*Vocal*Focals

Mariah Louderback, Jay Napp, Chris Odom, Haley Patterson, Taryn Trusty, and Andrew Zeller
Department of Mechanical Engineering, The University of Alabama
Dr. Dan Fonseca

College of Engineering
Mechanical Engineering

Background

Clients: Dr Beth Todd and Dr. Dan Fonseca

For visually impaired individuals, interacting with printed text of normal-sized font is a common annoyance. For nearly 70% of visually impaired individuals, printed text of any size font is inaccessible

Problem Description

Goal: design a proof-of-concept device to empower the visually impaired by allowing them to read text in their environment

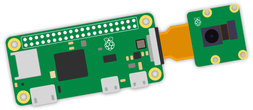
***Vocal*Focals solution:**

- Sunglasses with an embedded camera, processor, and audio output
- Identifies text in front of users and plays it back over audio

Design Requirements

Driven by user experience

- Inconspicuous
- Headpiece weight does not exceed 150 g
- Battery Life of 10 hours
- Durable
- Detect printed text with a minimum 75% accuracy
- Response time of less than 30 seconds



Design Concept Generation

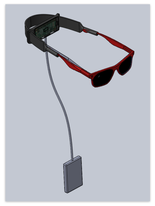


Figure 1
Initial assembly model

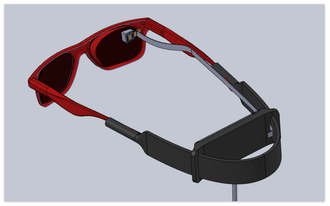


Figure 2
Initial design

Final Design

Figure 3
Croakies design

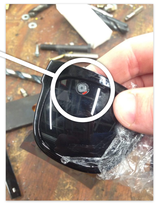


Figure 4
Integrated camera

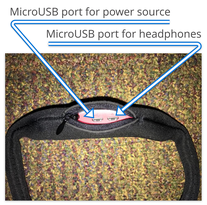


Figure 5
Raspberry Pi power and headphone ports

Testing Results

Camera

- Cable continuity
- Picture quality
- Assembly durability
- Reliability

Software

- User interface
- Reliability

Power Source

- Battery pack

Audio

- Quality
- Reliability

Glasses Assembly

- User experience
- Durability / quality

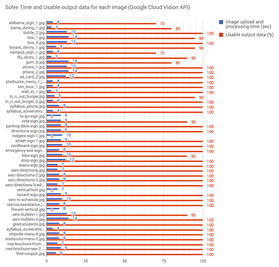
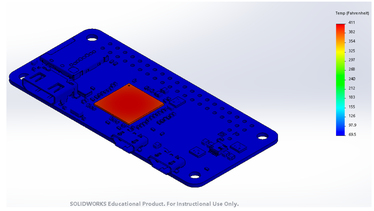


Figure 6
Google Cloud Vision API testing results

Thermal Analysis



Boundary Conditions:
Ambient room temperature: 70°F
Convection coefficient (h): 25W/m²K
Power input: 0.76 W

Figure 7
Thermal analysis of Raspberry Pi Zero W

User Experience

Device Startup

- User plugs device into power source
- User is greeted with audio welcome message
- Text detection automatically begins running

Detection

- Camera captures image
- Device plays *"scene captured, please wait"* audio message
- After a short wait, device plays back smooth audio message with any text, labels, or logos found in image
- Sequence repeats at a 30 second interval

Software

Version control and collaboration: GitHub
Open source linux platform: Raspbian OS on Raspberry Pi
Programming Language: Python
Google Cloud Platform APIs used:

- Google Cloud Vision
- Translation
- Text-to-speech

Application to run scripts when device starts up: Cron

Google Cloud Platform

Final Deliverables

Functioning Prototype

- Able to recognize printed text
- Automatic translation from other languages to English
- Object label detection
- Brand logo detection
- Average response time of 8 seconds
- Accuracy of 96 %
- Unit cost of $90.79

Codebase on Github for further development
Bill of Materials

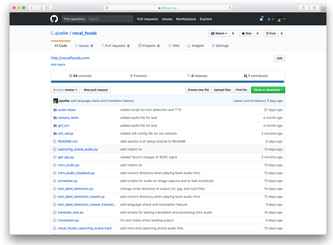


Figure 8 GitHub Codebase

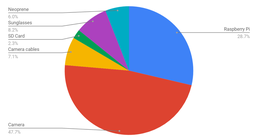


Figure 9 Budget breakdown

Other / References

Website: *vocalfocals.com*

In the process of patenting through UA OTT

Reference:

National Federation of the Blind. (2017, September). *Statistical Facts about Blindness in the United Statesction*