

# A Wearable System for Walking Cognitive Assistance Using People Flow Estimation and Vibrotactile Feedback in Crowded Sidewalks

Ayumu Okai, *Student Member, IEEE*, Mihoko Niitsuma, *Member, IEEE*

**Abstract**— In this paper, we develop a walking cognitive assistance system that enables visually impaired people to walk independently. Focusing on estimating the direction of people flow, we developed a wearable system by achieving wireless and minimalized vibrotactile presentation devices using two Spresense units and independence of image processing units using the Jetson Orin Nano. Additionally, by employing YOLOv8 for people detection, we have realized improved processing speed and expanded the range of people detection. In the accuracy evaluation experiments, the accuracy of estimating the direction of approaching people's flow improved, making safer walking possible.

## I. INTRODUCTION

Approximately 83% of environmental information is perceived through vision [1]. However, visually impaired people find it difficult to accurately obtain environmental information. According to a 2019 WHO survey, approximately 400 million people worldwide have moderate or severe visual impairments [2]. Visually impaired people primarily navigate using white canes, though the perceptual range afforded by them is typically constrained to a limited spatial proximity of 1-2 meters. A study by Larisa Dunai et.al has shown that visually impaired people need environmental information not only at close range but also at a distance [3].

Many technological supports have been considered to achieve safe walking. However, there has been little research focusing on understanding the environmental information necessary for independent walking by visually impaired people. Moreover, navigation [4][5] and obstacle warnings using sound or vibration [6][7] are unsuitable as methods for obtaining environmental information because they are only capable of acquiring single pieces of information. A study has been conducted on supporting environmental recognition by presenting static environments such as wall shapes [8][9], and a study using YOLOv8 to provide support for visually impaired people by obtaining information about surrounding objects through smart glasses [10].

Furthermore, in recent years, research on multimodal assistance technologies that integrate these various approaches has emerged [11], presenting a promising avenue for comprehensive environmental perception and navigation support for individuals with visual impairments. However, these studies present significant limitations in their practical application and have not adequately addressed the challenges of utilization in complex, crowded environments.

In actual walking environments, many dynamic objects exist, and people walk while predicting collisions with other pedestrians. During sidewalk congestion, smoother walking

may be enabled by following pedestrians moving in the same direction to create people flows and joining the flow [12]. However, visually impaired people have difficulty obtaining information about the surrounding people flow, which may lead to problems such as being overwhelmed by the flow and losing their sense of position.

The Previous study [13] proposed a method for estimating presenting information about the people flow, achieving dynamic obstacle avoidance focused on people flows. Additionally, a method for detecting and presenting sidewalk edges was proposed, expecting to enable movement with self-position awareness.

A depth camera-based method has been proposed to estimate people flow direction and detect sidewalk edges. The user wears a depth camera on their chest to perform people detection on the color image in front of them. The orientation of individuals is determined by calculating the change in distance between the person and the camera. A perfect flow is rare on sidewalks and the number and direction of pedestrian are expected to be non-uniform. Therefore, the system calculates congestion levels and variations in people flow directions in addition to determining the flow direction. This information is used to select the data to be presented.

For sidewalk edge detection, a method using binary semantic segmentation has been proposed. This technique distinguishes between sidewalk and non-sidewalk areas in color images for sidewalk estimation. When the user is walking near the edge of a sidewalk, the system provides information about which direction poses a potential danger due to the presence of an edge. This information is calculated by dividing the acquired image into four regions and processing each region separately. Based on these results, warnings are issued in areas deemed dangerous for walking, encouraging behavioral changes toward safer routes.

Vibrotactile feedback on shins has been proposed as a method to present information about the people flow and sidewalk edges. The visually impaired often rely on environmental sounds for obtaining information about their surroundings, and auditory stimuli may interfere with this process. The shin was chosen as the location for vibrotactile stimuli for three reasons. First, it consistently faces the direction of movement. Second, vibrations are less likely to propagate to other body parts. Third, the device enables inconspicuous wearing. An investigation of the tactile characteristics of the shin area revealed that it is possible to perceive vibrotactile stimuli at two locations on each leg [14].

Fig.1 shows the integrated system being worn. Walking experiments were conducted with sighted participants blindfolded to simulate visual impairment, using the integrated system. Participants successfully navigated a 100-meter

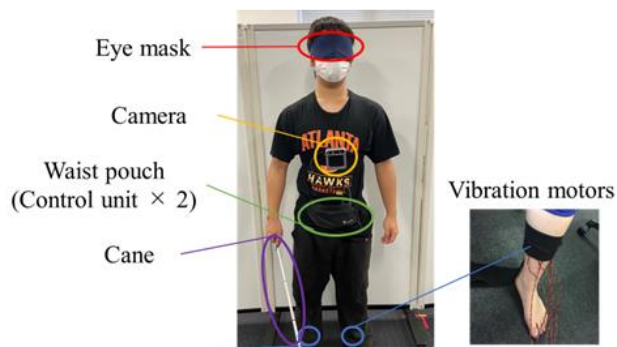


Fig. 1. View of the previous system installed

outdoor sidewalk course safely. They achieved this by following the provided information, aligning with people flow, and recognizing sidewalk edges. This suggests that the system holds promise as a useful tool for supporting environmental awareness while walking in crowded settings.

However, the previous study has several issues. The proposed method assumes that the user or an accompanying person holds a PC near the system. Therefore, it has not yet been fully made wearable. Additionally, the vibration motors (on shins) that present vibrotactile stimuli and the controller (on the waist) that controls them are connected by wires. This has been confirmed to cause discomfort due to wiring under clothing and wire breakage due to pressure from clothing and the body.

Therefore, this study aims to develop a practical system that enables faster processing and wider-range information processing by creating wireless devices, developing a wearable integrated system that does not use a PC, and partially modifying the previous system and algorithm.

## II. OVERVIEW OF THE PROPOSED SYSTEM

Fig. 2 shows the overall configuration of the proposed system. This system consists of two units: an image processing unit and an information presentation unit. The user wears a depth camera on their chest to detect the people flow from the color image in front of the user. The vibrotactile devices are worn on the shins for information presentation, while the image processor is attached to the waist.

### A. Hardware

Based on the information about walking danger areas calculated from image and depth information, vibrotactile stimuli are presented to the user's shin. For image acquisition, the Intel Realsense D435 (Intel) depth camera is used. The Jetson Orin Nano Development Kit (NVIDIA) is used for image processing, and Spresense (Sony) is used for the information presentation.

Fig. 3 shows the vibrotactile presentation device for the information presentation. The information presentation device was created by storing the Spresense, vibration motor drive board, and battery in a small bag, and sewing this bag together with a supporter. Spresense is equipped with a Wi-Fi board. Users walk by wearing these wireless vibrotactile presentation devices on each leg.

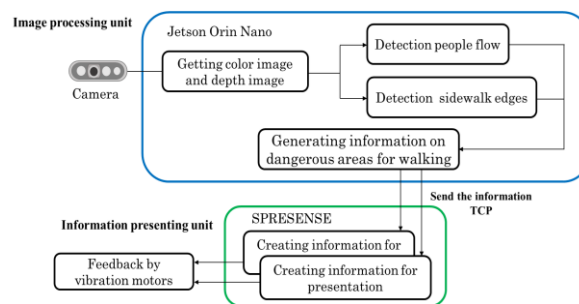
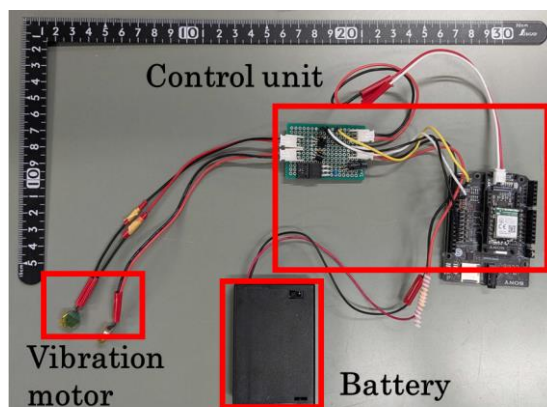


Fig. 2. Overall system configuration, illustrating the integration of image processing and vibrotactile feedback units.



(a) Outside



(b) Inside

Fig. 3. Vibrotactile feedback device, including Spresense modules, vibration motor drive boards, and batteries.

Furthermore, by using the Jetson Orin Nano for image acquisition and processing, there is no longer a need to carry a PC as in the previous study system, thus realizing a wearable system. The Jetson Orin Nano is stored in a bag worn on the waist.

### B. Software

As in the previous study [14], we calculate the estimation of people flow direction, congestion level, and variability in people flow directions using the image and depth information. During this process, the image is divided into four regions, and

calculations are performed for each region separately. Based on these calculation results, we determine the walking danger areas. The image division into four regions is determined based on the resolution at shin level in the previous research [13].

The direction estimation of people flow direction is carried out by combining people detection and tracking using YOLOv8 on the image with depth information acquisition from the RealSense camera.

Additionally, we utilize the multi-core programming capabilities of Spresense for communication with the vibrotactile feedback device and control of the vibration motors.

### III. METHOD OF INFORMATION PRESENTATION

#### A. Vibration Stimulation of Information Presentation

For the visually impaired, auditory information is crucial for recognizing the surrounding environment. Vibration stimuli are suitable for presenting environmental information as they do not interfere with auditory information and enable the simultaneous provision of information in multiple directions. As previously introduced in the introduction, the shin was used as the location for vibrotactile stimuli, following the precedent set by previous studies[13][14]. Based on the perceptual characteristics of the shin obtained in the previous study, two vibrotactile stimuli are used on each leg, presenting walking danger areas using vibrations with a cycle of approximately 1 second, synchronized with the walking cycle. Fig. 4 shows the attachment positions of the vibration motors on the shins.

#### B. Vibration Information Design for Presenting information

In this study, we use a total of four locations, the inner and outer sides of the shin on each leg, to present information about the people flow. People flow information is presented not only when there is an approaching flow of people, but also when there is congestion or when the distance to pedestrians moving in the same direction is close, even if no specific flow has formed. In such cases, the area is determined to be a dangerous walking area. Regarding the criteria for determining when the distance to pedestrians is considered close, based on the finding that 2 meters is the appropriate distance between a stationary obstacle and the user [15], we present information to the user when the distance is half of that, 1 meter or less.

#### C. Tactile Vibration Feedback devices

In this study, we developed a new compact wireless vibrotactile feedback device using two Spresense modules, one for each leg. The previous research used only one Spresense module for the vibrotactile feedback device, requiring the Spresense to be attached to the waist and wired to the shins of each leg under the clothing. This caused discomfort during walking and the wires could be cut by pressure from clothing and the body. However, two new wireless vibrotactile feedback devices were developed to simplify user operation and improve practicality.

Next, we discuss the improvement of the communication method and cycle of the new vibrotactile feedback device. In the previous device, the multi-core capabilities of Spresense

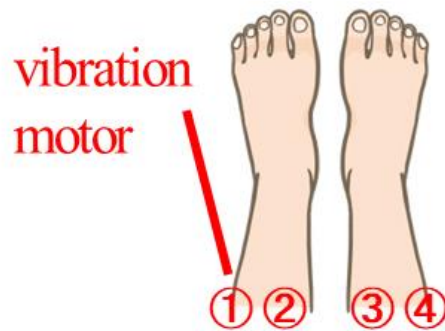


Fig. 4. Vibration motor locations

were utilized, with the main core receiving information through communication and the sub-core presenting vibration information. The sub-core performed an interrupt process every second to achieve a cycle of presenting information every second. However, this method was inadequate because the sub-core could not acquire multiple information from the main core during interrupt processing. As a result, multiple data receptions were missed consecutively, leading to incorrect information presentation. In this research, by utilizing shared memory for communication between cores, we achieved a method of receiving and presenting information at high speed without interfering with each other's operations. From this, the device proposed in this study has significantly improved processing speed compared to the previous device.

### IV. IMPROVEMENT AND ACCURACY ASSESSMENT OF METHODS FOR ESTIMATING THE DIRECTION OF PEOPLE FLOW

Our proposed system aims to improve the accuracy and expand the detection range of people detection by adopting the state-of-the-art object detection model, YOLOv8. Furthermore, to address the depth information errors caused by the expanded people detection range, we propose a novel method for the people flow direction estimation. This method combines two existing methods: one estimates the people flow direction based on depth changes [14], and the other estimates the people flow direction based on changes in the size of human regions in images [13].

#### A. People Flow Detection

This study focused on estimating the direction of people flow. As shown in Fig. 5, it aims to easily identify the people flow by dividing the color image in front of the user into four regions and detecting and tracking individuals in each region.

In this study, it was necessary to partially modify the algorithm due to hardware changes. Specifically, this involved improving processing speed and expanding the search range. With the conventional system, processing became too heavy to operate on an SBC (Single Board Computer), making real-time processing difficult. Additionally, even when depth information was available, delays in people detection sometimes caused slight delays in the people flow detection. Therefore, to improve the system's robustness, we newly employ YOLOv8 for people detection, enabling faster processing and people detection over a wider range. For a



Fig. 5. Camera image divided into four regions for estimating people flow direction

person's coordinates, we use the center coordinates of the bounding box of the person area detected by YOLOv8.

### B. Estimation of People Flow Detection

In this study, we assume that the people flow on sidewalks consists of two types: flows in the same direction as the user's progress or in the opposite direction. The movement direction of each person is determined by evaluating the change in distance from the depth camera to each person detected by YOLOv8. However, when depth information is unavailable or when a person is more than 20 meters away from the depth camera, we determine the direction by evaluating the change in area of the bounding box of the person region detected by YOLOv8. The direction of people flow in each region is estimated by the number of people moving in the same direction. When the numbers are equal, we compare the areas of the bounding boxes of the person regions, and the direction occupying a larger area becomes the direction of the people flow. There are three types of individual direction judgments: approaching, moving at the same speed, and moving away. Based on these individual direction judgments, we determine whether the overall flow is in the same direction as the user's movement, in the opposite direction, or if there is no significant people flow. When two or more individuals in the same area are judged to be moving at the same speed, we determine this as a flow moving away from the user. If only one person is moving at the same speed, we determine that there is no significant flow. For visualization purposes, approaching flows are indicated in red, moving away flows in blue, and same speed flows in green.

Fig.6 shows the flow of the people flow detection estimation method using depth information. Depth information becomes less reliable and contains larger errors when the distance is too great. Therefore, by adjusting the method for estimating people flow direction based on depth information, we use accurate the depth data where available and rely on the image data for people flow direction in other cases. This allows us to estimate the direction of people flow over a wide range while maintaining accuracy as much as possible. A low-pass filter was applied to reduce the subject's motion blur by referencing previous walking data

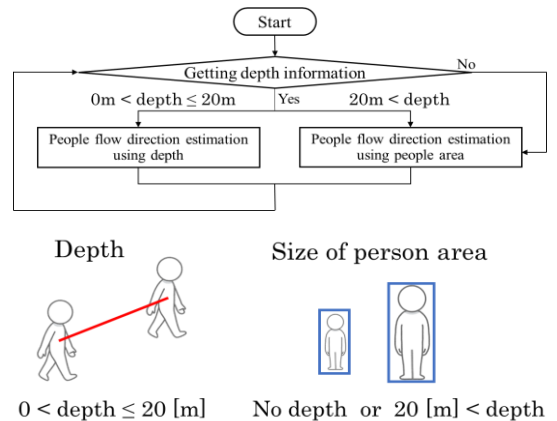


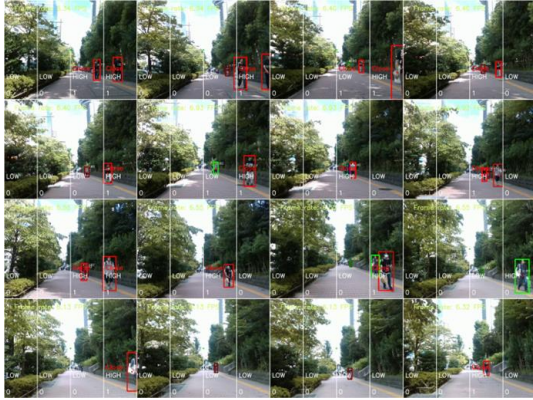
Fig. 6. Estimating the direction of people flow using depth

### C. Accuracy Evaluation of people flow direction estimation

Due to modifications in hardware and some algorithms, we conducted an accuracy evaluation of people flow direction estimation. Measurements were taken over a 30-minute period on the sidewalk in Kasuga Street Tomisaka, Bunkyo city, Tokyo. A sensor was attached to the chest of a 173 cm tall male to capture walking videos, detect pedestrians, and acquire distance information. The accuracy of the people flow direction estimation in each area was obtained by calculating the agreement rate between visual evaluation and the estimation results of the proposed method every 5 seconds. Visual evaluation was conducted as shown in Fig. 7, by comparing several frames before and after the target frame to determine whether the flow was approaching or moving away.

### D. Result and Discussion

TABLE I shows the accuracy of people flow direction estimation for the previous system and the proposed system, while TABLE II presents detailed accuracy evaluation results of the proposed method for the people flow direction estimation. The results indicate that the proposed system achieved accuracy comparable to that of the previous system. While the proposed system improved the accuracy of approaching flows compared to the previous system, the accuracy of away flows decreased. This reduction in accuracy is likely due to the specific implementation of the same-speed assessment algorithm and its parameter settings. A person was moving away, but the system judged them to be moving at the same speed. Since this person was the only one in that area, the system erroneously determined that there was no flow. Looking at Fig. 7 (b), we find several instances where a single person within a departing flow is partially judged to be moving at the same speed. This explains why such misclassifications occur. A closer examination of the results reveals that a large proportion of away flows were misclassified as equally fast flows. However, the primary goal of this research is to enable safe walking by allowing users to recognize the direction of potential walking hazards and modify their behavior. Therefore, the accuracy of information about approaching flows is considered most important. Consequently, we believe that distinguishing between away flows and equally fast flows



(a) Assessment of approaching flows



(b) Assessment of away flows

Fig. 7. Visual and system-based assessment

has a minimal impact on safe walking. From the perspective of walking safely with the flow of people, the proposed system is

considered to have improved accuracy compared to the previous system. Regarding the accuracy of people flow estimation at close distances, as the depth camera's precision increases with proximity, the flow estimation accuracy also improves. However, extremely rare misclassification events might still occur. In such cases, there is a potential risk of failing to notice a pedestrian and causing a collision. Nevertheless, for pedestrians in proximity, we have confirmed that misclassifications do not consecutively occur across multiple frames. Therefore, we consider this risk to be almost nonexistent.

Furthermore, the frame rate was used as an indicator of processing speed. The average frame rate of the previous system was 7.7 FPS, whereas the proposed system achieved an average of 22.94 FPS on a PC with almost identical specifications as the one used for the previous system and an average of 6.56 FPS on a Jetson Orin Nano. This indicates that the proposed system has significantly improved processing speed compared to the previous system.

Finally, we discuss the range of people detection. Fig. 8 shows people detection results of the previous research system and the proposed system when real-time processing was conducted at the same location. Comparing the people detection ranges, we find that while there are people unrecognized by the previous research system, our system is capable of accurately detect individuals even at greater distances. From this, we conclude that the people detection range has been expanded. The previous research system could only recognize people flow up to approximately 20 meters away, where the reliability of depth information could be ensured. In contrast, although the proposed system's accuracy is inferior to that of depth information, it succeeds in handling people flow beyond 20 meters by utilizing the changes in size of people regions between frames. As a result, we believe that our system has become more robust in terms of walking planning, information transmission, and adaptation to

TABLE I . Accuracy of people flow direction estimation for previous and proposed systems

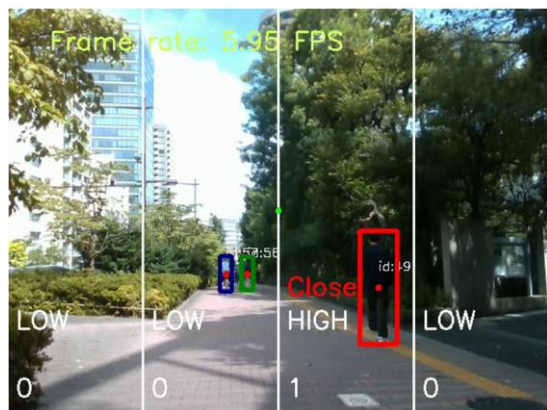
	<i>Previous method</i>	<i>Proposal method</i>
<i>People flow approaching [%]</i>	80.6	93.75
<i>Pople flow moving away [%]</i>	95.2	79.17

TABLE II . Detailed Accuracy Evaluation of the Proposed System for Each Flow

	<i>Estimated approaching people flow</i>	<i>Estimated moving away people flow</i>	<i>Estimated no people flow</i>
<i>People flow actually approaching [%]</i>	93.75	0.0	6.25
<i>Pople flow actually moving away [%]</i>	4.17	79.17	16.66
<i>Actually no people flow [%]</i>	0.56	0.0	99.44



(a) People detection range of the previous



(b) People detection range of the proposed system

Fig. 8. Comparison of people detection ranges between the previous and proposed systems, highlighting extended range capabilities.

complex environments. However, when considering rapidly moving objects such as cyclists or joggers, high-precision depth information beyond 20 meters remains crucial. Consequently, the method utilizing changes in people region sizes should be considered a supplementary approach, ideally applied only in situations where depth information is unavailable. This limitation suggests that while our proposed system expands detection capabilities, further refinement is necessary to ensure reliable tracking of fast-moving subjects across extended distances.

## V. CONCLUSION

In this paper, we proposed a method for estimating the people flow direction to support independent walking for visually impaired people. To create a wearable system compared to the previous systems, we implemented wireless functionality for the vibrotactile feedback device and made the image processing unit independent. Additionally, we improved parts of the algorithms for people detection and the people flow direction estimation. We conducted accuracy evaluation experiments and confirmed improved precision of approaching flow and speed of processing. In the future, we

plan to conduct walking experiments using the proposed system. Based on our experimental findings, we will focus on refining the algorithm and optimizing the presentation methodology to further enhance the system's accuracy.

## ACKNOWLEDGMENT

This study was supported by the Sony Semiconductor Solutions Corporation.

## REFERENCES

- [1] N. Sahoo, et al., "Design and Implementation of a Walking Stick Aid for Visually Challenged People," *Sensors*, vol.19, no.1, p.130, 2019.
- [2] WHO, "World report on vision", p44, 2019
- [3] L. Dunai, et al., "Real-Time Assistance Prototype - a new Navigation Aid for blind people," 36th Annual Conference on IEEE Industrial Electronics Society, pp.1173-1178, 2010.
- [4] S. Kayukawa, et al., "How Users, Facility Managers, and Bystanders Perceive and Accept a Navigation Robot for Visually Impaired People in Public Buildings," 2022 31st IEEE RO-MAN, pp. 546-553, 2022.
- [5] A. Xiao, et al., "Robotic Guide Dog: Leading a Human with Leash-Guided Hybrid Physical Interaction," 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 11470-11476, 2021.
- [6] V. V. Meshram, et al., "An Astute Assistive Device for Mobility and Object Recognition for Visually Impaired People," *IEEE Transactions on Human-Machine Systems*, Vol. 49, No. 5, pp. 449-460, 2019.
- [7] P. -J. Duh, et al., "V-Eye: A Vision-Based Navigation System for the Visually Impaired," in *IEEE Transactions on Multimedia*, pp.1567-1580, 2021.
- [8] R. K. Katzschmann, et al., "Safe Local Navigation for Visually Impaired Users With a Time-of-Flight and Haptic Feedback Device," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pp. 583-593, 2018
- [9] T. Watabe, et al., "Mental Map Generation Assistance Tool Using Relative Pitch Difference and Angular Information for Visually Impaired People", 2013 IEEE 4th International Conference on Cognitive Info communications, pp.225-260, 2013.
- [10] Y. -L. Poy, et al., "Standalone Smart Glass System for the Blind and Visually Impaired", 2024 IEEE 14th Symposium on ISCAIE, pp. 239-244, 2024.
- [11] Li. Guoxin, et al., "Sensing and Navigation of Wearable Assistance Cognitive Systems for the Visually Impaired," *IEEE Transactions on Cognitive and Developmental Systems*, pp. 122-133, 2023.
- [12] Y. Suma, et al., "Anticipation effect in pedestrian dynamics: modeling and experiments," *Physica A: Statistical Mechanics and its Applications*, Vol.391 Issues 1-2, pp. 248-263, 2012.
- [13] K. Fujino, et al., "Extraction and Presentation of People Flow Information to Assist Dynamic Environment Recognition for Visually Impaired People," 45th Annual Conference of the IEEE Industrial Electronics Society, pp.6883-6889, 2019.
- [14] Takumi Sato, et al., "Evaluation of Cognitive Assistance for Visual Impairment to Recognize Dynamic Walking Environments Using Vibrotactile Stimulation", the 49th Annual Conference of the IEEE Industrial Electronics Society, pp.1- 6, 2023.
- [15] M. A. Soto et al., "Drone Navigator: Using Leashed and Free-Floating Quadcopters to Navigate Visually Impaired Travelers," *International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 300-304, 2017.