Navigation Helper Helmet

NAU Northern Arizona University
College of Engineering, Informatics, and Applied Sciences

Jingwei Yang, Junlin Hai, Bo Sun, Alfred Ranasinghege Madushan Gunasekara

Introduction

Most of the information obtained by humans comes from vision. However, in 2015, 253 million people worldwide had visual impairments.[1]
Visual impairment can endanger people's lives and personal safety. Because they cannot complete many daily activities alone, these visually impaired people will rely more on others' help. So, our project is to design a device based on computer vision. This device is wearable and can use voice to describe the surrounding environment to the user. We hope that this device can help people with disabilities complete their daily activities.

Abstract

Abstract:
Our team designed a device based on deep learning to help visually impaired people improve their daily lives. The device can help the visually impaired obtain basic information about the surrounding environment. The equipment we designed is divided into five subsystems: recognition system, voice prompt system, motor system, LIDAR system, and user interface system. The voice prompt system can collect the recognition system's information, the motor system, and the LIDAR system, and then play the logically processed language to the user through the voice prompt method. In this way, the user can obtain basic information about the surrounding environment. This information includes the object's name or type, the direction, and the distance of the object. The recognition system is divided into face recognition and object recognition, respectively responsible for recognizing the type of target object and the name of the target object. The motor system is responsible for evaluating the target direction. The LIDAR system is responsible for measuring the distance between the target and the user. Its units are meters and centimeters

.**Most important requirement:**
- Object recognition for a variety of common objects. (table, chair, bottle, tree, car, bicycle, train, monitor, cat, dog) [See image 5 and image 6]
- Face recognition of familiar people. [See image 3 and image 4]
- The device has a voice prompt function, and it can describe the surrounding environment to the user. Which includes objects, people, and their position and distance.
- The device should be equipped with a user interface that allows users to control the device and customize some parameters through the user interface. [See image 1 and image 2]

Achievements Exhibition

User Interface (User login interface):

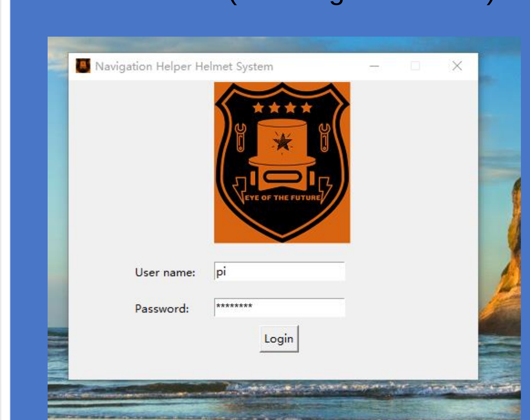

Image 1

User Interface （User control interface ）

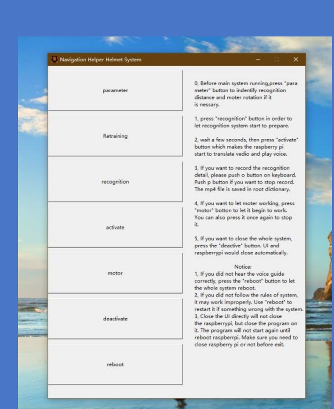

Image 2

Face Recognition

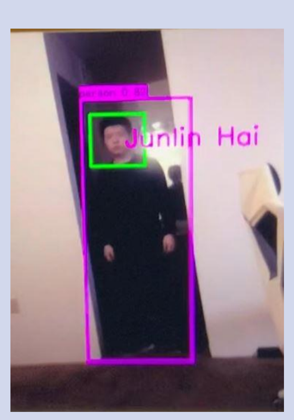

Image 3

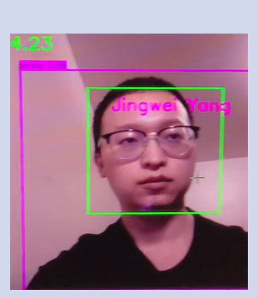

Image 4

Object Recognition

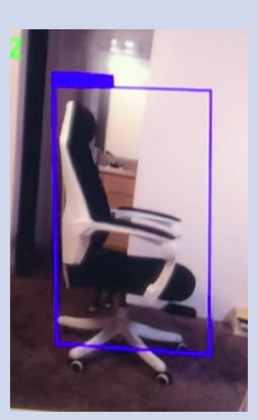

Image 5

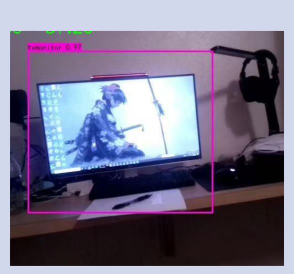

Image 6

Methods

Recognition System
Object Recognition:
The recognition system can recognize various familiar objects in life, such as tables, chairs, cars, bicycles, people, trees (Flagstaff common trees), cats, and dogs.

Face recognition:
The face recognition system judges the target person's name, but when a target is an unknown person, the face recognition system will mark the target person as "Unknown".

This part of the code is based on Python, YOLOv4 and HAAR.

User Interface
As the most important way for users to control equipment, the user interface has multiple functions.:
- Turn on the function
- Turn off the function
- Restart the system
- User-defined voice description detection and camera rotation speed
- Turn on or off the video recording function
- Add a new face recognition sample
- Turn the recoding function on or off
- A username and password are required to log in
- Control motor speed

This part of the code is based on Python.

LIDAR System
The LIDAR system can measure the distance between the target and the user. Then convert the distance data into meters plus centimeters, such as 1 meter 23 centimeters.

This part of the code is based on Python.

Voice Prompt System
The voice prompt system first collects the information of the recognition system, LIDAR system, and motor system. Then it will organize this information into a complete sentence according to the language logic that has been set. Finally, the voice prompt system will broadcast to the user through audio (the device supports Bluetooth audio and wired audio).

This part of the code is based on Python, GTTS.

Motor System
The motor system can ensure that the camera's swing range is +90 degrees to -90 degrees, and the angle of each rotation is 45 degrees. According to the angle of each rotation, the angle that the device can judge is divided into 5 types: "Left", "Right", "Left Front", "Right Front", and "Front".

This part of the code is based on Python.

.

Figure 1: Flowchart of five systems

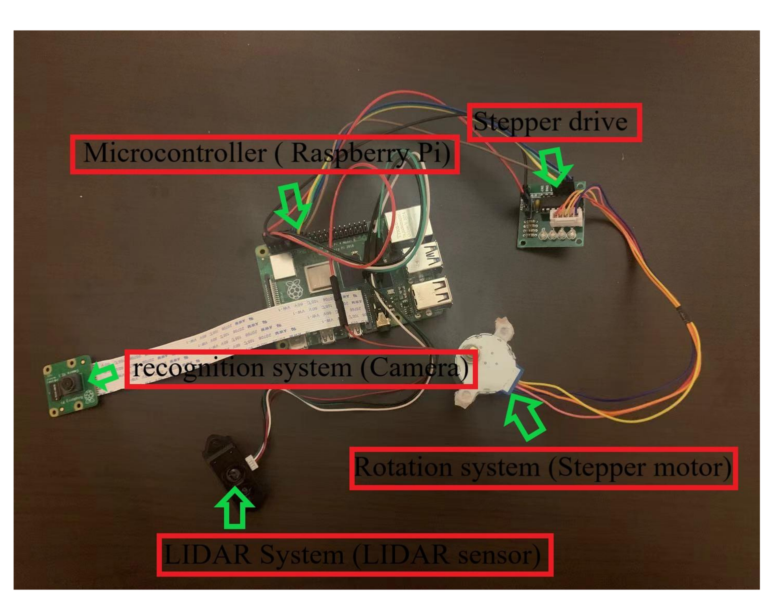

Image 7: Hardware

Image 8: Helmet

Results

The product can identify objects with a relatively high degree of recognition. According to the test, in the absence of external interference, the target recognition accuracy is as high as 90%. The product can also perform facial recognition. However, there are obvious differences in the accuracy of face recognition under different lighting conditions. According to the test, the accuracy of face recognition in outdoor conditions is as high as 85%. However, under indoor conditions or under low-light conditions, face recognition accuracy will be reduced. The accuracy of recognition will also be affected according to the dim degree of light—the darker the light, the lower the accuracy of recognition. According to many tests, its accuracy is 80% to 60% indoors. The LIDAR system can measure the distance between the target object and the user, but there are some errors. When the distance between the target object and the user exceeds 300 cm, an error of +/- 10 cm will occur. When the distance between the target object and the user is less than 300 cm, the error value is +/- 3 cm [See figure 3]. Therefore, we should choose to control the measurement distance within 300cm. The motor system can control the rotation of the camera and the lidar at the same time. When the motor is not equipped with a camera rotation device, it can be ensured that each rotation can maintain 45 degrees [See figure 2]. However, because the camera's data line is too wide, the rotation angle will be less than 45 degrees when the motor is equipped with a camera. The voice system can accurately broadcast target object information to users, such as object type, character name, distance, and direction. There will be a delay of 5 s before the start of each broadcast to ensure the voice system's stability. The user system can complete all the functions mentioned in the "method," There will be almost no bugs if the user operates as required.

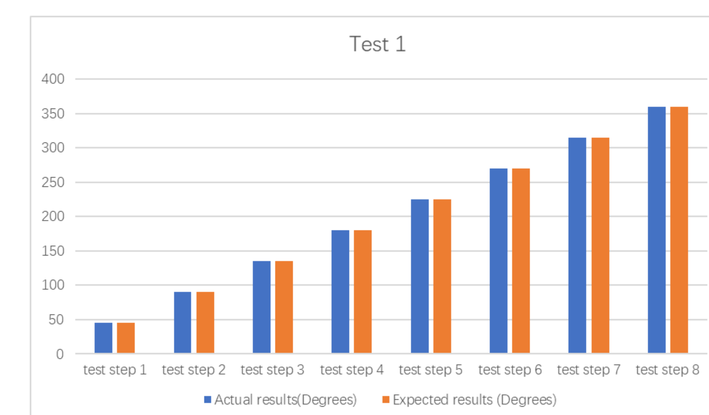

Figure 2 :Test the rotation angle of each step of the motor

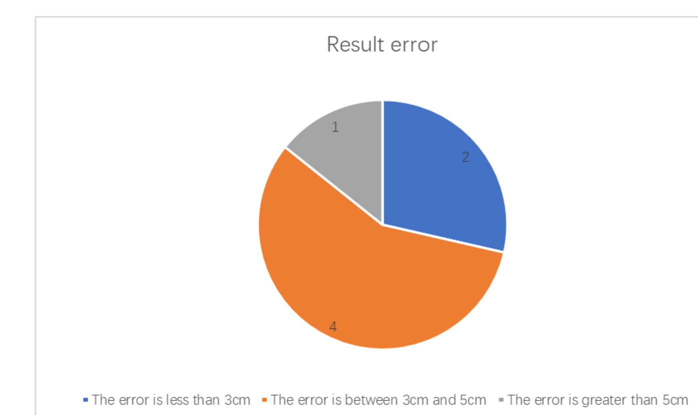

Figure 3: Error statistics of LIDAR distance measurement

Discussion

Evaluation of product design:
This product can help the visually impaired or need to obtain basic environmental information, including the type of the target object, the name, the distance to the user, and the direction relative to the user.
Main Problems:
The accuracy of face recognition is too low. When LIDAR is used in conjunction with a camera, LIDAR measurement accuracy does not meet the requirements. Also, due to the hardware system's limitations, the system response time and startup time are too long.
Envisaged solutions:
We need to use a more powerful USB camera or wireless camera instead of a Raspberry Pi camera. That can improve the clarity of the camera capture and the picture, thereby improving recognition accuracy. We can also use more samples for neural network training to improve the accuracy of recognition. To improve the accuracy of LIDAR measurement distance, we can use two LIDARs to measure distances simultaneously. When the distance error detected by the two lidars is within 5 cm, two non-parallel lidars are used to detect the same object and calculate the distance between the user and the target. This can reduce measurement errors. Since two LIDARs are used to measure the same object simultaneously and set that the difference between the two-measurement data cannot be greater than 5cm, we can be more certain that it is the measurement of the same object or one more plane. To improve the system's response speed and boot speed. We can use more powerful microprocessors.

Reference

[1] Peter Ackland, Serge Resnikoff, and Rupert Bourne, "World blindness and visual impairment: despite many successes, the problem is growing", Community Eye Health Journal, vol. 30,100: 71-73, 2017
[2] W. Boyuan and W. Muqing, "Study on Pedestrian Detection Based on an Improved YOLOv4 Algorithm," 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 2020, pp. 1198-1202, doi: 10.1109/ICCC51575.2020.9344983.
[3] Hengliang Tang, Yanfeng Sun, Baocai Yin and Yun Ge, "Face recognition based on Haar LBP histogram," 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE), Chengdu, China, 2010, pp. V6-235-V6-238, doi: 10.1109/ICACTE.2010.5579370.

Acknowledgements

We thank our Client, Professor Abolfazl Razi, for his support and guidance to our group. Dr. Razi taught us how to make our products better and also provided me with many ideas. We also want to thank our lecturer, Dr. Robert Severinghaus, for guiding and supervising our group activities. He taught us how to cooperate better and gave us a deep understanding of team spirit. Also we need to thank our GTA, Mr. Han Peng. Thank him for always patiently answering our various inquiries and for his guidance to us.