# A Research of Vision-Based Assistance Navigation Method for Visually Impaired People

Yibo Ding[1], Jeffrey Ma[2,*], Xinming Wu[3]

[1]The College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[2]University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China
[3] College of Mechanical Engineering, Tianjin University of Science and Technology, Tianjin, China

*jeffreygoodgood@sjtu.edu.cn

**Abstract.** With the rapid development of technology, life and travel have become more convenient. However, current navigation systems have limited functions, and visually impaired people still face difficulties in traveling. To address this issue and provide a convenient living environment for the visually impaired, this study proposes a navigation system for the visually impaired based on binocular cameras, aiming to enhance their independent travel ability and safety in complex environments such as indoor shopping malls and streets. This system integrates binocular camera visual input, text recognition technology, and path planning algorithms, enabling it to perceive the surrounding environment and provide precise direction guidance. By processing visual information, the system can recognize text information in the environment and associate it with the navigation system, providing clear feedback to help the visually impaired walk safely in complex environments. With the continuous advancement of artificial intelligence and smart hardware technology, this system is expected to be applied in more fields such as tourism and transportation, making greater contributions to improving the quality of life for the visually impaired.

**Keywords:** Sensor, Navigation, Image Recognition

## 1 Introduction

According to information provided by the World Health Organization in its article, at least 2.2 billion people worldwide have impaired near or far vision [1]. Population growth and ageing are expected to put more people at risk of visual impairment. At the same time, visual impairment will have a great impact on the development of young children and the psychology of adults.

For blind or visually impaired people, they can only rely on other sense such as hearing and touching to recognize direction and avoid obstacles. However, these


J. Iqbal (ed.), *Proceedings of the 2025 2nd International Conference on Mechanics, Electronics Engineering and Automation (ICMEEA 2025)*, Advances in Engineering Research 272,
https://doi.org/10.2991/978-94-6463-821-9_26

senses provide limited information and can be easily disturbed, hearing can be distorted by noise while touching is extremely insufficient in the sense of navigation. Lack of reliable directional information makes independent navigation a challenging task for visually impaired people.

Popular navigation strategies heavily rely on the usage of GPS, which are designed to assist users who can see roads and make decisions based on visible navigation aids such as maps. However, these strategies have limitations towards visually impaired people, current function of GPS systems are susceptible under more complex environment for its unstable signal, for example, high buildings can cause signal reflection and obstruction [2].

Therefore, an alternative assistance navigation method is urgently needed for visually impaired people. Past researches on this field have been focusing on mainly two types of sensors: sonar and camera [3]. A commercial sonar-based electronic travel aid is developed in past research; it is able to form landscape and provide auditory feedback to its user for navigation [4]. A bigger portion of research in this field lies in the usage of camera, a project from 2007 used strategy of matching the 3D projectile captured by camera with image to locate the user [5].

Now that much more portable devices are used, these functions can be implemented more easily. In our design, visual landmarks is a crucial factor, objects become landmarks based on their distinctiveness, stability, providing clear indicator of nearby target location, road sign can be a perfect example of landmark. Such landmarks can be used as associative cues, trigger relevant contextual information that aids in route planning and decision making [6]. Furthermore, this phenomenon can be applied to both outdoor and indoor environments as sign for directions.

This research focuses on a new vision-based navigation method that can be implemented into a navigation system in order to help visually impaired people, by capturing landmarks like road sign, this method will be able to guide the user through a real-time interaction, providing a more accurate navigation in both outdoor and indoor environment. Operating this method, the system will first generate the shortest path towards the destination, landmarks on the path are stored in the system. Users are notified once reach a landmark, guide into different directions through voice message.

# 2 Methodology

## 2.1 Image processing methods

Figure 1 is a schematic diagram of image preprocessing

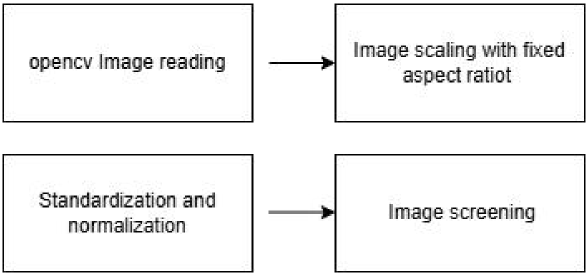


**Fig. 1.** Schematic diagram of image preprocessing (Picture credit : Original )

This part introduces the method of road sign recognition. In this paper, the features of road signs are segmented from the aspects of color and shape, and the Histogram of Oriented Gradients (Hog) feature extraction algorithm is used to extract the features of road signs. Among them， HOG is a local feature extraction algorithm based on edge feature of image shape [7].

In terms of the color segmentation,, the red, yellow and blue traffic signs were segmented by HVS (Hue, Saturation, Value) color model algorithm, and the images were morphologically processed and filtered [7]. The image denoising morphology is simplified, the image is enhanced, and the filtered non-target areas are removed.

The current street signs contain circles, triangles, and squares. There are fixed shape data standards for different forms. It is possible to calculate the roundness of C, the rectangle R, and the elongation E according to the parameter standard [4]. As shown in Figure 2, H is the height of the road sign, W is the width of the road sign, S is the area of the road sign, L is the perimeter of the road sign, and the roundness: C=4π S/L2, Rectangle: R=S/W*H, Elongation: E=min（W,H）/max（W,H）[7].

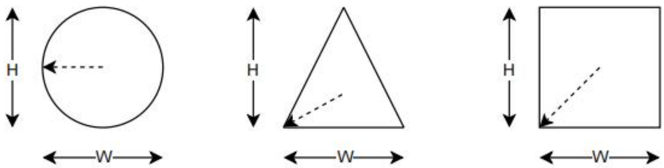


**Fig. 2.** Schematic diagram of road sign parameter feature display (Picture credit : Original )

This part introduces the ranging principle of binocular camera. The binocular camera uses the image data collected by the two cameras to calculate the distance between the image and the camera through the principle of parallax, Figure 3 shows the ranging principle of the binocular camera, p is the arbitrary point, b is the camera baseline, PR is the right camera imaging point, PL is the left camera imaging point, XL and XR are the distances from the left and right imaging points to the left plane respectively [8].

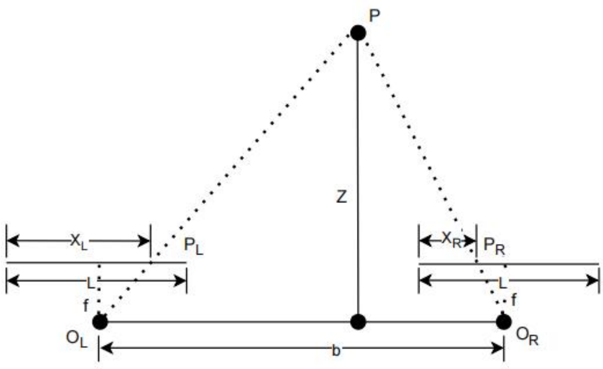

**Fig. 3.** Binocular camera captures schematics (Picture credit : Original )

The distance between the two imaging spots, $P_L$ and $P_R$:

$$P_LP_R=b-(X_L-L/2)-(L/2-X_R)=b-(X_L-X_R) \quad (1)$$

According to the principle of similar triangles:

$$b-(X_L-X_R)/Z-f=b/Z \quad (2)$$

The distance from the point p to the center surface of the projection can then be obtained[2]:

$$Z=b*f/X_L-X_R \quad (3)$$

## 2.2 Interaction between the outcome of the image processing and navigation system

### 2.2.1 Logic Design.

Our vision-based navigation method consists of two major parts, image processing and navigation.

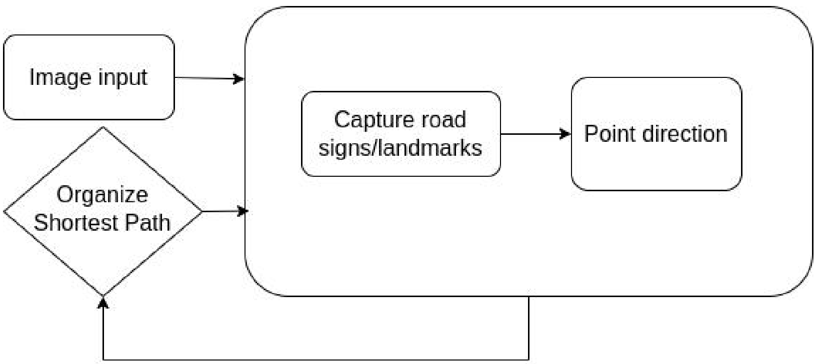


**Fig. 4.** Flow Diagram (Picture credit : Original )

Figure 4 processing process captures characters on road signs and landmarks. The navigation system operates in two steps, first, a map is formed, second, the system organizes a shortest path and navigates in accordance to the image processed.

### 2.2.2 Program Design.

The design of this method enhances the usage of OSMnx, a python package that allocates OpenStreetMap data [9].

First, to build map, using graph_from_point command in OSMnx, a map is formed consisting of nodes and edges specifically for walking routes that expand to the surrounding one kilometer of the current location, collect road names and landmarks from the OpenStreetMap database, store them into the map.

Second, the find route function find the shortest path from the starting point to the destination using Dijkstra's algorithm [10].

Third, a function to find direction at each crossroad is built, by capturing road signs at each crossroad, the navigation system match the next road with edges in the map, prompt direction information accordingly. Three general directions are defined, L as left, R as right, F as forward.

This function is achieved by calculating the bearing angle of two edges with three consecutive nodes using dot product.

For points A, B, and C with coordinates $((A_x, A_y))$, $((B_x, B_y))$, and $((C_x, C_y))$, the magnitude of vectors:

$$|AB| = \sqrt{\left\{(B_x - A_x)^2 + \left(B_y - A_y\right)^2\right\}} \tag{4}$$

$$|BC| = \sqrt{\left\{(C_x - B_x)^2 + \left(C_y - B_y\right)^2\right\}} \tag{5}$$

Bearing angle Calculation Using Dot Product:

$$\theta = \arccos\left(\frac{AB \cdot BC}{|AB| \cdot |BC|}\right) \tag{6}$$

If $45 \leq \theta \leq 135$, direction is pointed towards left, if $225 \leq \theta \leq 315$, direction is pointed towards right, else, direction is pointed forward.

# 3 Results

## 3.1 Result Presentation

Figure 5 shows the expected outcome on a visualized map:

In the visualization, the outcome of Build Map is presented by white edges and nodes, representing roads and crossroads. The red edges and nodes show the result of the Find Route, the shortest path is presented on the map.

The direction navigation is presented according to each crossroad. In the visualization, the three directions: left, right, and forward is labeled as L, R, and F. In application, the direction at each crossroad is prompt by voice message.

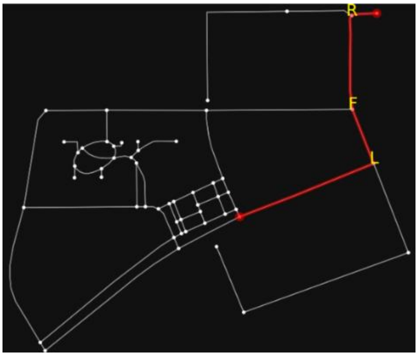


**Fig. 5.** shortest path with direction guidance on each crossroad (Picture credit : Original )

### 3.2 Implementation in indoor environments

Our navigation method has several advantages compared to traditional GPS navigation systems in indoor environment. First, it does not require the usage of any wireless input, making it exceptionally stable. Second, the recognition of location of user is based on visible objects, it is more accurate than GPS positioning.

To apply this method in indoor environments, several challenges are met and solved:

Indoor environment lacks clear signs like road signs compared to outdoor environment. Our solution to this problem is to expand sources of landmarks to plates of specific stores. At each crossroad, represented by nodes, store name or store ID is matched with the nodes, making it possible to locate the user and guide direction.

Another major challenge of navigation system design is the recognition of different floors in buildings. Our system can solve this issue by using a new map building strategy.

Instead of one map, different maps for each floor are built and stored separately in different layers following the order of floors. The system jumps to the next layer of map entering a specific node where the staircase or escalator is located.

## 4 Conclusion

This research is dedicated to exploring the application of binocular camera-based navigation methods for visually impaired individuals when they travel in indoor shopping malls and on streets, and proposes an innovative solution that combines visual input with feedback. The research mainly focuses on text recognition based on binocular camera visual input and a navigation system. Through the results of text recognition output and path planning algorithms, it guides visually impaired people in the right direction when they pass through intersections or move in indoor shopping malls, enhancing their autonomy and travel convenience, and improving their mobility in complex environments such as intersections and shopping malls. Future research can further optimize the system, improve the adaptability of the algorithm to different road conditions and indoor scenarios, and build a more universal model, especially in indoor navigation and precise navigation in dynamic environments. In addition, with the continuous progress of intelligent hardware and artificial intelligence technology, vision-based navigation methods are expected to be promoted in more fields such as tourism and transportation, making greater contributions to improving the quality of life of visually impaired people.

In conclusion, the binocular camera-based navigation system for visually impaired people not only provides a convenient and safe travel method for the visually impaired but also opens up a new direction for the application of intelligent technology in disadvantaged groups in society. It is expected to become an important force promoting inclusive social development in the future.

## Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

## References

1. World Health Organization: Vision Impairment and Blindness, World Health Organization, 2023. Available at: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment.
2. Chan, E., Baumann, O., Bellgrove, M. A., Mattingley, J. B.: "From objects to landmarks: The function of visual location information in spatial navigation," Frontiers in Psychology, 2012, 3, pp. 304.
3. Real, S., Araujo, A.: "Navigation systems for the blind and visually impaired: Past work, challenges, and open problems," Sensors, 2019, 19(15), pp. 3404.
4. Hersh, M. A., Johnson, M. A.: Assistive Technology for Visually Impaired and Blind People, Springer, London, UK, 2008, ISBN 9781846288661.
5. Treuillet, S., Royer, E., Chateau, T., Dhome, M., Lavest, J.-M.: "Body Mounted Vision System For Visually Impaired Outdoor And Indoor Wayfinding Assistance," 2007.
6. Hasan, A. M., Samsudin, K., Ramli, A. R., Azmir, R. S., Ismaeel, S. A.: "Review of GPS/INS integration in navigation systems," Australian Journal of Basic and Applied Sciences, 2009.
7. Zhao, J. D., Bai, Z. M., Chen, H. B.: "Research on Road Traffic Sign Recognition Based on Video Image," 2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA), 2017, pp. 110-113.
8. Ai, C., Geng, D., Qi, Z., Zheng, L., Feng, Z.: "Research on AGV Navigation System Based on Binocular Vision," 2021 IEEE International Conference on Real-Time Computing and Robotics (RCAR), 2021, pp. 851-856.
9. Boeing, G.: Modeling and Analyzing Urban Networks and Amenities with OSMnx, Working paper, 2024. Available at: https://geoffboeing.com/publications/osmnx-paper/.
10. Dijkstra, E. W.: "A note on two problems in connexion with graphs," Numerische Mathematik, 1959, 1(1), pp. 269-271.