HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 4 (2023)

Taking Control: A Novel Galvanic Stimulation Device for the Visually Impaired

Jing Quan Peng[1], Elizabeth Strehl[2], Peter Mbua[3], Susan Leong[#] and Mert Kaval[#]

[1]David Thompon Secondary School High School, Canada
[2]University of Michigan, USA
[3]University of Florida, USA
[#]Advisor

ABSTRACT

The white cane has been the prominent and widely used mobility aid by visually impaired persons for many years, however, there are some limitations associated with the white cane mobility aid device. Primarily, the white cane exhibits restricted capability in detecting ground-level obstacles in proximity and does not provide reliable detection of aerial obstacles. Our work proposes a device for the safe navigation of visually impaired persons utilizing a non-invasive galvanic vestibular stimulation (GVS) technique. By delivering a signal (1-1.2mA) delivered behind the ear via electrode pads to the vestibular system, we induce a sensation of steering thereby facilitating navigation. Our proposed device utilizes a combination of Intel's D435 depth camera and an object detection model, YOLOv5, to identify and process the detection of obstacles within a 3-meter range. Within 0.2 seconds, the object is detected, and the algorithm assesses the situation and sends instructions via User Datagram Protocol (UDP) packets wirelessly to the GVS device. The device receives the packets and steers the subject (human) autonomously. In total, four tests with different scenarios have been conducted. Through experimentation, it was observed that the device can recognize common objects and safely steer the user out of a collision >60% of the time without the need of the user's control within the ranges of 0 – 3 meters. The system could successfully detect, process, and transmit geospatial information to the GVS module; and steers the user into the correct trajectory to avoid any hazards obstructing the user's path.

Introduction

Today, an estimated 1.5 million Canadians[1] (approximately 3.95% of Canada's total population) have trouble seeing with corrective lenses, and many of the visually impaired use white canes. The traditional white cane can be easily thought of as the effective primary device that enables the visually impaired to stay mobile and independent. However, there are some important limitations of the white cane that can leave the user prone to injury. Incoming objects moving at high speeds are impossible to detect due to the white cane design nature. Additionally, they are limited to a short range, which gives users little time to react to their surroundings and increases the risk of colliding with other objects. To steer clear of an obstacle, users will not only know where the object is located but also approximate the safest direction to move toward. Several devices provided limited information to users in the form of echolocation, where object detection was done by a single range-finding sensor in a simple manner. Conversely, other designs have had an inefficient signaling method, using vibration as the only form of communication for the user. These modified navigation aids, along with the white cane, also require the users to provide direct physical interaction with the ground, introducing a multitude of hazards and techniques the users must be competent with.

An important feature of a visual impairment aid device is the ability to react and transmit the information fast and accurately to the user so that the person will be able to react and steer clear of the obstacle.

Presently, there are many electronic devices available to aid the visually impaired and prevent injuries. The UltraCane[2], introduced in 2010, is a white cane equipped with two ultrasound beams that can detect incoming obstacles and vibrate when an object comes within proximity. However, there is a major flaw in this design. If multiple obstructions were to be presented in front of the user, Ultra Cane will only be capable of vibrating to notify the user that there are incoming objects. The user will not be able to determine the location nor the number of objects obstructing the user to prevent colliding safely.

In a study[3] conducted by students from the University of California, 300 legally and completely blind people who are considered expert users of white canes were asked about injuries related to colliding with other objects. 98% of those interviewed indicated they have experienced one or more collision injuries. Furthermore, 23% of those collisions required medical attention.

Our proposed design attempts to circumvent these challenges from the approaches. The purpose of this device is to create an innovative prototype that can provide semi-autonomous collision avoidance for the visually impaired. The next section will present the methods used, after which the results will be highlighted and point out some areas of future work and lastly the conclusion.

Methods

Galvanic Vestibular Stimulation

Galvanic Vestibular Stimulation[4] is a method to stimulate the vestibular system. Sets of tiny structures behind the ears (vestibular nuclei, sensors...) just behind the ear is what keeps a person upright and steady. There are three fluid-filled canals, known as the circular canals, that sense the ratio of the head, while the other structure, the otoliths, sense the direction of gravity. Signals from the vestibular nerve travel to the brain, causing the sensation of being “steered”.

GVS Device

The main components of the GVS device consist of a: PCB, a 3.5mm Jack to the Terminal, an MT3608 booster module, and an ESP8266[5] module.

When the device is turned on, the ESP8266 creates a UDP server, or in other words, a Wi-Fi connection. Connecting a device to the UDP server creates UDP communication between the ESP8266 and the client (device). The ESP8266 constantly monitors its IP and port for new UDP packet instructions from the client. This is how it receives commands. Once a packet instruction is received, the ESP8266 sensor filters the packet and compares it with its own list of commands. It will execute the command that corresponds to that filtered packet. Using GPIO0 and GPIO2 (a digital signal pin) from the ESP8266 Chip, it then feeds the input from the client to an H-bridge circuit. In simple terms, the H-bridge simply controls the direction of the flow in the current. The H-bridge circuit then drives the output electrodes to a series of transistors and resistors, limiting the maximum current to 3 mA. This means any voltage input from 12V – 18V will still always provide a consistent current across the electrode pads.

Below is an illustrated diagram of the GVS device. There are in total three distinct sections: the power supply, current limiter, and H-Bridge Circuit. See the block diagram below.

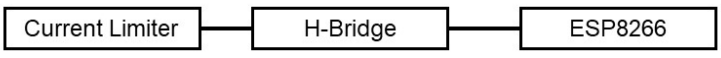

Figure 1. Illustrated Diagram of GVS Device

The power supply consists of two AA batteries, combining a total voltage of 3.3V. The power supply is directly connected to the 3.3V J1 Male pin. The J1 3.3V pin utilizes the 3.3V to power the ESP8266. After supplying 3.3V of voltage to power the ESP8266 module, an external MT3608 Booster Module is then utilized to multiply the voltage to roughly 12 volts. This provides power to the circuit and supplies current to the two electrode pads.

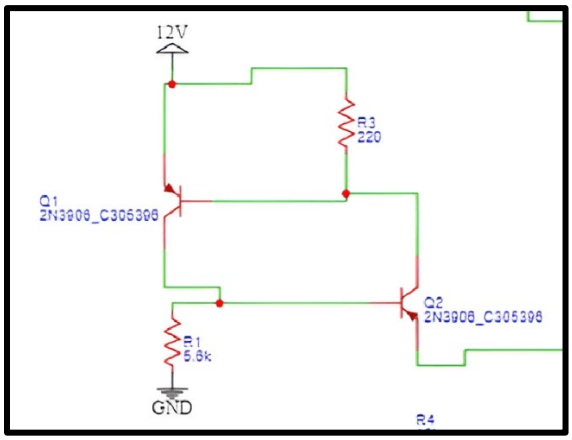

Figure 2. Diagram of the Current Limiter

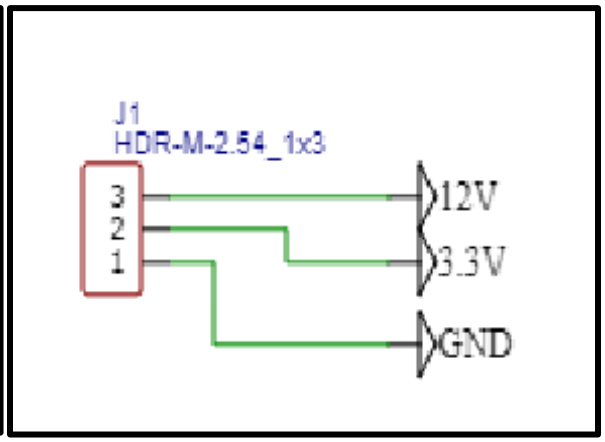

Figure 3. Diagram of the Power Supply

To impart an extra layer of current security, a current limiter circuit section was implemented. After being multiplied, the voltage passes through section 1 (Figure 2). The current limiter consists of two PNP transistors and one 220OHM resistor. This creates a feedback loop. When enough current flows through the 220 OHM resistor, such that the voltage drop located in the intersection becomes 0.65 volts, transistor Q1 begins to initiate. By turning on the node located near the bottom left corner, transistor Q2 begins to turn off. Therefore, the voltage drop across this resistor will never exceed approximately 0.65V.

Utilizing Ohm's Law[6] (Figure 4), it was found that 0.003 amps or 3 mA are never going to be passed through into the H-bridge circuit.

$$\frac{V}{R} = I \qquad \frac{0.65\,V}{220\,Ohm} = 0.003\,Amps$$

Figure 4. Ohm's Law

Section 2 (Figure 5) consists of an H-bridge, which will provide the circuit with the ability to switch the polarity of the 3 mA of current that is being passed through. Here, 4x 2N3906 transistors were used so that there is as little voltage drop as possible. The 4x resistors are then connected to NPN resistors, which are in turn connected to the ESP8266 module's GPIO 2 (A) and GPIO 0 (B).

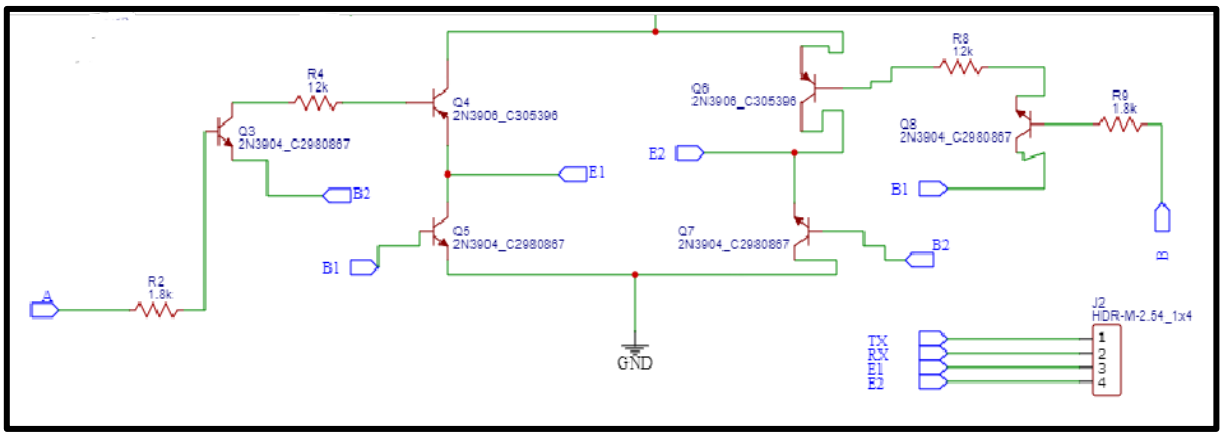

Figure 5. Diagram of the H-Bridge Circuit

When "A" and "B" are set to low, transistors Q3 and Q9 are turned off. The N-Channel transistors remain on, but the P-Channel transistors remain off, ceasing all current flowing through the load. When a UDP packet is received instructions from the device to turn left, "A" is set to high, and "B" is set to low. When "A" becomes high, transistor Q3 turns on and brings the node down to the ground. It then proceeds to turn transistor Q4 off and turn transistor Q5 on. The same thing happens on the other side.

The Intel RealSense D435 Depth Camera

Cameras that provide distance measurement along with RGB data have increasingly been appearing in the market as alternatives to the more expensive setup of LIDARs and webcams. While products such as the Kinect have existed in the past, their weight and form factors have been demanding constraints for projects that involve mobility. The Intel RealSense[7] D435 camera is a 3D depth-sensing camera currently available on the market. Its small form factor is used in many applications, such as robotics, machine learning, depth estimation, and automation. The depth camera consists of four highly compacted sensors: Right Imager, IR projector, Left Imager, and RGB Module. The combination of a stereo camera and an RGB module allows it to simultaneously provide RGB data and accurate depth values per pixel while in motion; bringing 3D to devices and machines that only see 2D today. The device developed in this research report utilizes the Realsense camera to perform navigation, specifically its collision avoidance pathing algorithm.

Stereo Vision

Stereo Vision[8] is the extraction of 3D information from digital 2D images. By comparing information about a scene from two different vantage points, 3D information can be collected by examining the relative positions of objects in the two panels. This works similarly to the two eyes humans have. Human binocular vision perceives depth by the difference in image location of an object seen by the left and right eyes. Objects closer seem to move significantly from eye to eye, whereas an object in the far distance would appear to move very little. The brain uses this binocular disparity to extract depth information from 2D images. The Intel RealSense camera simulates this technique of human vision to perceive depth.

The Intel RealSense D435 sensor uses an approach called Active Stereo. Active stereo vision is a form of stereo vision that actively emits a light such as a laser to simplify the stereo matching problem.

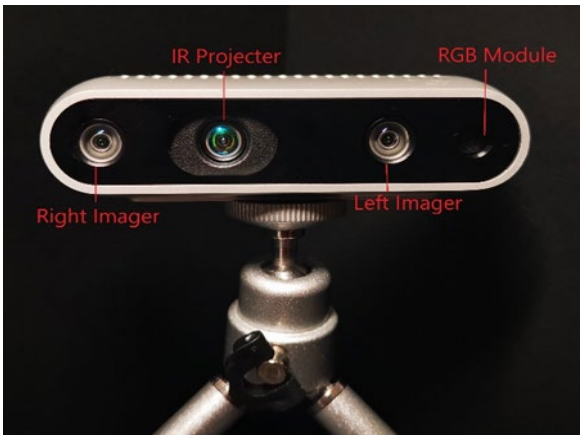

Figure 6. Sensors in the Intel Realsense

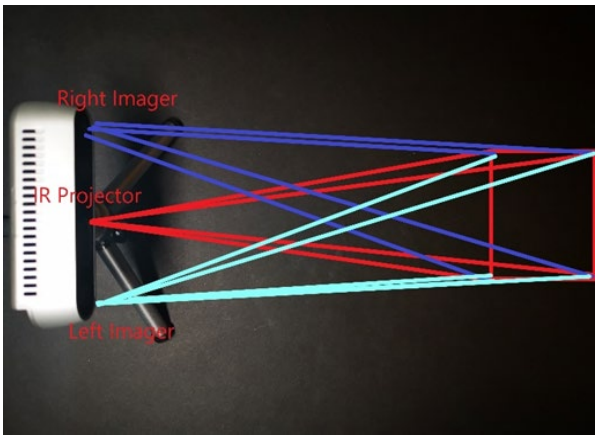

Figure 7. Active Stereo

The addition of an infrared projector allows the camera to perform exceptionally well in low light, allowing it to function normally under feeble light conditions. This also offers a significant advantage over traditional passive stereo systems, which are practically unusable under lower light conditions.

Stereo Triangulation

Stereo Triangulation is the principle underlying stereo vision. It refers to the process of determining a point in 3D space by finding the intersection of two lines passing through the center of the projection and the projection of the point of each image. The Intel RealSense D435 camera uses this method to quantify depth. The graph illustrated on the below represents the basic stereo camera setup. It is assumed that the D435 camera uses a similar setup to Figure 21. A 3D point is captured where ${}^{c1}\boldsymbol{p}$ and ${}^{c2}\boldsymbol{p}$ intersect. As ${}^{n}\boldsymbol{x}$ is the projection of the target, if a ray was projected from camera 1, it would be able to pass through the target. Meaning that any light rays coming from the target are going to travel down this ray and intersect the image plane and into the optical center. This will happen for any points in our scene that this camera can see. Simultaneously, rays travel and intersect the same points on camera 2. To calculate the depth, simply determine whether the light rays intersect.

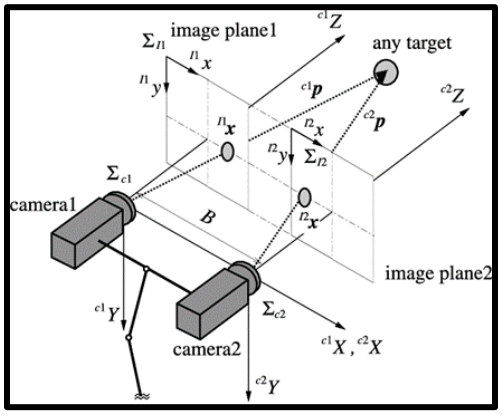

Figure 8. Model of a Parallel Stereo Camera[9]

Epipolar Geometry

The main difficulty is to find the exact intersection point from camera2 to camera1. The search base is quite large, and reliably finding the exact point like this in a different image when it might have rotated and changed

slightly is a lot of work in two dimensions. Every single pixel in both cameras will be needed to compare and attempted to be matched. To increase efficiency and solve the problem, Epipolar Geometry is utilized.

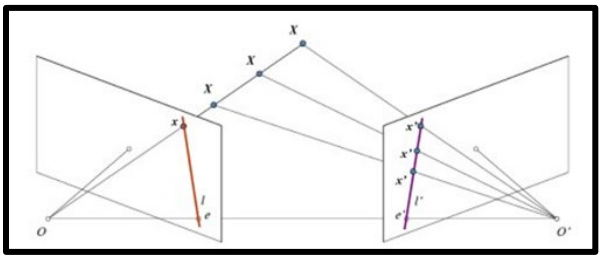

Figure 9. Model of Epipolar Geometry[10]

Suppose that this is a triangle. From ***o'***, rays can be passed through from the optical center and intersect with the ray ***o.*** This creates a line called the epipolar line (drawn in purple). The points along the line are all the possible projections of the ***o'*** ray into the ***o*** image, significantly reducing the list of pixels required to be examined.

YOLOv5 (You Only Look Once)

YOLOv5[11] is a CNN-based high-speed object detector, known for taking little computational resources and achieving high accuracies. YOLO is an object detection model, where it's trained to look at an image and search for a set of object classes. When the specific class is found, a bounding box is illustrated, and their class is identified. YOLOv5 prioritizes real-time object detection, making it very suitable for this project. The dataset used in this project is from COCO[12], which contains a broad range of 100+ common objects.

Collision Avoidance Pathing Prediction Python Algorithm

The collision avoidance pathing prediction python algorithm was then constructed to be used in conjunction with the YOLOv5 object detection model. The algorithm utilizes Intel's open-source RealSense library as well as several other Python modules, including pyttsx3[13] (text-to-speech function), socket[14] (UDP packet sender/GVS information transmitter), NumPy[15] (computing tool), etc. Below is a flow chart illustrating how the algorithm makes its path prediction decisions.

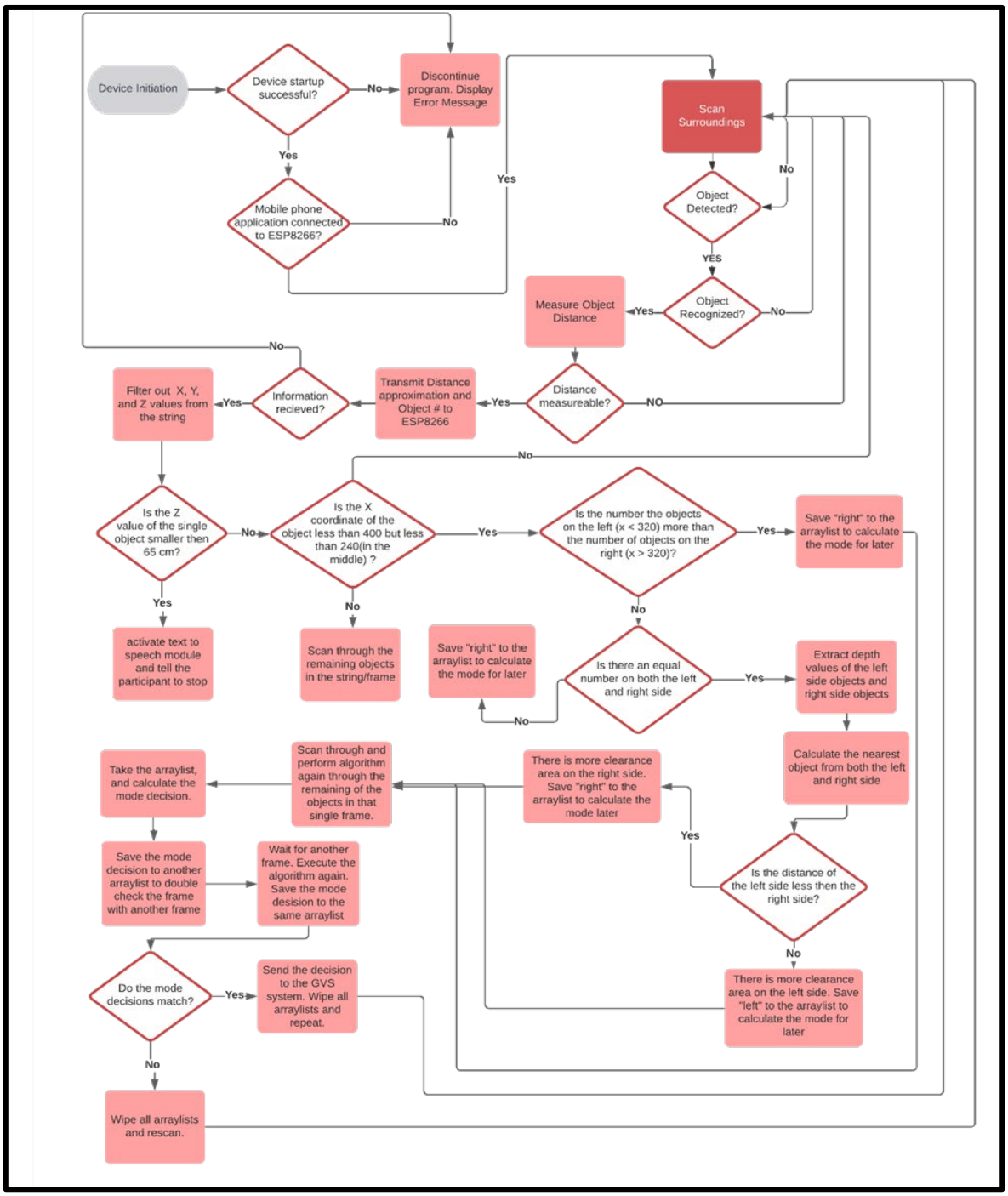

Figure 10. Collision Avoidance Pathing Algorithm

Discussion

Part 1: Experimenting with the GVS System

Part 1 of the experimentation consisted of finding the relationship between the voltage outputted and the head angular change. A total of 10 trials, each trial with a duration of 5 seconds, was conducted. Utilizing a simple

"for-loop" control flow statement, directions were sent from a computer to the GVS device through UDP packets.

The voltage output from the GVS device was measured by an external multimeter, and the roll degrees were calculated by analyzing recorded camera footage. In this experiment, the participant was standing upright and still in a room-temperature environment. Figure 11 below shows the relationship between the voltage output from the GVS and the roll degree. According to Figure 11, the voltage output continues to rise and reaches around 11 volts, the change in the roll degree begins to plateau or flatten. This suggests that the vestibular system's response becomes less sensitive to further increases in voltage. At this point, additional increments in voltage might not lead to substantial changes in the roll degree, indicating a saturation effect. The saturation effect could be due to various factors, such as physiological limits of the vestibular system's response, neural adaptation to continuous stimulation, or limitations of the GVS device itself.

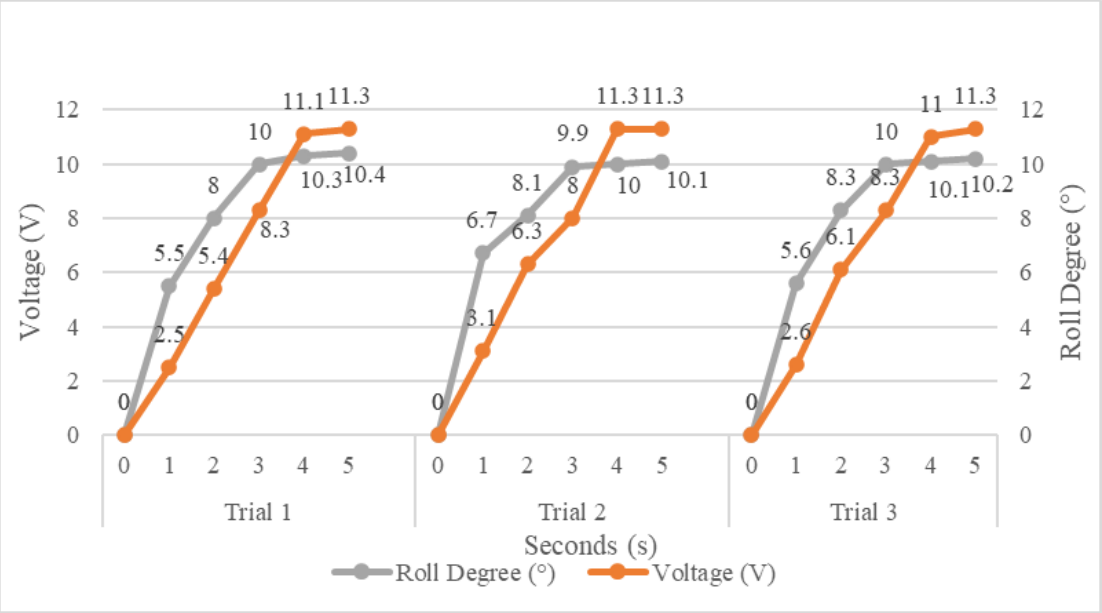

Figure 11. Effect of Voltage Output on Roll Degree

Part 2: Trialing with Intel RealSense D435 Depth Camera, YOLOv5 Detection Algorithm, and Collision Avoidance Prediction Pathing

Part 2 of the experiments consisted of determining the accuracy of the pathing decision algorithm. A total of 6 objects were utilized to conduct the experiment. Each object was split into 3 separate categories: large, medium, and small. The controlled environment consisted of a wall on the left side, which obstructed the user's path; the objects were placed in the middle, leaving the only possible clear path to the right. The accuracy was calculated based on the number of successful decisions the algorithm made; in this case, it should have outputted "right". Each object was trialed 10 times from 1 meter, 2 meters, and 3 meters away. In total, 180 trials were conducted. It was found that large and distinctly colored objects were more easily identified, therefore increasing the accuracy rate of the algorithm.

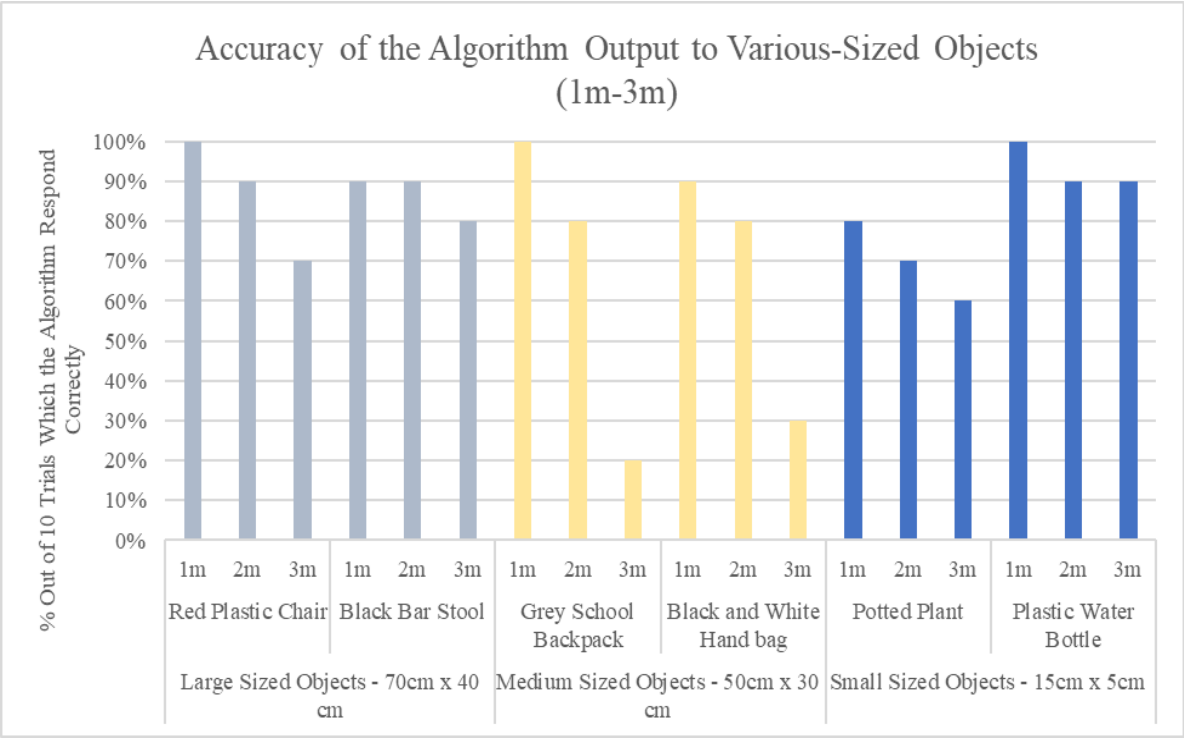

Figure 12. Accuracy of the Algorithm Output to Various-Sized Objects (1m-3m)

Part 3: Merging the GVS System with the Camera, Object Detection, and Algorithm

The final part of the project consisted of experimenting with the system when both devices were merged. In total 4 trials were conducted, each with a different scenario. The experiments took place in a large open space measuring 230 cm in width and 345 cm in length. This space was cut into a 4x6 grid, each approximately 60 cm by 60 cm. Each trial was recorded using a smartphone and was analyzed and mapped out after the trials were finished (refer to Appendix A for recording footage). Prior to the scenario setups, the participant, Justin Peng, was blindfolded. This was to minimize bias in data collection. The adult supervisor afterward created a randomized scenario and arranged the objects.

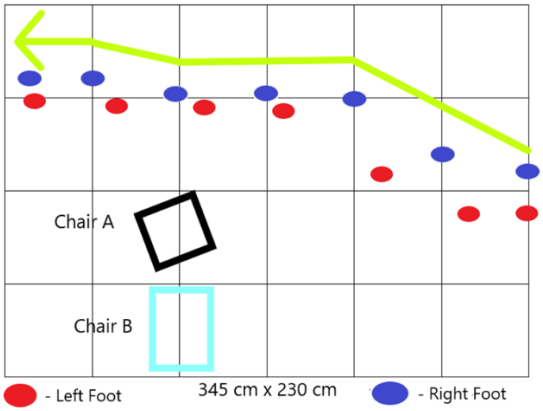

Figure 13. The device performed as expected. Chair A & B were detected, and the decision made was to steer right.

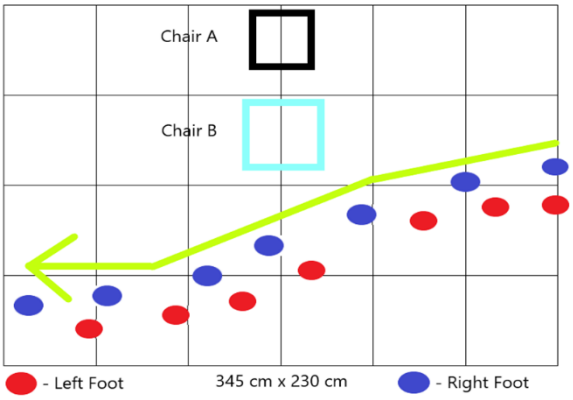

Figure 14. Chair A & B were moved and rearranged to the right side. The device successfully detected two objects and steered the participant to the

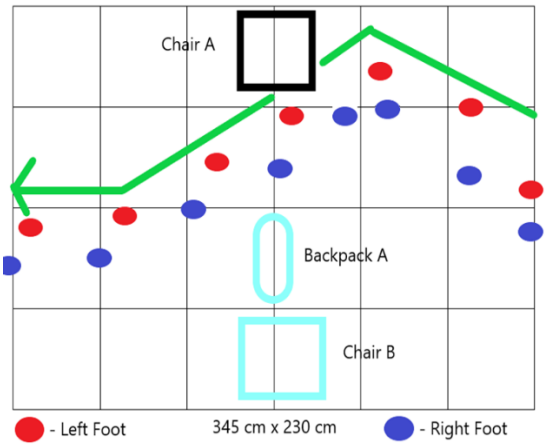

Figure 15. Chair A & B were moved and rearranged to the right side. The device successfully detected two objects and steered the participant to the left.

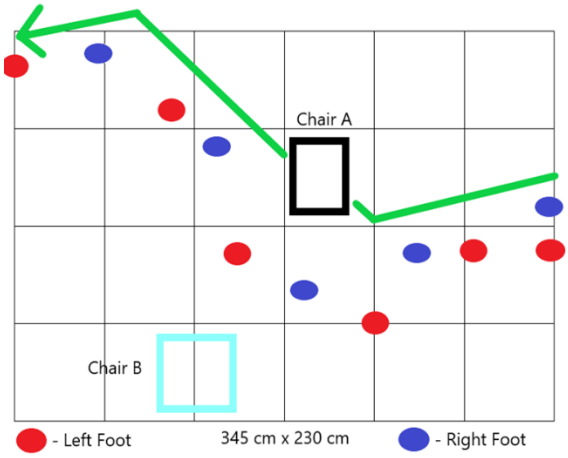

Figure 16. Three objects: Chair A & B and Backpack A, were picked up by the camera detected. For an unknown reason, the device decided to steer the participant to the right. But this decision was changed afterwards by immediately turning the user to the left, therefore successfully avoiding the objects.

Conclusion

Based on the results, it can be observed that the device can recognize common objects and safely steer the user out of a collision >60% of the time without the need of the user's control within the ranges of 0 – 3 meters. The purpose, which was to develop a novel approach to a semi-autonomous collision avoidance device for the visually impaired and allow it to be used in most real-life situations where a white cane is proven to be hazardous and limited, was achieved. Based on our results, it is hopeful that the implementation of a four-pole GVS device can provide fully autonomous control to the human's walking directions. This includes left roll, right roll, front pitch, back pitch, left yaw, and right yaw. Overall, this project was proven to be successful and can provide an alternative innovative solution to traditional visual impairment aid devices utilizing GVS and object detection technologies, which is implemented on custom hardware.

Limitations

Current Decrease for GVS Output

Like the voltage decrease for the ESP8266 module, the maximum current output to the two electrode pads steadily decreased as the number of trials conducted and usage increased. Although the output voltages were examined prior to every experiment, the exact voltages were not monitored during the experiment. This may have led to a major difference in left and right anodal output and roll angle.

YOLOv5 Detection

Some objects that were not included in the pre-trained dataset, was not capable of being detected. This resulted in several misleading path predictions. But because of the double-frame-checking method implemented in the algorithm, these mistakes were mostly avoided.

Lack of CUDA Computational Cores

The lack of CUDA[16] cores majorly affected the performance and results of the experiment. According to OpenCV[17]'s official website, OpenCV's GPU module runs up to 30x faster than the CPU module. The GPU module was written using CUDA, therefore OpenCV's "DNN" module only supported NVIDIA GPUs. The computer that was utilized to conduct contained both an AMD[18] CPU and an AMD GPU, therefore only the CPU module was employed. The speeds of both the algorithm and detection were extremely slow, sometimes even skipping whole frames. The average speed of the algorithm was approximately 0.25 seconds per frame from detecting an object to calculating its decision.

Due to the algorithm's slow speeds, the error in motion heavily affected the results. If a frame was captured and calculated during a high activity movement (e.g., spinning around in a circle very fast), it would have been impossible to calculate the best path to avoid a collision. Part of the python algorithm consisted of a double-checking method. The double-checking method compared results from two frames and identified if both results were the same. Because of the long amount of time in between each of the frames, the double-checking method sometime malfunctioned, resulting in no output.

Voltage Decrease for the ESP8266 Module

Through visual observation, it was found that the flashing blue LED location on the ESP8266 module appeared dimmer every time the device was turned on. The ESP8266 module required a 3.3V power level to properly supply the module for communication. A battery case containing two AA batteries was able to output exactly 3.3V assuming the batteries were brand new. As the device was being used more often, the voltage gradually decreased, dropping under the required voltage for the module. In addition, the packet collection speed visually appeared to be twice as fast as batteries that have been used for a long time. This may have affected the results, as the reported time frames may be slightly delayed.

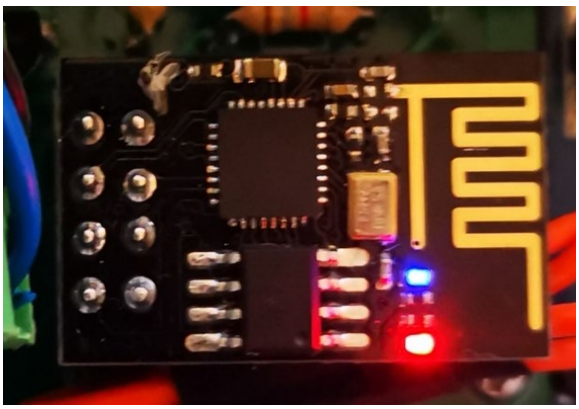

Figure 17. ESP8266 LED with an old battery

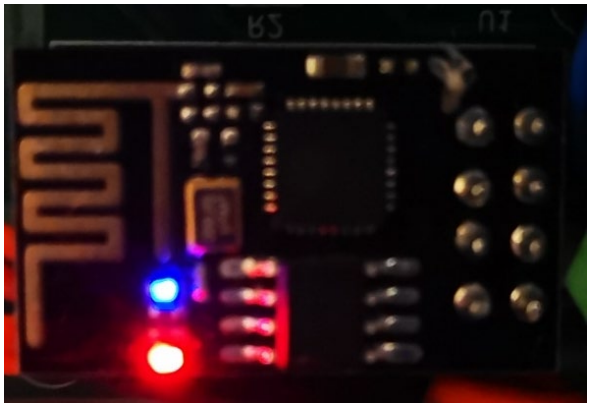

Figure 18. ESP8266 LED with a new battery

Acknowledgments

I would like to express my deep gratitude to my mentor, Mert Kaval, for his feedback and guidance throughout the project. I would also like to thank Ms. Leong for helping me proofread most of the parts related to science fair. Finally, I wish to thank my parents for their support and encouragement throughout my studies. I would also like to thank Elizabeth Strehl and Peter Mbua for their diligent editing of the paper. Their keen attention to detail and editorial finesse have undoubtedly elevated the quality of this work.

References

[1]*Blindness in Canada*. CNIB. (n.d.). Retrieved January 18, 2022, from https://cnib.ca/en/sight-loss-info/blindness/blindness-canada?region=bc#:~:text=Today%2C%20an%20estimated%201.5%20Million,that%20could%20cause%20sight%20loss.

[2]*Ultracane - a safer way to travel....* UltraCane .... putting the world at your fingertips. (n.d.). Retrieved January 18, 2022, from https://www.ultracane.com/

[3]*Mobility accidents - users.soe.ucsc.edu*. (n.d.). Retrieved January 18, 2022, from https://users.soe.ucsc.edu/~manduchi/papers/MobilityAccidents.pdf

[4]Dlugaiczyk, J., German Center for Vertigo and Balance Disorders, Gensberger, K. D., II, D. B., Straka, H., C, A., DE, A., ST, A., RA, B., J, B., S, B., M, B., S, B., EA, B., JH, B., J, B., TC, B., SF, B., E, B., … Mikhail, Y. (2019, June 10). *Galvanic vestibular stimulation: From basic concepts to clinical applications*.

Journal of Neurophysiology. Retrieved January 24, 2022, from https://journals.physiology.org/doi/full/10.1152/jn.00035.2019
[5]*ESP8266*. ESP8266 Wi-Fi MCU I Espressif Systems. (n.d.). https://www.espressif.com/en/products/socs/esp8266
[6]*Ohm's law - how voltage, current, and resistance relate: Ohm's law: Electronics textbook*. All About Circuits. (n.d.). https://www.allaboutcircuits.com/textbook/direct-current/chpt-2/voltage-current-resistance-relate/
[7]*Python*. Intel® RealSense™ Developer Documentation. (n.d.). Retrieved January 24, 2022, from https://dev.intelrealsense.com/docs/python2
[8]Read, J. C. A. (2014, December 5). *Stereo vision and strabismus*. Nature News. https://www.nature.com/articles/eye2014279
[9]Model of Parallel Stereo Camera | Download Scientific Diagram. (n.d.). https://www.researchgate.net/figure/Model-of-parallel-stereo-camera_fig5_3832506
[10]*Epipolar geometry*. OpenCV. (n.d.). https://docs.opencv.org/3.4/da/de9/tutorial_py_epipolar_geometry.html
[11]Xu, S., Guo, Z., Liu, Y., Fan, J., & Liu, X. (2022, September 11). *An improved lightweight YOLOV5 model based on attention mechanism for face mask detection*. arXiv.org. https://arxiv.org/abs/2203.16506
[12]Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015, February 21). *Microsoft Coco: Common Objects in Context*. arXiv.org. https://arxiv.org/abs/1405.0312
[13]*Text-to-speech x-platform¶*. pyttsx3. (n.d.). https://pyttsx3.readthedocs.io/en/latest/
[14]*Socket - low-level networking interface*. Python documentation. (n.d.). https://docs.python.org/3/library/socket.html
[15]*Numpy user guide#*. NumPy user guide - NumPy v1.25 Manual. (n.d.). https://numpy.org/doc/stable/user/index.html#user
[16]*Cuda Zone - Library of resources*. NVIDIA Developer. (2023, June 27). https://developer.nvidia.com/cuda-zone#:~:text=CUDA%C2%AE%20is%20a%20parallel,harnessing%20the%20power%20of%20GPUs.
[17]*Home*. OpenCV. (2023, August 17). https://opencv.org/
[18]AMD | together we advance AI. (n.d.-a). https://www.amd.com/en.html
Dougherty, B. E., Kehler, K. B., Jamara, R., Patterson, N., Valenti, D., & Vera-Diaz, F. A. (2011, November). *Abandonment of low-vision devices in an outpatient population*. Optometry and vision science: official publication of the American Academy of Optometry. Retrieved January 18, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3204005/
Aoyama, K., Iizuka, H., Ando, H., & Maeda, T. (2015, May 11). *Four-pole galvanic vestibular stimulation causes body sway about three axes*. Nature News. Retrieved January 24, 2022, from https://www.nature.com/articles/srep10168
Dlugaiczyk, J., German Center for Vertigo and Balance Disorders, Gensberger, K. D., II, D. B., Straka, H., C, A., DE, A., ST, A., RA, B., J, B., S, B., M, B., S, B., EA, B., JH, B., J, B., TC, B., SF, B., E, B., … Mikhail, Y. (2019, June 10). *Galvanic vestibular stimulation: From basic concepts to clinical applications*. Journal of Neurophysiology. Retrieved January 24, 2022, from https://journals.physiology.org/doi/full/10.1152/jn.00035.2019
Python. Intel® RealSense™ Developer Documentation. (n.d.). Retrieved January 24, 2022, from https://dev.intelrealsense.com/docs/python2
RC; W. D. L. F. (n.d.). *What does galvanic vestibular stimulation stimulate?* Advances in experimental medicine and biology. Retrieved January 24, 2022, from https://pubmed.ncbi.nlm.nih.gov/12171101/
Utz KS;Korluss K;Schmidt L;Rosenthal A;Oppenländer K;Keller I;Kerkhoff G; (n.d.). *Minor adverse effects of galvanic vestibular stimulation in persons with stroke and healthy individuals*. Brain injury. Retrieved January 24, 2022, from

https://pubmed.ncbi.nlm.nih.gov/21879800/#:~:text=Results%3A%20The%20most%20frequent%20symptoms,10.7%25)%20underneath%20the%20electrodes.