On the Use of Computer Vision to Assist Visually Impaired People Achieve Social Distancing

Lucas G. Nadolskis
University of Minnesota

Vinod Namboodiri
Wichita State University.

Final project report of the Distributed Research experiences for Undergraduates (DREU, summer 2020)

Introduction

This year the world saw the rise of the new corona virus known as covid19. Bringing much distress to the human society in general, the virus redefined what we knew as normal. Some of the changes will be temporary but others will define a new normal which we will need to learn how to deal with in the upcoming years. Some of the most common guidance from medical and political authorities are the use of face covers and the practice of social distance. After studying the way that the covid19 virus propagates it was understood that if people keep a distance of around 2 meters from another the chances of virus transmission reduces drastically. [21]. Therefore, many governments in different spheres as well as private business have developed mandatory use of face covers and the practice social distance. As confusing and inconvenient these measures may seem to a regular member of the society, the implications are much bigger for those of us who have some medical condition or disability. In special visual impairments that reduces the capacity of one to determine the distance to another person, or in the case of blindness makes it totally impossible. Visually impaired people are able to have a life very close to normal in a non-pandemic society, but with the current tools that are most widely used to aid visual impaired people with navigation social distance is almost impossible to be practiced.

The issue of independent navigation for blind and low vision people is one that we as a society are trying to address for years. As it is mostly commonly known blind and low vision people mainly uses two different ways of walking around, those are traditionally a white cane and a Gide dog. The classic tapping of the cane has many different uses on the navigation for blind people, some of them can be understanding the hardness of objects when tapping [8]. Also, we can see in [14] that the use of the white cane makes the independent navigation much more efficient and safer for blind and low vision people. On the question of the guide dogs, the many benefits of its uses can be found in [12], here an extensive work presenting how using a guide dog brings more independency and safety for blind and low vision people. The efficiency of the dog and the cane are clear in the works presented above, however, with the technological developments more advanced devices have been tempted, more information about those can be found in the next section. However, we can see a common issue with all of them that are specially more relevant in the current scenario. Any of them provides the user with an accurate information about distance to the obstacle. The devices are mainly focused on avoiding the obstacles, and for that one may argue that providing the precise distance until collision is quite unnecessary. However, with the social distance measures the information about distance becomes much more relevant for the user. Not only to not be hurt but to avoid going against rules and also complying with the same norms as other member of society. Here it is important to remember that accessibility goes in both ways, as people with disabilities wish to become regular members of society, being subject to the same rules comes with that. This is the reason for the development of Blind distance.

On Blind Distance we aimed to find the fine equilibrium between giving enough information for the user and not giving too much information to the point that becomes another difficulty. This is one issue present in many devices that aim to help the blind and visually impaired on navigation, this fine line is very hard to find and may often ends in an unusable device if it is not correctly tuned. So, the decision on Blind Distance was to use a different form of information. Giving the distance with a voice would be more information than needed and vibrations may often be confused with the vibration of one's phone or watch. So, beeps were used. Blind Distance understands when the user is getting ever closer to another person and based on the social distance rules beeps to avoid crossing that line of 2 meters.

Related Work

In this review we will see some different ideas that were applied throughout time with the technology that was available to develop more advanced ways to aid on the navigation of blind and low vision people.

Here two main categories will be observed, infrastructure less technologies and infrastructure-based technologies. Infrastructure based (ISB) is where besides the navigational device itself, installations on the environment are necessary to guide and inform the device, on the case of infrastructure less (ISL) only the navigational device itself is required.
Let us explore more the ISB case. In one example in [5] a system using RFID is proposed. It is cost effective since the tags can be reused from animal identification. A RFID detector is inserted on the cane which is used to identify the tags fixed on the ground. In another example of ISB in [4] Proposes a method to implement IR and ultrasound systems in the room. Through cap locate those and using a path finding algorithm find the best rote. The interesting thing about [4] is the use of wearable technologies, those are going to be present later on ISLs too. In [13] a similar not different from [5] is proposed, but here the device is connected to the cane of the user. In all of the ISBs we could see a concern with portability and in some cases cost effectiveness, this is important to keep in mind as we move forward into new technologies.
Entering now on the ideas of ISL, as stated previously this is the idea of self-sufficient devices. The first example in [6] it is possible to understand the transaction from ISB to ISL considering that the idea of IR and sonars are still used but in a way that is independent of the installation it is an interesting paper to understand the differences between ISL and ISB. In [2], [6] and [16] a similar idea of using RIFDs is applied. However, not only radio signals and ultrasounds have

been tried on the task of independent navigation for blind and low vision users. In [7] an interesting idea of using a dynamic map that will refresh depending on the localization of the user is proposed. And in [1] instead of developing a device for the user to carry around an innovative idea of developing a guide for indoor navigation is proposed. Here a robot would guide the user to where he or she wants to go, an interesting system to be used in large areas such as malls and airports.

All of the systems studied above both in ISL and ISB have one common question, how does a blind person perceive a new environment? Either by RFIDs, robots that guide people or dynamic maps the question of how a blind person can explore and understand the environment of the place that he or she is walking to is the same. An excellent work aiming to answer this question is presented in [9]. Here investigators developed a virtual 3d environment where blind people could use a force stick to navigate around and understand what is happening. The position of the objects in the room or were the entrance and exit is. This shows how important the question of how we can make more accessible for a blind person to walk around in a new environment is.

One of the newest techniques trying to answer this question is the use of computer vision. Computer vision is the idea of processing multiple images and obtaining information's about that image such as which objects are present or where in the image something is. The uses of CV for accessibility can be found in papers such as [3]. This paper goes in details on different ways that researchers tried to create systems that used computer vision to aid the blind people navigation. Also work aiming to help the blind person to identify places and objects can be found in [15] where investigators used vgg16 as their database to identify places on the go. In object detections products such as OrCam or cybereyez are trying to achieve the same goals. However, one of the most common issues on using computer vision on accessibility is the difficulty that blind people have to position the camera correctly to detect the image. This was well explored in [10]. This is most definitely one of the issues we will try tackling in our own work.

This leads us to the final part of this review. In our own research we seek to answer one specific question. In all of the papers above we did not see one example where the authors answered the question of how distance an object is. Certainly, works on the proximity of the object were done, but in a much general way. We will focus on answering the specific question of determining the distance to an object. The implications of these are huge, in [14] for example we saw how the efficiency and speed of blind people work. In that case knowing that you are at a certain distance to an object could make all of the difference on improving these two factors. Also, the relevancy of this work in 2020 where social distance is not only recommended but required in many cases, it most definitely became a factor in the independent life of blind and low vision people. If we can tackle this question, exploring new places will be less stressful and safer not only for the user of this technology but for everybody else.

The Blind Distancing System

Since it is essential for the device to have a good calculation of depth, it was decided to use an intel real sense d 435 camera, which appears to be the best camera available to this kind of task so far. To do the calculation and run the algorithms the camera was hooked to a raspberry pi 4 with 4gb of memory, first it was attempted on a raspberry pi 3b+, however the need of a USB 3.0

for the camera requires the raspberry 4. In order to make the application mobile, the battery is provided by a hat using two 18650 batteries. The sound is produced by a 5v active buzzer using a 10k potentiometer to regulate the volume of the sound. The raspberry is also hooked to a standard cooler for raspberry pi. The cooler is plugged in the GPIO 1 and the buzzer on the 23. The entire setup may be seen in the figure below.

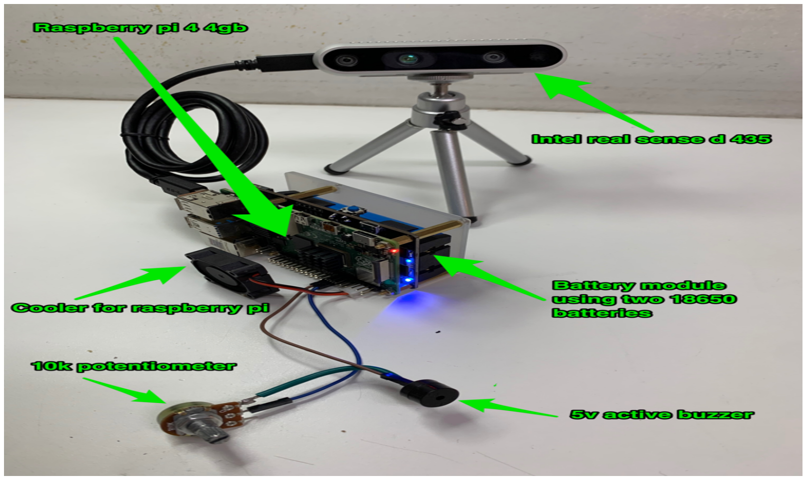

Figure 1: Prototype System

The standard approach to calculate object's distances is using disparity maps. A real world application of this may be found in [18]. Disparity maps use the concept that, on the image Pixels that are closer would move faster and those farther away would have a slower motion. From that it is possible to make an intensity map of the pixels therefore estimating the distance to the object in question. It was understood that this approach would not be time effective and besides the difficulty on program this with a good level of precision using disparity maps turned out to be unnecessary. It was decided to use the aid of a set of preprogrammed apps from AlwaysAI, one of which is called realSense_objectDetector. This program uses a camera intel real sense d 435 to give distances to the framed objects. The intel real sense d 435 camera has depth recognition making possible to have a precise calculation of the distance between the camera and the object without the necessity of any further calculations. The camera returns the distance to the framed object in an array. The next step was to run the framed image in a pre-trained classification model. These models can be obtained from many repositories, but in this case the one trained by alwaysAI themselves, "alwaysai/mobilenet_ssd", was chosen for compatibility purposes. The pre-trained model returns the prediction label for the image and the confidence that the code has on the result.

However, the program had some adaptations that would be required to achieve the goal of the project. One of them was the results object from the code. Originally the results are displayed in a browser window that the user needs to navigate too. On this window the image that is currently displaying on the camera is shown together with the distance to the object and the prediction label that the code gave using the pre-trained model. For the intended purpose this was not very helpful, since the goal was the actively use the result of the distance + prediction label. The code was then modified in a way that instead of sending the live video, the distance and the prediction label to the browser the distance and the prediction label would be output by the terminal. This made manipulations with the results much more intuitive. Once that was obtained the question of how to alert the user was raised. Now the code could detect images and classify than as person with quite good accuracy, more in the next section, but this needed to arrive for the user in an informative way. To solve this issue a 5v active buzzer was implemented. The decision for the buzzer over other solutions like a speaker, was made considering the directness of the information provided by the buzzer together with its low cost and simplicity to be implemented. This simplicity comes on the form of the GPIO python library. In which it is possible to either activate or deactivate certain GPIOs of the raspberry. As the Active buzzer is simply a matter of turning it on and off the simple on and off functions of the GPIO library paired with the information that the buzzer was on the GPIO 23 of the raspberry was enough. The function was programmed in a way that every time that the distance returned by the code was < 2.0 or 6.7 feet the buzzer was activated. This is a safe social distance since it is the larger one adopted by government and business. As it is shown in the next sections. Since some range is required for the camera so it was preferred to use a larger distance. As the prediction label was also returned by the code adding the condition: "if prediction label = people and distance <= 2.0 then activate buzzer" was implemented. The potentiometer is used to control the volume of the beep, as the GPIO is a on or off function some kind of external volume control was needed and the 10k potentiometer was a cheap and very effective solution. The only step left was the mobility. The device would not matter if it could not be taken on the go. The solution was found on the module with two 18650 batteries. It provides at least one hour of use which is enough for testing.

Algorithm Details

In order to understand the approach taken here it is interesting to explore in depth the complexity of the way in which the device calculates the distance to the object. This will also be important later on when understanding the limitations and results obtained on testing. The explanation here will be a more high-level approach to the problem. More details may be found in [19] and [20]. On performing the calculation of the distance, the algorithm uses the idea of bounding boxes. These are rectangular areas traced around the image which the code will work with. Therefore, after the real sense camera detects the image and focuses on that it will trace a bounding box around it. Than this image will be classified using the object detector [ref]. The second parameter to the function will be the depth channel. This is provided by the intel RealSense as a separate channel. Using this information, it is a matter of combining them to print the result as the distance to the detected object. More information about the test methodology for the precision of the intel real sense d 435 may be found in [17].

Preliminary Findings

Using this information, the device was tested on different scenarios, some of which are in the Results section. It is relevant to observe however, that the internal lights and the infrared projector of the intel real sense made it possible for results with a good precision even in situations like complete darkness and different environments. On the other hand, limitations of the system including portability and practical uses will need to be addressed in the future. The weight and size of the intel real sense make a fully portable and convenient device a difficult task, and it uses of energy reduces significantly the energy life of the batteries. Also, the sensibility of the raspberry pi is a problem, it is known that blind people tend to need devices that a little more durable, more of that in the future work section. Also, the beeping may be inconvenient in large groups of people or in situations where other beeping is happening, for example, a truck backing up. But the findings are still first step.

Results

On the introduction section the question of how does a blind person may be able to complain with the same rules as sighted people for social distance. In this section it will be demonstrated how to use Blind Distance to do just that. The device was positioned in a glass table and pointed straight forwarded into an empty area. It will be demonstrated that whenever people get to a range of less than 2 meters from the device the 5v beep will activating alerting the user for the proximity of another person. The code is run remotely using ssh and the results are obtained. Different people walk on different directions from the camera until the distance is reached. It was found that when programmed for 2 meters on the code the beep activates when the person is anywhere between 1.9 and 1.95 meters from the camera facing straightforward. The results whenever the person is coming from the left or from the right in different angles differ a little. It was found that the closer that the beep took to buzz was 1.26 meters coming in a 30 degrees angle and on totally dark condition. This was recorded around midnight outside with the lights off. Regarding rotation to the left or to the right, it was found that the device detects people in a maximum distance of 30 degrees angle to the left or to the right. During the day in perfect conditions the detection happened around 1.8 meters on a 30 degrees angle, which is considered the worst-case scenario. This test was performed with different people in different light conditions and the results gave very similar results. Other conditions included testing on a dog positioned at the same distance and the beep did not activated as it did with the presence of people. Asking for the subject to sit down was also a conditioned in which the beep activated in the same distances. Other trials were performed with people standing behind lower objects such as a chair or bench and the beep activated around the same distance. For training purposes, the maximum distance that the camera would register a person was registered around 3.15 meters. Asking for people to slowly walk from 3.15 meters to the distance that beep would activate was the test methodology to study the precision of the beeping.

Conclusions

Developing the Blind Distance device was both a personal and academic achievement. During the times that the society in general is living on 2020 it is very challenge for everybody to go on with normal routine, and especially for those with disabilities. Blind people are struggling more than ever to reduce the gap between living as a blind person and as a non-blind person. It is not only the challenges that normally are raised by the simple fact that we are blind that is a complicated factor now but also the new laws that are being imposed on society. As it was mentioned in the introduction in order for blind people to have the same rights as everybody it is a good action from our part to do whatever possible to complain with the same laws as everybody else. It is understood that the device proposed here achieves this goal. But it stretches a little more. By letting the user know when someone is close it makes possible for blind people to proceed respecting the same laws brining a more equal society for both blind and sighted people. The hope is that on the future the device may be upgraded to include a speaker instead of a buzzer and provide more complex information to the user, therefore, being a more technological answer to the navigational challenge that today is answered by the white cane and the guide dog. By using a raspberry pi, the possibilities that raised by future programs that may be implemented may bring the blind community closer to the technological advances of the 21st century.

Future Work

Because of the limitations of quarantine and lockdowns it was not possible to test the device in real world situations like walk on a street or going to a more crowded area. This is a future area of testing that may present interesting results. Also testing with different people to understand the response to the beep stimuli, it would be a useful concept to understand if everybody perceives the beeps in the same way or if they ended up becoming more distracting to some people and less to others. All of these testing are not difficult to do once the restrictions of quarantine are lifted and they may bring conclusions for the future updates mentioned in the conclusion section.

Bibliography

1. Feng, Catherine & Azenkot, Shiri & Cakmak, Maya. (2015). Designing a Robot Guide for Blind People in Indoor Environments. 107-108. 10.1145/2701973.2702060.

2. Nieto Martelo, Luis & Padilla, Carlos & Barrios, Mauricio. (2014). Design and implementation of elctronic aid to blind's cane. 10.1109/CIIMA.2014.6983468.

3. Sivan, Shankar & Darsan, Gopu. (2016). Computer Vision based Assistive Technology for Blind and Visually Impaired People. 1-8. 10.1145/2967878.2967923.

4. Islam, Md.Iftekhar & Rahman, MD & Raj, MD & Hossen, Sabbir & Nath, Shantanu & Imam, Mohammad. (2018). An Indoor Navigation System For Visually Impaired People Using a Path Finding Algorithm and a Wearable Cap. 10.1109/I2CT.2018.8529757.

5. D'Atri, Alessandro & Medaglia, Carlo & Serbanati, Alexandru & Ceipidor, Ugo & Panizzi, Emanuele & D'Atri, Edoardo. (2007). A system to aid blind people in the mobility: A usability test and its results. 2nd International Conference on Systems, ICONS 2007. 10.1109/ICONS.2007.7.

6. Ando, B.. (2009). A Smart Multisensor Approach to Assist Blind People in Specific Urban Navigation Tasks. Neural Systems and Rehabilitation Engineering, IEEE Transactions on. 16. 592 - 594. 10.1109/TNSRE.2008.2003374.

7. Maingreaud, F & Pissaloux, Edwige & Velázquez, Ramiro & Gaunet, Florence & Hafez, M. & Alexandre, J-M. (2005). A Dynamic Tactile Map as a Tool for Space Organization Perception: Application to the Design of an Electronic Travel Aid for Visually Impaired and Blind People. Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference. 7. 6912-5. 10.1109/IEMBS.2005.1616095.

8. Nunokawa, Kiyohiko & Seki, Y & Ino, S & Doi, K. (2014). Judging hardness of an object from the sounds of tapping created by a white cane. 2014. 5876-9. 10.1109/EMBC.2014.6944965.

9. Lahav, Orly & Mioduser, D.. (2004). Blind persons' acquisition of spatial cognitive mapping and orientation skills supported by virtual environment. Virtual Reality & Assoc. Tech. 4. 10.1515/IJDHD.2005.4.3.231.

10. Jayant, Chandrika. (2010). MobileAccessibility: camera focalization for blind and low-vision users on the go. ACM SIGACCESS Accessibility and Computing. 96. 10.1145/1731849.1731856.

11. Wong, F. & Zelek, John. (0001). Tactile & Inertial Patterns from a Long White Cane. 2006. 519 - 524. 10.1109/BIOROB.2006.1639141.

12. Whitmarsh, Lorraine. (2009). The Benefits of Guide Dog Ownership. Visual Impairment Research. 7. 27-42. 10.1080/13882350590956439.

13. Shiizu, Yuriko & Hirahara, Yoshiaki & Yanashima, Kenji & Magatani, Kazushige. (2007). The development of a white cane which navigates the visually impaired. Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference. 2007. 5005-8. 10.1109/IEMBS.2007.4353464.

14. Clark-Carter, D. D., Heyes, A. D., & Howarth, C. I. (1986). The efficiency and walking speed of visually impaired people. *Ergonomics*, *29*(6), 779–789. https://doi.org/10.1080/00140138608968314

15. Jafri, Rabia & Ali, Syed & Arabnia, Hamid & Fatima, Shameem. (2013). Computer vision-based object recognition for the visually impaired in an indoors environment: a survey. The Visual Computer. 30. 10.1007/s00371-013-0886-1.

16. Mahmud, N. & Saha, R.K. & Zafar, R.B. & Bhuian, M.B.H. & Sarwar, Syed. (2014). Vibration and voice operated navigation system for visually impaired person. 2014 International Conference on Informatics, Electronics and Vision, ICIEV 2014. 1-5. 10.1109/ICIEV.2014.6850740.

17. A. G. Hapsani, D. Syauqy, F. Utaminingrum, P. P. Adikara and S. Adinugroho, "Onward movement detection and distance estimation of object using disparity map on stereo vision," 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), Dubai, 2017, pp. 41-44, doi: 10.1109/ISCBI.2017.8053541.

19. https://www.alwaysai.co/docs/edgeiq_api/real_sense.html

20 https://www.alwaysai.co/docs/edgeiq_api/bounding_box.html

21. Aslam, Fahim. (2020). COVID-19 and Importance of Social Distancing. 10.20944/preprints202004.0078.v1.