Assistive Technology for the Visually Impaired to Navigate Cluttered Environments

Devesh Nath, Electrical Engineering
Mentor: Pavan Turaga
School of Arts Media and Engineering; School of Electrical, Computer and Energy Engineering

INTRODUCTION

There are many products and technologies that have been developed to assist the blind and the visually impaired for everyday tasks by using methods to deliver information to the user through senses other than vision. Such devices developed in the past assist the visually impaired through auditory, tactile, and haptic feedback yet, many have failed to make a big difference in the world of the visually impaired because of drawbacks like high cost, low environmental resistance, slim usability scenarios and portability. This research. This research explores a combination of various inexpensive existing technologies driven by advances in open-source machine-learning tools for visual analysis. Our goal is to develop a scaled down model using machine learning on edge devices like raspberry pi and the Arduino microcontroller.[1,2,3]

RESEARCH QUESTION

What are the drawbacks of current commercially available (and other proposed technologies) to assist the visually impaired and how can these technologies be made better?

RESEARCH FOCUS

● CURRENT SOLUTIONS ● DISTANCE ESTIMATION ● TEXT RECOGNITION
● TEXT-TO-SPEECH ● FEEDBACK ● INTEGRATION AND FUTURE WORK

CURRENT SOLUTIONS

- **OrCam Devices** [4,5] **:** Very Portable, suited to read written text, only Text-to-Speech functionality, very pricy($1500)

[13]

- **WeWalk Smart Cane** [7,8] **:** Portable, only obstacle avoidance functionality through haptic and auditory feedback, very prices($500)

[14]

- **eSight:** [9,8]**:** Provides an AR (augmented reality) expe -reince, several functions (magnification, contrast and brightness adjust), very pricey($5500)

[15]

TEXT RECOGNITION[10]

Text recognition is a very important part of this research as it can help the visually impaired understand text through an auditory input like headphones. 2 modules have been found to best suit this use-case, Google's Tesseract and EasyOCR. In general, results on a GPU have been faster as compared to CPU because GPUs can perform multiple simultaneous computations unlike the CPU. GPU implementations can be complex because of the various compatibility issues and CPU implementations are relatively simpler. After testing many sample images, Google's Tesseract has been more consistent with results and does not break as often as EasyOCR does. Both modules were found to be more comfortable towards .png type images in general. Both modules fail when they received relatively complex images with different fonts, text orientations, emojis etc. The modules work well, generally, only in ideal conditions. Gray scaling, thresholding and median-blurring are required for noise reduction.

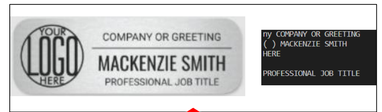
Processing Time	CPU	GPU
Tesseract	0.3 seconds/image	0.25 seconds/image
EasyOCR	0.82 seconds/image	0.07 seconds/image

[10]

Error Rate	Error Rate on Numbers	Error Rate on Alphabets
Tesseract	5.50%	0.7%
EasyOCR	1.90%	4.3%

[10]

Examples (Images to the Left, Results to the Right):



TEXT-TO-SPEECH

The pyttsx3 module is an easy module to implement in code. It is a text to speech conversion library that can work offline and can also run with both Python 2 and Python 3. Another option is the gTTS library or Google Text To Speech. This library is also very easy to implement however it cannot run offline and requires to save the results to an MP3 file instead of directly playing it. Playing a MP3 file repeatedly through code can add some extra unnecessary lines to the code and can be slow due to repeated opening and deletion of a separate file. Although gTTS has a more natural voice, pyttsx3 is better suited for this implementation. [11]

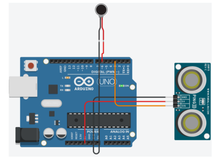
DISTANCE ESTIMATION[12]

Distances can be estimated to an accuracy of 5mm for a range of 30cm to 12m using a TF Mini Lidar Module with an Arduino. The result from this sensor can be very reliable due to very high accuracy and low-light functionality. It is relatively expensive compared to other sensors, however its pros outweigh the cons. It has been used to generate a 2-D point cloud of its surroundings for a visual look. This technology can be expensive for better versions and does not wok well in extreme weather.

FEEDBACK

Feedback can be provided to the user using haptics and regular headphones. Wired 3.5 mm audio jack connections ports on edge computing devices like the Raspberry Pi can be used to send audio output directly to the user. Vibration motors can be used to send haptic data to the user. Such motors can be directly powered by a PWM pin on an Arduino and have the capability to modulate vibration intensity depending on any input parameter.

A simple Arduino hookup to a Distance sensor →



INTEGRATION AND FUTURE WORK

All these technologies can be integrated using powerful boards like the Nvidia Jetson or the Raspberry Pi. The Raspberry Pi was found to be suitable for this research because of great I/O connectivity, community support, portability and efficiency. The raspberry pi can be programmed in python 3 to perform text recognition, convert text to audio feedback, distance estimation using LiDAR and GPIO pins, and haptic feedback through vibration motors. This all-in-one device can then be battery powered through a USB-C port and be attached as a portable, battery powered installation to a glove.

Implementation of the mentioned software and hardware on the Raspberry Pi can be relatively cheaper and a more all rounded solution than currently available technology for the visually impaired. More capable edge devices like the Jetson Nano can also be used instead of the Pi however, they can also be hard to find during the current chip-shortage.

Further research in computer vision can bring more efficient and more accurate learning models suited even for non-ideal situations.

REFERENCES

[1] Li-Chi Huang, Kuldeep Kulkarni, Anik Jha, Suhas Lohit, Suren Jayasuriya, Pavan K. Turaga: *CS-VQA: Visual Question Answering with Compressively Sensed Images*. ICIP 2018: 1283-1287 **[2]** Suhas Lohit, Ankan Bansal, Nitesh Shroff, Jaishanker K. Pillai, Pavan K. Turaga, Rama Chellappa: *Predicting Dynamical Evolution of Human Activities From a Single Image*. CVPR Workshops 2018: 383-392 **[3]** Divya Mohan, Sarmeeksha Katoch, Suren Jayasuriya, Pavan K. Turaga, Andreas Spanias: *Adaptive Video Subsampling For Energy-Efficient Object Detection*. ACSSC 2019: 103-107 **[4]** Waisbord, Michael, et al. "The Impact of a Novel Artificial Vision Device (Orcam) on the Quality of Life of Patients with End-Stage Glaucoma." *Investigative Ophthalmology & Visual Science*, The Association for Research in Vision and Ophthalmology, 11 June 2015, https://iovs.arvojournals.org/article.aspx?articleid=2335212. **[5]** AdaptVision. "5 Things to Consider before Buying the Orcam Read Reading Pen." *AdaptiVision*, 7 Oct. 2021, https://lowvisionsource.com/spotlight-orcam-read/. **[7]** Matchar, Emily. *Smithsonian.com*, Smithsonian Institution, 17 Oct. 2019, https://www.smithsonianmag.com/innovation/smart-cane-helps-blind-people-navigate-180973279/. **[8]** Stepko, Barbara. *Fantastic New Tech for People With Low Vision or Blindness*, 18 Feb. 2021, https://www.aarp.org/health/conditions-treatments/info-2021/low-vision-technology.html. **[9]** Wittich, Walter, et al. "The Effect of a Head-Mounted Low Vision Device on Visual Function." *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, Lippincott Williams & Wilkins, Sept. 2018,https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6133226/. **[10]** Kanjee, Ritesh, Augmented Startups, Youtube, https://www.youtube.com/watch?v=j_gl45Vrz8&ab_channel=AugmentedStartups} **[11]** The Python Code, https://www.thepythoncode.com/article/convert-text-to-speech-in-python **[12]** The Drone Bot Workshop, https://www.youtube.com/watch?v=VtbFbxvO1k **[13]** OrCam, https://www.orcam.com/en/en/product-comparison/ **[14]** Time, https://time.com/collection/best-inventions-2019/5733049/wewalk/ **[15]** eSight 4, https://esighteyewear.com/low-vision-device-for-visually-impaired/

FURI

Ira A. Fulton Schools of Engineering
Arizona State University