World Navigation Hat – Visual to Audio and World Mapping

Jason D'Souza*
University of San Agustin
jdsouza@usa.edu.ph

Ethel Herna Pabito
University of San Agustin

ChenLin Wang
University of San Agustin

Vince Ginno Daywan
University of San Agustine

Abstract

We present a device that is capable of acting has the user's eyes allowing the visually impaired to navigate the world and the general population or those with spatial awareness issues to more easily keep track of their surrounding and their operating system. The device works by taking environmental data such has visual, audio, gyroscopic, etc. sensors and using it to make a virtual work map on which the user especially those with visual imparities can more easily navigate, while also allowing more addition features like hand gesture recognition to allow the user to navigate their operating system removing the need to hold a phone, or integrating language models for artificial assistants and so on. The device is not yet developed has it is still in the prototype phase

Table of Contents

1. Introduction

{State the growing or current number of visually impaired, their status (partially or completely impaired, etc.}. There are a lot of methods that attempts to provide them techquies to navigate the world which I categorized into two groups: Primary where the visual data is primary processed by the person themselves and Secondary where the visual information is processed by some other entity before provided to the person. Examples of pre-existing solution include something has simple has a guide dog (secondary), cane (primary), or advance has AI powered visual recognition devices (secondary), and our target Sensory Substitution Devices (primary). Here the paper's two primary philosphies come into play

a. Sensory Substitution of Visual and Depth to Sound

SSDs(Sensory Substitution Devices) more specifically when it comes to Visual to Sound such has "The vOICe" (Doward, 2014), "EyeMusic", and "SoundSight App" shows that users with enough training can identify shapes, letter, objects, and even navigate environments from substituted sound alone. What makes this theory different is that it attempts to solve key challenges and limitations of these devices, the theory states that by incorporating depth projected into a virual environment then partially translated into a spectrogram (Previous limitations being limited spatial and temporal bandwidth is solved by this theory by only translating the direct vertical view of the user) with taking account of the Equal-loudness contour(solves the limitation of sounding unpleasant or harsh over load periods of time) and changing heatmaps(solves the limitation of Sensory overload and cognitive load) than the users can more easily learn to use the device (Previous limitations being complex training requirements). The existing SSDs although promising results haven't gained widespread traction because of all of these issues plus being high cost and uncomfortable (solved by being a hat instead of near the sensitive body parts like eyes and ears). Hopefully this theory could possibly allow true perceptual integration by not only focusing on hardware or the software aspects but also human consciousness and psychology.

b. Argumented Modular Operating System

This objective is that the operating system the raspberry pi will implement will not be limited to the primary feature but since it already includes capabilities to handle applications since users all have different needs and wants, for example if the user is unable to use their phones either through visual imparements or inconvenience of holding a phone an application can be installed that uses the Hat's camera to detect the hand guestures of the user to navigate the system. This objective also allows applications to provide the user a personal argumented assistant that could observe the environment, keep track of previous situations, and much more to aid the user in their unique needs and wants. This also opens the door to IoT and automation systems has applications like for health keeps tracks of specific sensors to a centralized systems, or automation applications with combination like with the hand guesture applications allows the user to automatically help them with their work with minimal effort.

2. Related Work

What separates this project from other similar project is that the visual data that is captured from the environment is directly projected into audio for the person to develop their own mental map of the world, this is based on this paper's theory that if visual information primarily depth information is projected into a spectrogram with few human calibration like taking into account the Equal-loudness contour or Changing heatmaps than the person can make their own mental map of the world without a middle man to process the data for them like Object detect or assistant. Another unique feature that this paper presents is modularity has not everyone has the same problems or needs as everyone else, since the main controller of this device is a Raspberry PI then we take full advantage of this by allowing applications like to message or call people, applications to use the built-in cameras to also detect hand gestures to navigate the operating system without voice, and basically any feature of an operating system.

a. Object Detection and Voice Assisted Navigation: Smart Hat

https://journal.iba-suk.edu.pk:8089/SIBAJournals/index.php/sjcms/article/view/1535/469

The visual data processed by the device undergoes first with an Object Detection model and even incorporates several audio cues, while this is helpful for the person it doesn't completely emulate the eye's of the users, this basically acts like another people helping the visually impaired rather than the visually impaired having a secondary eye, which is the primary theory of this paper. This device is also limited by the microprocessor which is an ESP32-CAM limiting the features of modularity and the paper doesn't seem to take into account different conflicting interest of the users.

b. Clearway Companion – an AI powered AI for Visually Impaired

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4809356

This paper is similar to the previous paper but has a similar idea to ours by also taking into account distance (depth), environmental changes (PIR motion sensor), and even a Raspberry PI has the main microcontroller. This paper however doesn't attempt to prove our paper's theory of mind map has it also uses Object detection with its estimated distance has the primary method of observation, and also doesn't allow applications for people's unique needs despite using a Raspberry PI

3. Methodology

This research method used is the Qualitative Methodology, we will {fill in the blanks}

4. Results

5. Discussion/Analysis

6. Conclusion

7. References

Doward, J. (2014, December 7). *vOICe: the soundscape headsets that allow blind people to 'see' the world*. Retrieved from The Guardian: https://www.theguardian.com/society/2014/dec/07/voice-soundscape-headsets-allow-blind-see