International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Design of Intelligent Guide Cane Based on SLAM Technology and BeiDou Navigation

Lu Kaiyue, Chang Hao*, Huang Shuqin
School of Management Science and Engineering, Anhui University of Finance and Economics
No. 962 Caoshan Road, Bengbu City, Anhui Province, China 233030
*Corresponding author: Chang Hao, Email: 007changhao@163.com

Abstract—*As a special group in society, visually impaired individuals face many inconveniences in their daily lives, especially when it comes to mobility, where they require more care and assistance. However, factors such as the frequent occupation of blind paths, the limited number of guide dogs, and the inconvenient use of ordinary guide canes make traveling a significant challenge for them. In order to provide more safety assurance and convenience for visually impaired individuals, this paper proposes an intelligent guide cane based on SLAM (Simultaneous Localization and Mapping) technology and BeiDou navigation. This product integrates multiple technologies and systems, including SLAM technology, BeiDou navigation, voice interaction, and some others, to realize three major module functions: environmental perception and path planning and path planning, positioning and navigation, and voice interaction. It aims to offer more precise positioning and navigation services for visually impaired individuals. Through this product, the visually impaired will be able to move more independently and safely, enhancing the safety and convenience of going out. The design and implementation of this product will also provide a more humanized and comfortable experience for the visually impaired, which not only improves their quality of life, but also helps to promote social care and concern for the visually impaired, having significant social value and practical application significance.*

Keywords—*BeiDou navigation: guide cane: SLAM technology: visually impaired individuals: voice interaction.*

I. BACKGROUND

According to data from the China Disabled Persons' Federation, there are over 17. 31 million visually impaired individuals in China. Currently, most visually impaired people walk with the aid of ordinary blind canes and guide dogs. However, there are many limitations of those ordinary blind canes, for example, they struggle to detect hanging obstacles, can't locate specific positions quickly and accurately, and find walking by knocking on the ground to navigate slow and laborious. Although guide dogs can assist visually impaired individuals well, their numbers are extremely limited due to long training periods, high costs, and limited life span, and China has the largest number of visually impaired individuals in the world, so just relying on guide dogs alone is far from meeting the vast market demand. The above situations make it exceedingly difficult for visually impaired individuals to go out.

To provide visually impaired individuals with more safety and convenience during going out, this paper proposes an intelligent guide cane based on SLAM technology. By integrating multiple technologies and systems such as SLAM technology, BeiDou navigation, and voice interaction, it can help the visually impaired go out more independently and safely. The product is not only able to perceive the surrounding environment and autonomously plan routes but is also equipped with an intelligent voice interaction system for real-time communication, guiding users safely and providing a more humanized and comfortable user experience.

II. INTELLIGENT GUIDE CANE SUPPORT TECHNOLOGY

A. SLAM Technology

Simultaneous Localization and Mapping (SLAM) technology, first proposed by Smith, Self, and Cheeseman in 1988, is widely used in autonomous driving and robot navigation [1]. This technology can rely on sensors, such as cameras, Light Detection and Ranging(LiDAR), and inertial measurement units (IMU) [2] to collect data from the surrounding environment, and then build a model of the environment in real time through the collected data, as well as estimating its own movement state and position for autonomous navigation and mapping, so as to realize fully autonomous mobility. Compared to traditional Structure from Motion (SfM), SLAM's ability to process data in real-time or near real-time makes it particularly suitable for applications requiring immediate feedback.

B. BeiDou Navigation

BeiDou Satellite Navigation System (BDS), independently developed and designed by Chinese teams, is the world's third mature global satellite navigation system following GPS and GLONASS. The system provides highly reliable, high-precision positioning, navigation, and timing services[3], along with short message communication capabilities. One of its significant advantages is delivering high-precision services globally, around the clock. The use of a mixed orbit satellite configuration and multi-frequency navigation signals has significantly improved the accuracy of its services.

With the successful global networking of the BeiDou system, its global application prospects have expanded, playing a crucial role in national security and economic and social development. To further promote the application and development of the BeiDou system, China has integrated satellite navigation and positioning reference station resources within the natural resources system under the "one network" requirement. A countrywide BeiDou high-precision navigation and positioning service platform has been effectively initiated for trial operation [4]. The establishment of this platform signifies a further enhancement in the high-precision service capability of the BeiDou system, providing more robust technical support for broader applications.

C. Voice Interaction Technology

Voice User Interface (VUI) is a technology that realizes the exchange of information and the execution of commands between humans and machines through the use of speech as a medium. This technology involves five key processing stages: Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Dialogue Management (DM), Natural Language Generation(NLG), and Text-to-Speech(TTS)[5]. It converts human speech into understandable commands for machines and then transforms machine responses back into spoken output, achieving human-machine interaction. Currently, this technology is widely applied in various fields such as smart homes, vehicle systems, medical scenarios, and industrial environments.

With the application and development of new technologies like artificial intelligence, big data, and cloud computing, intelligent voice interaction technology is becoming the next-generation mode of human-machine interaction. This technology not only enhances the efficiency of cultural dissemination, but also advances the intelligent transformation across multiple industries, promoting barrier-free access to information for special groups such as the visually impaired.

III. INTELLIGENT GUIDE CANE SYSTEM DESIGN FRAMEWORK

A. Functional Module Framework of Intelligent Guide Cane

Our team proposes an intelligent guide cane based on SLAM technology and BeiDou navigation. By deeply studying SLAM technology, the BeiDou system, and voice interaction technology, we integrate these technologies with the guide cane to enable its intelligence. At the same time, continuous optimization ensures that each functional module works closely together, helping visually impaired individuals to travel independently and safely.

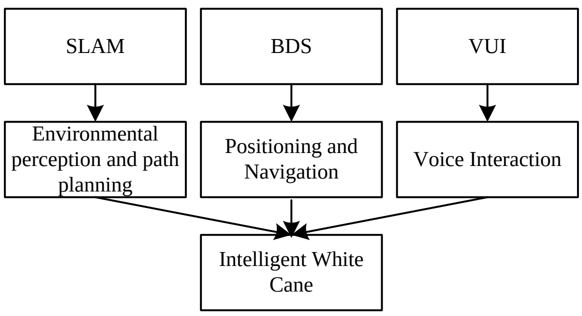

Fig. 1. Functional modules and supporting technologies of intelligent guide cane.

As shown in Fig. 1, this product includes three main functional modules:

- The environment perception and path planning module. Reasonable use of SLAM technology to complete the environment perception and path planning of the intelligent guide cane to help users optimize the travel path, avoid obstacles, and protect the user's travel safety.
- The positioning and navigation module. Utilizing China's self-developed BeiDou system to make up for the lack of SLAM technology in positioning and navigation accuracy, and improve the accuracy of positioning and navigation.
- The voice interaction module. Adopting voice interaction technology and appropriate hardware to complete voice recognition and voice broadcasting to improve the user's sense of use and increase the convenience of using the guide cane.

B. Intelligent Guide Cane Module Design

1. Environment Sensing and Path Planning Module design

As illustrated in Fig. 2, the environment perception module is implemented by introducing SLAM technology. Using specific sensors, the guide cane extracts environmental information about its current location, processes the collected information, and constructs a three-dimensional map model. In the process of traveling, the information is constantly updated, the map model is corrected, and the current position is determined according to the information collected in the process of traveling to realize the path planning. In addition, if the traveling time is too long, feature matching will be conducted to detect any route repetition, which is known as loop detection, and finally realize the environment perception and path planning function of the intelligent guide cane.

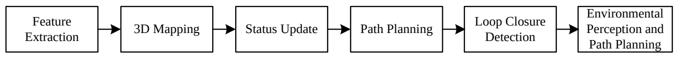

Fig. 2. Environmental perception and path planning module.

(1) SLAM Global Map Construction Based on RGB-D

SLAM technology uses cameras to capture images to build environmental maps. The RGB-D SLAM algorithm in SLAM can directly obtain the depth information of objects in images and build a global environmental map. The process of the RGB-D SLAM algorithm is shown in Fig. 3.

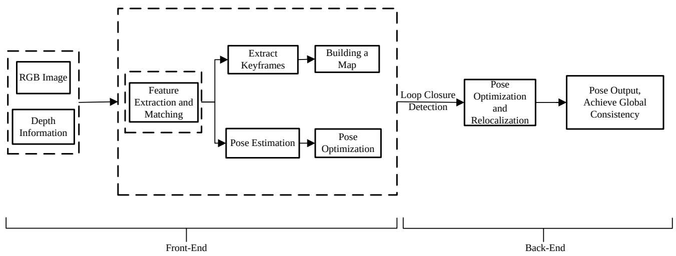

Fig. 3. RGB-D SLAM algorithm process.

RGB-D SLAM combines RGB images and depth information for simultaneous localization and mapping. The algorithm is mainly composed of two parts: the fore-end and the back-end. The fore-end is mainly responsible for processing RGB images and depth information, getting feature points for matching, completing key frame extraction and pose estimation, and then realizing initial map construction and position optimization. The back-end further optimizes the maps and poses transmitted from the fore-end to achieve pose optimization and relocation[6]. In addition, for the drift phenomenon that may exist in the fore-end data, the back end performs loop detection for noise processing, ensuring consistency in the global motion trajectory. For map construction, a pinhole camera model is used for camera modeling, obtaining the camera intrinsic matrix K and extrinsic matrix H through projection transformations, thus obtaining the transformation between the world coordinate system, the camera coordinate system, and the image coordinate system. The transformation relation between the world coordinate system, the camera coordinate system, and the image coordinate system is obtained as follows[7]:

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \rightarrow \begin{bmatrix} m \\ n \\ 1 \end{bmatrix} = F_w(X) = KH_{3\times 3}\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad (1)$$

where X_w, Y_w, and Z_w denote the coordinates of the feature point in the world coordinate system; μ and ν denote the projected coordinates of this point in the image coordinate system; $F_w(X)$ denotes the mapping function[8].

(2)Path Planning and Obstacle Avoidance Based on A* and DWA Algorithms

In path planning and obstacle avoidance strategies, both global path planning and local path planning methods are commonly used. Global planning has limited adaptability but can achieve better path selection compared to local planning. The most commonly used algorithms are the Dijkstra algorithm and the A* algorithm. The A* algorithm incorporates the breadth-first search idea of the Dijkstra algorithm, prioritizing nodes closer to the target node, thereby efficiently finding the optimal path and greatly reducing the time cost and resources. The evaluation function of this algorithm is:

$$F(n) = G(n) + H(n) \quad (2)$$

where G(n) is the actual cost from the start node to the current node, H(n) is the estimated cost from the current node to the target node, and F(n) is the combined cost[9].

Local path planning has strong adaptability and can be dynamically adjusted according to real-time sensor information, thus effectively improving the stability of the system. The Dynamic Window Approach (DWA) algorithm in local path planning has high real-time performance and relatively low complexity, which can adjust and plan in real time according to the data collected by the sensors to adapt to the uncertainty of environmental changes. The algorithm not only realizes the obstacle avoidance function, but also calculates the optimal collision-free speed to reach the target point and the acceleration during operation. Through different acceleration sizes, the range of speeds during operation is obtained, and then different speeds in the velocity space are sampled using the dynamic window method, and the coordinate information obtained from the sampling is plotted in the coordinate system, to simulate the trajectory of the user's action, and scored by the evaluation function[10]. The trajectory with the highest score is the planned optimal action route, and the corresponding linear velocity and angular velocity of this trajectory are the current optimal traveling speed. The evaluation function of the algorithm is:

$$G(n,w) = s[aP(n,w) + bD(n,w) + gV(n,w)] \quad (3)$$

$$\begin{cases} P(n,w) = \sqrt{(x_d - x_t)^2 + (y_d - y_t)^2} \\ D(n,w) = v_{\max} - v_t \\ V(n,w) = \dfrac{1}{\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}}; i = 1,2,L\ ,n \end{cases} \quad (4)$$

where α, β and γ represent the weighting coefficients. $P(\nu,\omega)$ denotes the angular difference between the end direction of the trajectory and the current target point; $D(\nu,\omega)$ denotes the nearest safe distance when the velocity is (ν,ω); $V(\nu,\omega)$ denotes the evaluation function of the current velocity, which is used for evaluating the size of the running velocity under the current trajectory; σ denotes the smoothing function; (x_t, y_t), (y_d, y_d), and (y_i, y_i) denote the (x,y) coordinate positions of the predicted node, target node, and the ith obstacle, respectively; and $v_{\max}$ and v_t denote the maximum and predicted velocities of the velocity space, respectively[7].

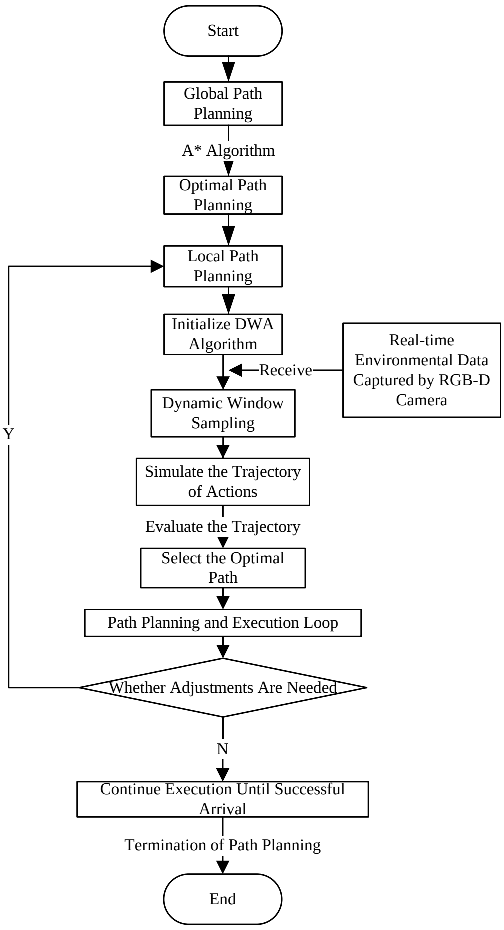

Fig. 4. A*-DWA algorithm process.

In conclusion, considering the coordination of path planning strategy and obstacle avoidance strategy, A* algorithm and DWA algorithm will be used for the overall and local planning of the path respectively, to better realize the path planning and obstacle avoidance function of the guide cane. The flow of the A*-DWA algorithm is shown in Fig. 4.

The A*-DWA algorithmic process is divided into six main steps:

- Conduct global path planning based on known map information, setting a preliminary path from the start point to the end point;
- Optimize the global path using the A* algorithm, taking into account the path length and obstacles;
- Extract key nodes from the optimized path to form the final path;
- On the basis of the optimal path, conduct local planning and initialize the DWA algorithm;
- Based on the environmental data captured by the RGB-D camera in real time, simulate and evaluate trajectories using the DWA algorithm;

Enter the path planning and execution loop, according to the optimal path and real-time environmental data, make real-time adjustments to ensure that the current path is the optimal path. Finally, determine whether the target point has been reached to confirm task completion.

2. Positioning and Navigation Module design

The BeiDou Satellite Navigation System consists of three main parts: the space segment, the ground segment, and the user segment[11]. The space segment of the BeiDou system consists of three orbital satellites, namely, geostationary orbit satellites (GEO), inclined geosynchronous orbit satellites (IGSO), and medium-Earth orbit satellites (MEO), which form a hybrid navigation constellation[12]. Fig. 5 shows the distribution of BeiDou satellites in orbit on December 3, 2024, as well as the movement trajectory. The ground segment mainly consists of three parts: the main control station, the time synchronization/injection station, and the detection station. The ground segment is responsible for the operation and control of the system's navigation tasks, ensuring the precise operation of the satellites and the accurate transmission of signals. The user segment includes basic products such as chips, modules, and antennas for BeiDou, which are compatible with other satellite navigation systems, as well as terminal products, application systems, and services[11].The main function of these products is to decode satellite signals using the integrated calculation module, providing speed, position, and time information. This enables users to achieve navigation, positioning, speed measurement, timing, and information exchange.

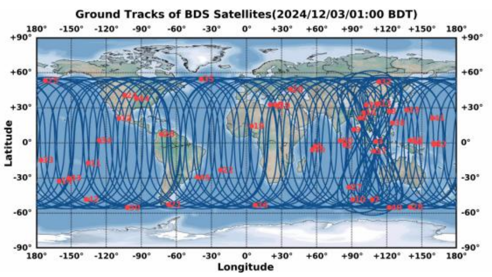

Fig. 5. Distribution of BeiDou satellites in orbit and their trajectories on December 3, 2024.

The BeiDou system provides two main services: navigation and positioning, as well as communication and digital transmission, which are complementary to SLAM technology, thereby realizing the positioning and navigation module of the guide cane. Fig. 6 gives the structure of the intelligent guide cane positioning and navigation module. First of all, the BeiDou system provides real-time location information for the guide cane, combines with voice interaction technology to build an electronic map of the user's starting point and end point, and uses the path planning module to find the optimal path and provides voice navigation; then, it uses the satellite positioning, satellite communication SMS and other functions to receive the environment perception information and import it into the electronic map to further optimize the path. Finally, it integrates real-time location information with environmental perception data, uploads the combined data to the backend for detection and recording, and completes the path.

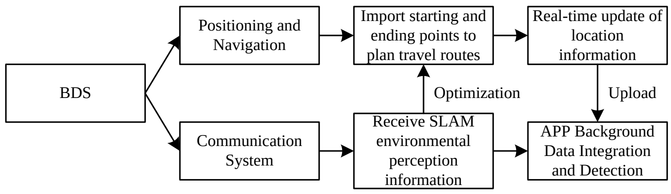

Fig. 6. Positioning and navigation module.

3. Voice Interaction Module design

As shown in Fig. 7, the voice interaction module of the smart guide cane integrates two major functions of voice recognition and broadcasting[13]. The voice recognition sub-module collects the user's voice through the microphone array and utilizes the voice chip in the central processing system of the cane to complete natural language processing, recognition, and conversion. The voice broadcasting sub-module utilizes speech synthesis technology to transform navigation information into audible guidance, thus realizing an all-round voice interaction experience. The application of voice interaction technology not only improves the intelligence level of the guide cane, but also brings more convenient and accurate navigation and walking experience to users.

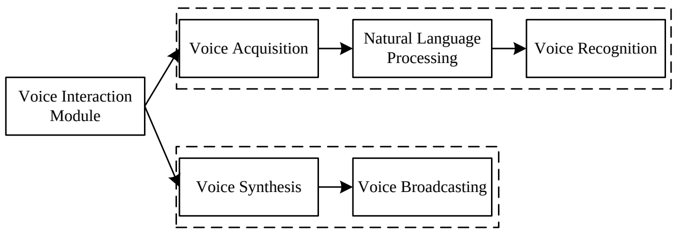

Fig. 7. Voice interaction module.

In terms of speech recognition, the real-time speech recognition WebSocket API provided by Baidu will be utilized. This speech recognition interface is based on the WebSocket protocol and enables real-time speech recognition, transforming audio streams into text in real time while achieving full-duplex communication. The specific steps are: first, establish a WebSocket connection through the URI; then, immediately send a start parameter frame containing key parameters such as application authentication information, recognition model selection, audio format, and sampling rate; next, send audio data frames to the server in real time in binary format; subsequently, the server returns recognition results in the form of text frames; after all audio data have been sent, an end frame is sent to notify the service; finally, the server will close the connection itself after the recognition is completed[14]. The flow of speech recognition is illustrated in Fig. 8.

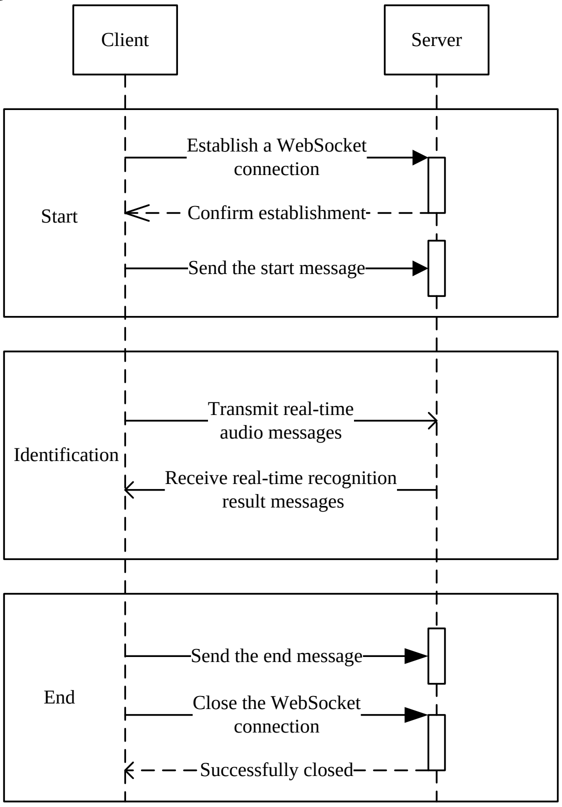

Fig. 8. Speech recognition process.

In terms of voice broadcasting, considering that the product's user group is mainly visually impaired people who prefer concise and direct feedback, the short text online synthesis service provided by Baidu was chosen. This service, based on the REST API interface for HTTP requests, quickly converts text information into audio files[15] and boasts a fast response time, enabling visually impaired users to obtain real-time feedback with ease. Furthermore, the service supports customized pronunciation annotations for polyphonic words, ensuring correct pronunciation of special words. It also provides options for speed, pitch, volume, and speaker, allowing voice output characteristics to be adjusted according to user preferences.

In specific applications, the Baidu Short Text Online Synthesis Service can convert key information such as navigation instructions, obstacle warnings, and neighborhood details into speech output in real time. For instance, it can broadcast navigation commands like 'go straight,' 'turn left,' and 'stop', as well as warning messages such as 'be careful, there are obstacles ahead!' Additionally, based on the names of stores, bus stops, and other relevant information, the service provides users with useful information about their surroundings. Users have the flexibility to adjust the speed, pitch, and volume of the voice output according to their personal preferences, thereby enhancing the overall user experience.

IV. DESIGN

The core of the appearance design of the smart guide cane lies in enhancing the user experience and precisely addressing the daily needs of visually impaired individuals through a simple, intuitive, and humanized design concept. The design of the guide cane's appearance is illustrated in Fig. 9.

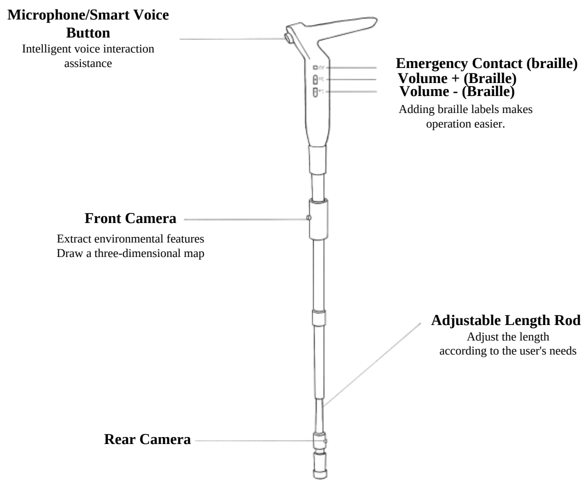

Fig. 9. Overall appearance design of the guide cane.

The pole is designed with adjustable length to accommodate the physical characteristics and preferences of diverse users. The primary structure is constructed from lightweight yet robust alloy materials, ensuring stability and durability throughout frequent telescopic adjustments. The integration of a multi-layer telescopic structure with high-strength connectors further augments the product's reliability.

The top microphone and intelligent voice button design enable users to interact with the device through voice commands, realizing the voice interaction function, which greatly enhances the convenience of use. In addition, the layout of the function buttons also takes into account the inconvenience of user operation, so the volume adjustment and emergency contact buttons are labeled in Braille to ensure that the user can recognize and carry out the corresponding operation. The Braille markings use the internationally recognized Braille system, making it easy for users to touch accurately when holding the device. This detail not only enhances the product's ease of use, but also demonstrates a meticulous grasp of the needs of the visually impaired. In case of emergency, the user can press the emergency contact button, and the cane will send a message to the preset emergency contact, and at the same time send out an alarm to the surrounding crowd for help, thus ensuring that the user can get assistance quickly in times of crisis. The emergency call button provides a solid guarantee for the safety of the user.

The strategic placement of the high and low cameras allows the guide cane to acquire more detailed environmental data, thereby significantly improving the accuracy of global map construction and enhancing the obstacle avoidance function of the guide cane.

To sum up, this product aims to bring users a more convenient, comfortable, and personalized user experience. In the design process, the special characteristics of the visually impaired are taken into full consideration, and through the design of various details, this product has the potential to become a powerful assistant in the daily lives of users.

V. Conclusion

This paper proposes an intelligent guide cane design program that integrates SLAM technology and BeiDou navigation. By incorporating SLAM technology, BeiDou navigation, voice interaction, and other technologies, the guide cane can perform environment sensing, path planning, positioning, and navigation, as well as voice recognition and broadcasting functions. On this foundation, the humanized appearance design fully takes into account the unique characteristics of visually impaired individuals and aligns with the product's functional requirements. The research and design of this intelligent guide cane aim to facilitate the application of new technologies in the field of assistive devices for the blind, ultimately enhancing the quality of life for visually impaired people in China. This product holds a promising market potential and is expected to generate significant social benefits.

Acknowledgment

This work was supported by the Anhui University of Finance and Economics Innovation and Entrepreneurship Program under Grant No. 202310378232, in part by the Undergraduate Research and Innovation Fund Project of Anhui University of Finance and Economics under Grant No. XSKY24161. This work was supported in part by the National Natural Science Foundation of China under Grant No. 61704001, in part by the Anhui Provincial Natural Science Foundation under Grant No. 1808085QF196.

References

[1] S. Zheng, J. Wang, C. Rizos, W. Ding, and A. El-Mowafy, “Simultaneous localization and mapping (SLAM) for autonomous driving: concept and analysis,” Remote Sensing, vol. 15, issue 4, pp. 1156, 2023.

[2] W. Yang, D. Li, J. Mu, and M. Liu, “Inertial-based multi-source fusion SLAM technology for complex environments,” AI-View, vol. 1, pp. 8-16, 2023.

[3] A. Chen, “Implementation of the BeiDou Navigation System in Underground Pipeline Surveying and Mapping Work,” China CIO News, vol. 5, pp. 11-14, 2024.

[4] L. Wang, “National BeiDou High-Precision Navigation and Positioning Service Platform Launched,” China Surveying and Mapping, vol. 09, pp. 22, 2024.

[5] B. Wang, Y. Wang, J. Cui, and E. Meng, “Progress in Intelligent Voice Interaction Technology,” AI-View, vol. 05, pp. 14-28, 2020.

[6] S. Zhang, L. Zheng, and W. Tao, “Survey and evaluation of RGB-D SLAM,” IEEE Access, vol. 9, pp. 21367-21387, 2021.

[7] L. He, J. Li, Y. Sun, and G. Ren, “Research on path planning and obstacle avoidance based on SLAM technology,” Automation & Instrumentation, vol. 10, pp. 108-112, 2023.

[8] Z. Li, L. Li, Z. Li, D. Kong, and X. Song, “Experimental Research on Autonomous Obstacle Avoidance and Path Planning of Intelligent Vehicle Based on RGB-D SLAM Technology,” Automobile Technology, vol. 09, pp. 55-62, 2021.

[9] X. Chen, D. Zhao, and X. Wang, “Research on Path Planning of Obstacle Crossing Search and Rescue Platform Based on Improved A* Algorithm,” Journal of Hubei University of Technology, vol. 01, pp. 18-22, 2024.

[10] X. Chi, H. Li, and J. Fei, “Research on robot random obstacle avoidance method based on fusion of improved A* algorithm and dynamic window method,” Chinese Journal of Scientific Instrument, vol. 03, pp. 132-140, 2021.

[11] N. Li and Q. Zhang, “On the Emergent Maritime Search and Rescue System based on the Beidou System,” China Maritime Safety, vol. 07, pp. 49-51, 2019.

[12] S. Ye, Y. Yan, and D. Chen, “Performance analysis of velocity estimation with BDS,” The Journal of Navigation, vol. 70, issue 3, pp. 580-594, 2017.

[13] L. Chen and L. Chen, “Design Of Multi-Sensor Blind Guidance System Based On Gps/Bd Navigation,” Computer Applications and Software, vol. 10, pp. 90-93+119, 2023.

[14] S. Niu, C. Zhang, Z. Dong, and Y. He, “Baidu AI Speech Recognition Based on WM8978 Audio Processor,” Integrated Circuits and Embedded Systems, vol. 02, pp. 42-44, 2023.

[15] Y. Yang, “Independent Research and Development of Broadcasting Virtual Newscaster System Based on Hybrid Engine,” Radio & TV Broadcast Engineering, vol. 03, pp. 40-44, 2022.