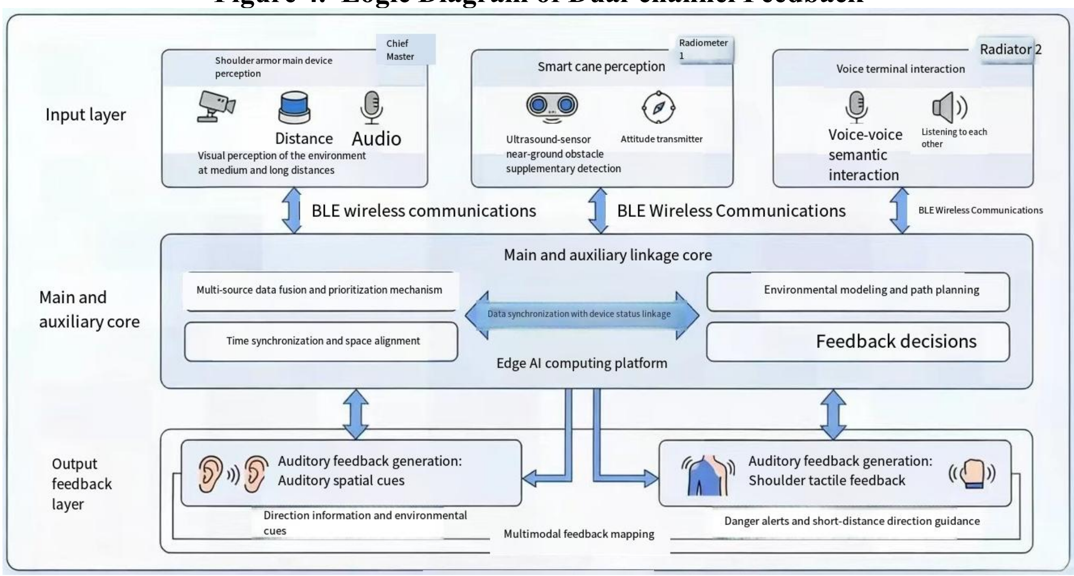

Figure 5. Main and Auxiliary Linkage Structure Diagram

6. Design of the Primary-Secondary Coordination Mechanism

6.1 Integrated Design Approach for Primary and Auxiliary Systems

This paper proposes a collaborative working approach based on a “primary device + auxiliary device” architecture. As illustrated in Figure 5,in this framework, the scapular primary device is primarily responsible for integrating environmental information and processing multimodal data; the smart white cane handles near-ground obstacle detection; and the voice terminal manages user semantic interactions and navigation requests. By distributing functions among different devices, this approach theoretically expands the range of environmental perception to some extent and reduces the computational and hardware burden on any single device.

6.2 Smart Cane Assistance Module

Although the shoulder-mounted main device can acquire environmental information within a certain range, it may still have perceptual blind spots when it comes to targets such as low obstacles, edge of steps, and ground protrusions. Therefore, retaining the close-range detection function of traditional white canes remains practically meaningful. By integrating a distance-sensing module into smart white canes, we can help obtain more detailed information about the immediate surroundings. For example, ultrasonic ranging technology can detect low obstacles ahead within a specific range, thereby providing the system with supplementary environmental data. In addition to obstacle detection, the motion state of the cane itself can also reflect the user's current walking conditions. For instance, the amplitude of the cane's swing, its tilt angle, and its contact status with the ground-all these factors may be closely related to the user's current walking environment. By incorporating a posture-sensing module into the cane's structure, we can further assist in analyzing the user's motion state. This additional information can serve as a valuable complementary reference for environmental assessment-for example, helping to identify changes in terrain such as uphill or downhill slopes, step transitions, or sudden stops. However, the results of posture analysis can easily be influenced by users' individual usage habits; therefore, the relevant logic still needs further optimization in conjunction with actual use cases.

6.3 Collaborative Design of Voice Terminals

Voice terminals are primarily used to handle semantic interaction tasks between users and the system. Users can issue navigation requests to the system via voice, such as destination queries, route selection, and inquiries about environmental information. Some basic interaction functions can be completed locally, while more complex semantic analyses can be processed in conjunction with cloud-based models.

7. Conclusion and Outlook

7.1 Study Summary

This paper focuses on the assistive navigation needs of visually impaired individuals in complex urban environments and proposes a conceptual design for a shoulder-mounted guide system based on multimodal perception and a master-slave coordination mechanism. The study conducts a theoretical analysis of this novel wearable guide device from several perspectives, including system architecture, perception logic, information interaction, and device collaboration. In terms of system architecture, the scapular region serves as the primary information carrier, and a collaborative working model comprising a "scapular main device + smart white cane + voice terminal" has been established. Specifically, the scapular main device is responsible for medium- and long-range environmental perception and information integration; the smart white cane is used to supplement near-ground obstacle detection; and the voice terminal undertakes certain semantic interaction tasks. Compared with traditional single-device architectures, this design approach places greater emphasis on functional division and information coordination among different modules. In terms of environmental perception, this paper discusses multimodal fusion approaches among visual perception, distance perception, and audio information, and analyzes the complementary relationships among different perception modalities. Multimodal collaboration can, to a certain extent, mitigate the limitations of single sensors in complex environments, providing richer environmental information for visually impaired assistance navigation. In the field of human-computer interaction, this paper proposes a dual-channel interaction approach that combines auditory feedback with tactile feedback. Specifically, auditory feedback is primarily used for directional guidance and environmental information presentation, while tactile feedback serves to provide danger alerts and short-distance directional cues. This design aims to reduce users' reliance on single-speech announcements to some extent and enhance their ability to discern information in complex environments.

This paper primarily explores the scapular-based guide system from the perspectives of conceptual design and system logic, aiming to provide a new design approach for future research on intelligent guide devices. Currently, this paper remains largely at the stage of theoretical analysis and architectural design; a complete engineering implementation and large-scale scenario validation have yet to be carried out. Therefore, the related functions and...

The actual effects still need to be further tested in subsequent studies.

7.2 Limitations of the Study

Although this paper presents a preliminary design and theoretical analysis of a multimodal scapular-based assistive guidance system, certain limitations remain due to constraints in research time and engineering conditions. First, the paper has not yet completed the construction of a full-fledged hardware prototype; thus, the overall system performance still lacks validation in real-world scenarios. Second, the discussion on multimodal fusion algorithms and information synchronization mechanisms in this paper is primarily based on existing research findings and theoretical derivations, and no systematic testing of specific algorithm performance has been conducted yet. Consequently, the system's real-time responsiveness and stability in complex, dynamic environments remain somewhat uncertain. Furthermore, given the stringent requirements for wearable assistive devices in terms of weight, power consumption, and battery life, striking an optimal balance between system performance and device lightweighting will be a key issue that future research must address. Finally, since users may differ in their ability to adapt to auditory and tactile feedback, the effectiveness of the dual-channel interaction approach in practical applications still needs further verification.

7.3 Future Research Prospects

Future research could further focus on areas such as system engineering implementation, interactive optimization, and practical scenario applications. In terms of hardware implementation, subsequent research could aim to complete the development of a prototype for the scapular-based blind navigation system and conduct tests in real-world road environments to further analyze the system's stability and applicability in complex scenarios. Meanwhile, by integrating lightweight materials with low-power computing platforms, it would also be possible to optimize the device's weight and battery life. In terms of perception and interaction, future research could further explore more natural non-visual feedback methods. For example, by integrating more refined spatial audio cues with tactile encoding techniques, we can enhance users' efficiency in understanding environmental information. Additionally, we could investigate the synergistic mechanisms among different feedback modalities to reduce cognitive load in complex environments. In terms of system coordination, future research could further analyze the issues of data synchronization and task allocation between primary and auxiliary devices, and explore more stable wireless communication methods to enhance the overall coordination capability of the system.

References

[1] Feng JinYu, Zhang Kuixing, Zhang Tielin, et al. Research Progress on Path Planning for Assisting Visually Impaired Individuals in Navigation with Integrated Obstacle Detection[J]. Computer Systems Applications, 2024, 33(04):50-59.

[2] Zhao Qingshuai, Teng Hao, Liu Qian, et al. Research Status of Intelligent Guide Devices [J]. Shandong Industrial Technology, 2022, (02): 22-28.

[3] Wang Bingyun, Wang Bingxiang, Xie Yinzhong. Design of an Intelligent Guide Cane with Multi-Sensor Information Fusion [J]. Daily Appliances, 2024, (06):77-81.

[4] Zhang Fangfang, Li Xiaoxuan, Yang Shuqiang, et al. Design of an Intelligent Navigation Cane for the Blind [J]. Electronic Quality, 2024, (01):51-57.

[5] Lu Nai-xun, Luo Kai-yun, Wang Qi, et al. Research on the Design of an Intelligent Navigation System for Blind People's Travel [J]. Internet of Things Technology, 2024, 14(01):109-113.

[6] Liu Qingtao, Yu Zhiqiang. Design of an AI-powered voice-interactive, foldable guide cane integrated with a seat-based on ToF radar ranging and GPS positioning [J]. Information Recording MaterialsMaterial, 2023, 24(03):132-134+138.

[7] Wang Yuanxin, Yan Kaihang, Deng Juxin, et al. Research and Design of Smart Voice-Based Obstacle-Avoiding Blindfold Glasses Based on CNN and LiDAR Fusion [J]. Electronic ManufacturingZuo, 2024, 32(12):45-47+44.

[8] Shi Yali, Gai Kanghao, Zhang Xinli, et al. Design of a Guide Dog Device Based on MaixPy and the YOLOv5 Model [J]. Modern Information ScienceTechnique, 2022, 6(04):185-188. DOI: 10.19850/j.cnki.2096-4706.2022.04.049.

[9] Lu Xiang, Wu Mian, Yang Xinyu, et al. Design of a Virtual Surround Sound Guiding

System for the Blind Based on Audiovisual Perception Substitution [J]. Journal of Shandong University of Science and Technology (Natural Sciences)Edition), 2022, 41(06):120-128.

[10] Zhang Fengjun, Dai Guozhong, Peng Xiaolan. A Review of Human-Computer Interaction in Virtual Reality [J]. Chinese Science: InformationScience, 2016, 46(12):1711-1736.

[11] Ding Yi, Guo Fu, Hu Mingcai, et al. A Review of User Experience Research at Home and Abroad [J]. Industrial Engineering and Management, 2014, 19(04):92-97+114.

