INDIVIDUAL PROJECT IDEA REPORT

CONCEPT OUTLINE

For my individual project idea, I have chosen to augment the sense of vision, more specifically emotion recognition for the blind and visually impaired. The concept is a tool which detects non-verbal cues of subjects on a computer screen (online meetings, films, TV shows, YouTube etc) and provides real-time feedback via braille to the user.

As of March 2020, there are over 250,000 people in the UK registered as blind or partially sighted (1) and while tools such as guide-canes can help with interactions in the physical world, human and social interactions can be difficult without visual and non-verbal cues. Knapp et al find that up to 65% of expression in a two-way interaction can be non-verbal (2) and for the visually impaired this impacts their ability to socialize and interact effectively. For many blind people, rates of anxiety and depression are significantly increased as a result (3). In an increasingly online world, this issue extends to the digital world with scenarios like online meetings where audible cues such as tone may be distorted and unclear. The proposed device aims to close this gap and help the user more accurately detect emotions and expression.

BACKGROUND

Existing computational methods to detect and classify emotions rely on computer vision and deep learning techniques such as those implemented in the Competition on Affective Behaviour Analysis in-the-Wild (ABAW), where participants compete to build computational models to detect human emotions in real-time (4). However, despite the rapid advancement in the efficiency and accuracy of these models, one of the key issues lies in accurately defining human expressions.

In 1969 Paul Ekman proposed a discrete system which distinguishes between six, universal, 'basic emotions': happiness, sadness, anger, fear, disgust, and surprise (5). Many find this to be an oversimplified representation of human expression and so continuous systems were developed to categorize dimensions of expression as opposed to specific emotions. One of the most widely used is James Russell's valence-arousal model which proposes two axes of expression: pleasant/unpleasant and calm/aroused (6). However, despite the wide-spread use of both classification systems, including in the ABAW competition, some studies highlight their limitations in describing the extremely complex and dynamic nature of human expression (7).

Alternatively, Alan Cowen of Hume.ai finds that there are dozens of varieties of emotions perceived in expressions of the face, body, and voice across over 30 dimensions (8). His company provides a framework for building custom models, which can detect 48 different emotions and output the intensity of each. However, translating 48 different emotions onto a physical interface for the visually impaired may be quite complex (see Figure 1). I believe a good solution is to use a refreshable braille display in combination with the software to relay to the user which emotions are currently being detected in real-time.

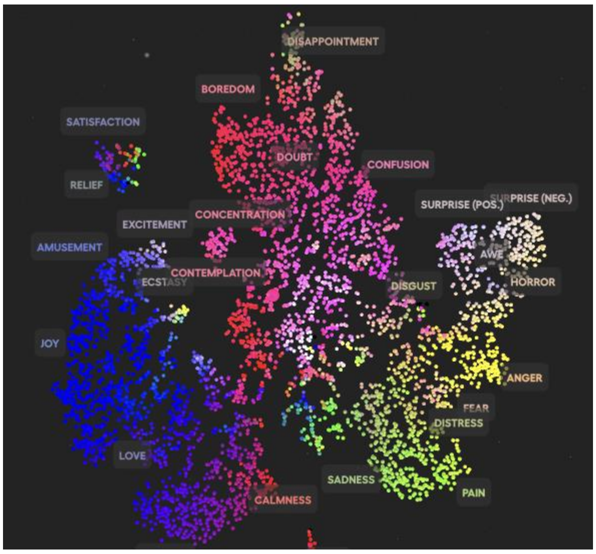

FIGURE 1: A DIAGRAM SHOWING THE COMPLEX SPACE OF HUMAN EMOTIONS DEVELOPED BY HUME.AI (11)

DEVICE OUTLINE

SOFTWARE

The software aspect comprises of an application based on the models developed by Hume.ai. This application will make use of Hume.ai's streaming API which performs real-time analysis and outputs the respective intensity of all 48 emotions. The API outputs a csv file containing the relative strengths of each emotion at any one time but to simplify the device, I propose that only the three most strongly detected emotions should be outputted to the user. This also allows for a reduced size of the braille display and a more ergonomic device for the user. I believe the intensity of the emotion is also important for the user to understand how confident the model is in its analysis and so the device will produce a braille output equivalent to the one in Figure 2. The rate at which the application outputs emotions to the user is also an important aspect due to the limitations of braille reading speeds and avoiding an unnecessarily distracting device. I think the user should have control over this refresh rate to adjust it to their preferences and abilities, and this could be in the form of a simple dial on the device. However, despite these considerations, a constantly refreshing output may be distracting to the user if they are trying to focus on other aspects of the computer content and so an alternative 'manual mode' could prove useful. In this mode, the emotion output is refreshed each time the user presses a button on the device.

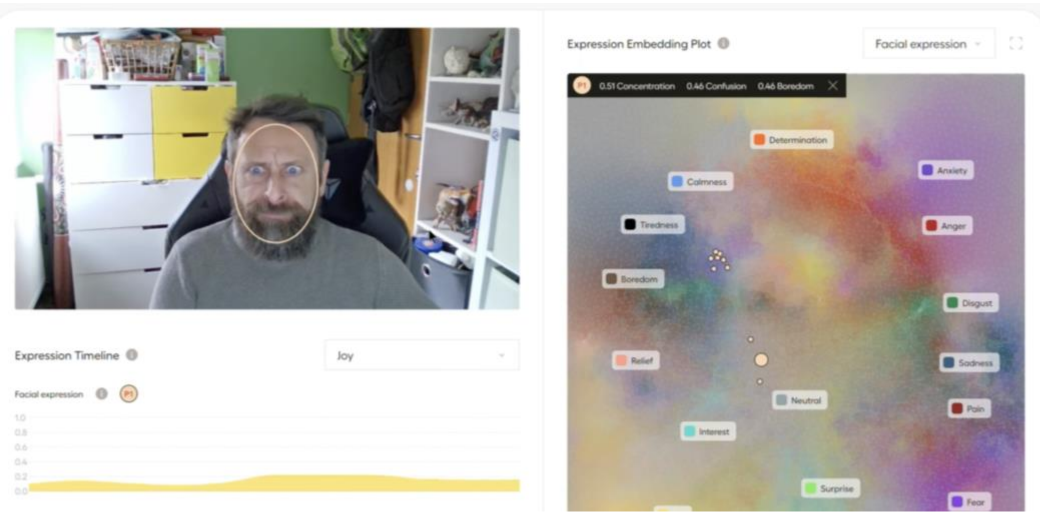

FIGURE 2: AN EXAMPLE OUTPUT OF THREE DETECTED EMOTIONS USING HUME.AI'S STREAMING API (10)

Hardware

There are already existing devices which provide refreshable braille displays to assist the visually impaired with tasks such as reading, but these can cost thousands of pounds due to the expensive piezo-electric actuated pins. This not only exceeds the budget of this project but can provide a barrier to lower-income users who could also benefit from such devices. An alternative solution for braille cells has been developed by Vijay Virada who makes use of 3D-printed materials and electromechanical actuators to achieve much lower costs (9). These are therefore the desirable solution to build a cheap and custom length braille module for my device. Other required parts include a dial/knob to vary the refresh rate, a button for manual mode, and a 3-position toggle switch for selecting the mode or turning the device off, all of which are relatively inexpensive. A USB connection to the computer can be used to simultaneously power the device and allow it to communicate with the application.

Figure 3 shows a sketch laying out a rough design of the device.

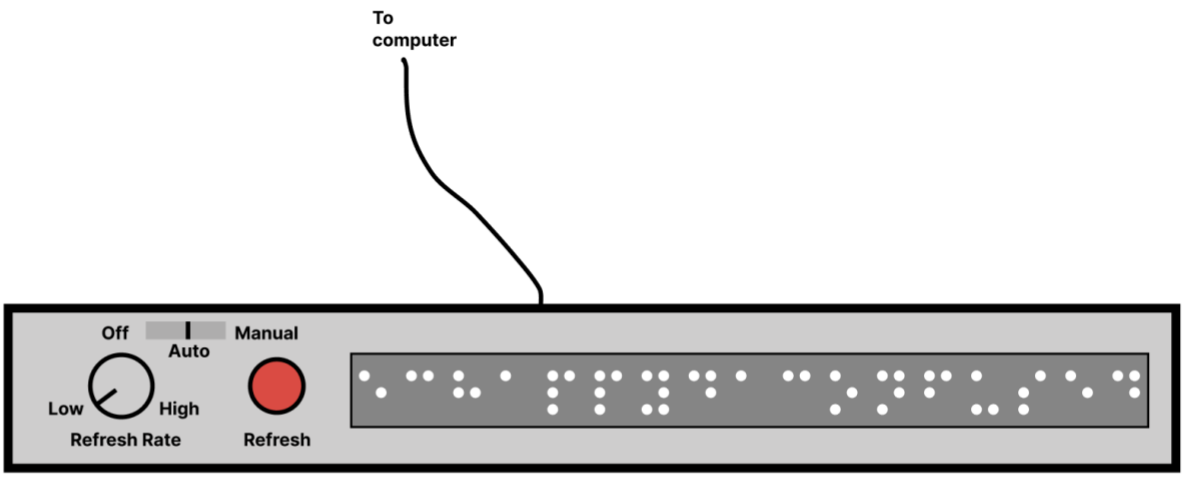

FIGURE 3: A ROUGH SKETCH OF THE DEVICE (NOT-TO-SCALE)

References

1. *Registered blind and partially sighted people, england 2019-20*. NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/registered-blind-and-partially-sighted-people/registered-blind-and-partially-sighted-people-england-2019-20 [Accessed 31st January 2024].

2. Knapp ML, Hall JA. *Nonverbal communication in human interaction*.. 6th ed. Belmont, CA: Thomson/Wadsworth; 2006.

3. Klauke S, Sondocie C, Fine I. The impact of low vision on social function: The potential importance of lost visual social cues. *Journal of Optometry*. 2023;16(1): 3–11. https://doi.org/10.1016/j.optom.2022.03.003.

4. Kollias D, Tzirakis P, Baird A, Cowen A, Zafeiriou S. Abaw: valence-arousal estimation, expression recognition, action unit detection & emotional reaction intensity estimation challenges. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Vancouver, BC, Canada: IEEE; 2023. p. 5889–5898. https://doi.org/10.1109/CVPRW59228.2023.00626. [Accessed 1st February 2024].

5. Ekman P. Universal facial expressions in emotion. Studia Psychologica. 1973 Jan 1;15(2):140-7.

6. Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology*. 1980;39(6): 1161–1178. https://doi.org/10.1037/h0077714.

7. Keltner D, Brooks JA, Cowen A. Semantic space theory: data-driven insights into basic emotions. *Current Directions in Psychological Science*. 2023;32(3): 242–249. https://doi.org/10.1177/09637214221150511.

8. Cowen A, Sauter D, Tracy JL, Keltner D. Mapping the passions: toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*. 2019;20(1): 69–90. https://doi.org/10.1177/1529100619850176.

9. *Electromechanical refreshable braille module*. https://hackaday.io/project/191181-electromechanical-refreshable-braille-module [Accessed 1st February 2024].

10. *Hume ai facial expression*. https://www.youtube.com/watch?v=5ALNG6fm4XQ [Accessed 1st February 2024].

11. *Facial expression | hume ai*. https://hume.ai/ [Accessed 1st February 2024].