SENIOR DESIGN I

Sean

Sound Enhancing Autonomous Network

DEPARTMENT OF ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF CENTRAL FLORIDA
Dr. Samuel Richie and Dr. Lei Wei

Initial Project and Group Identification Document
Divide and Conquer

Group 14:

Annette Barboza	EE	netb@knights.ucf.edu
Ayanna Ivey	CpE	a.ivey@knights.ucf.edu
Brandon Kessler	CpE	brandon.kessler44@knights.ucf.edu

Advisors/Contributors: Jonathan Tucker, *Signal and Image Processing*
Wasfy Boushra Mikhael, *Analog and Digital Signal Processing*

I. Project Narrative

The direction technology has been taking recently makes the lives of consumers more convenient in *almost* every conceivable way. However, there have been very few attempts to make products for those who live with disabilities as accessible as products that are simply for convenience. In the case of hearing impairment, approximately 15% of American adults report having some trouble hearing [1].

Hearing loss presents itself in three different types: conductive hearing loss, sensorineural hearing loss, and a mix of both. In permanent cases, conductive hearing loss generally affects the overall loudness of a sound, while sensorineural can affect loudness and perception of tone [4]. For the majority of the people who live with any of these, their options are limited to potentially invasive procedures in an attempt to correct the hearing or using hearing aids.

Hearing-aids are one of the most widely available removable solutions to hearing loss. They help users by converting sound to digital signals, amplifying those signals, and passing the amplified signal back to the user as sound [2]. Although some higher end hearing-aids can take the user's environment into account and try to reduce noise, a common complaint among hearing-aid users is that all sound, including unwanted background noise, is amplified and does nothing to help them hear what they want to hear as displayed in figure 1. A few groups have attempted to correct this same problem. Gupta et al. from the University of Massachusetts approached this problem in their senior design project. Their solution allowed the user to independently control the volume of different people in a room. They used an Xbox Kinect to detect individuals and a microphone array to perform beamforming and assign sound to a person while eliminating noise. However, their system was stationary and could only work in the room it was calibrated to [3].

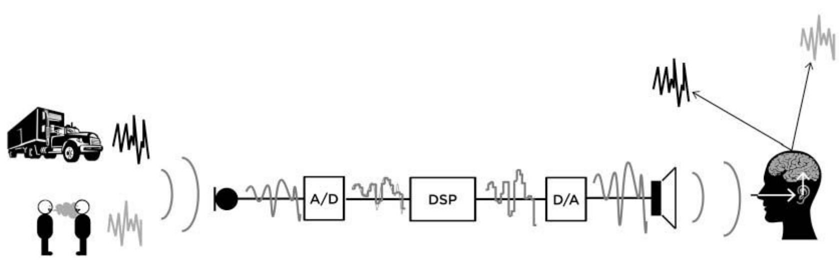


Figure 1. Example of speech and noise being combined and confused during processing. This leads to the amplification of both signals and hard for the user to hear the sought after conversation [9].

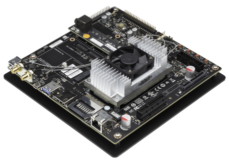
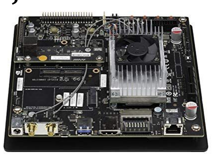
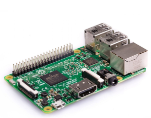
Our solution to the aforementioned issues is to design and create a noninvasive alternative to hearing aids, Sean. Sean is a sound enhancing autonomous network that utilizes practices of deep learning to detect a person in view of the user and analyzes audio to focus on human voices. The goal is to build a portable device that users can easily configure and start using to accurately amplify the sound they want to hear in bluetooth connected headphones or earbuds. The targeted users are people who have permanent conductive hearing loss and/or sensorineural hearing loss, as Sean raises the intensity of voices and diminishes background noise and helps clarify voices.

Sean aims to improve the quality of life of those who are hearing impaired by replicating the experience of being able to focus in on a person speaking in an indoor environment. This device works in real-time to lower the background noise in an indoor environment with a moderate to high signal-to-noise ratio to clearly make out what a person is saying. Sean will use computer vision to help detect when a person is in view of the user and digital signal processing methods to separate the background noise and the voice of the detected person. This will consequently lower the amplitude of the background noise and raise the amplitude of the voice so that it becomes the dominant signal.

II. Hardware

Sean will take input in two ways. The first will be audio captured through microphones, and the second will be video captured by cameras. The data taken though these devices will have to be processed through processors and outputted to wireless headphones. The processing will occur with a latency of 30ms at most as this has been characterized as the maximum amount of time a human will not notice a sound discrepancy [7]. This whole system will be portable and will therefore need a portable power supply. A trade study is being conducted as shown in table 1.

Hardware Trade Study

Table 1 -- Trade study on different processor options paired with compatible hardware				
Processor	**Specs**	**Compatible Cameras**	**Compatible Microphones**	**Compatible Power Supply**
NVIDIA JETSON™ TX2 	• GPU: NVIDIA Pascal™, 256 CUDA cores • CPU: HMP Dual Denver 2/2 MB L2 + Quad ARM® A57/2 MB L • Memory : 8 GB 128 bit LPDDR4 59.7 GB/s • Bluetooth 4.0 enabled	**TBD**	**TBD**	**TBD**
NVIDIA JETSON™ TX1 	**TBD**	**TBD**	**TBD**	**TBD**
NVIDIA JETSON™ TK1 	**TBD**	**TBD**	**TBD**	**TBD**
RASPBERRY PI 3 	**TBD**	**TBD**	**TBD**	**TBD**

III. Block Diagram

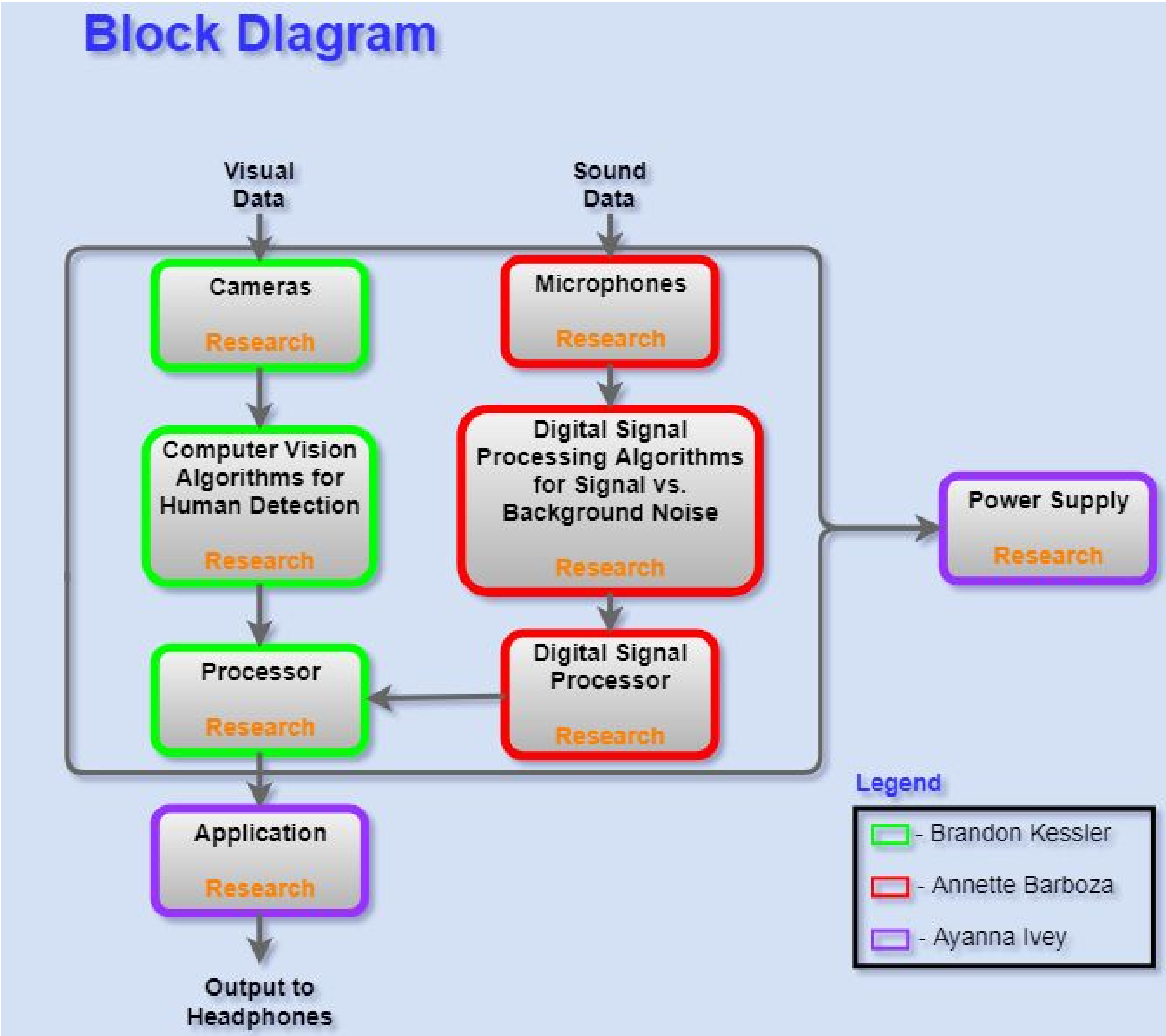


Figure 2. System Block Diagram

IV. Specifications

Any specification preceded with "Option:" are being considered and will be decided based on feasibility and time constraints

- Cameras
  - Forward facing camera with respect to the system
    - 30 frames per second (fps) minimum capture rate
    - 720p minimum resolution
    - 60 deg minimum diagonal Field Of View (FOV)

        - 44.2 deg minimum horizontal FOV
        - 25.8 deg minimum vertical FOV
        - Compatible with chosen processor

    - *Option:* Backward facing camera with respect to the system
        - 30 frames per second (fps) minimum capture rate
        - 720p minimum resolution
        - 60 deg minimum diagonal Field Of View (FOV)
        - 44.2 deg minimum horizontal FOV
        - 25.8 deg minimum vertical FOV
        - Compatible with chosen processor

- Computer Vision (CV) Algorithms for Human Detection
    - Autonomously detect humans within the FOV of the camera
        - 10 fps minimum processing rate
        - Detects humans up to 20 ft away
        - Detects up to 20 humans per frame
        - Correct detection rate of 90%
    - *Option*: Autonomous lip reading recognition
        - 10 fps minimum processing rate
        - Reads lips of one human up to 5 feet away
        - Correct reading rate of 80%
    - *Option*: Voice generation
        - 10 fps minimum processing rate
        - Generates voice of one human up to 5 feet away

- Processor
    - Embedded for real-time processing (Including Development Kit)
        - 15W maximum power consumption
        - 1.1lb maximum weight
        - 7.1in. x 7.1in. maximum size
        - 4-core @ 1.5GHz minimum CPU
        - 180-core minimum GPU
        - 4GB minimum memory
        - Bluetooth 4.0 or greater enabled
        - Latency of no more than 30ms
- Microphones
    - Array of microphones to convert sound to digital signals
        - 8-20 MEMS microphones in array

        - ~-26 dB @ 94 dB SPL (normal for digital microphone)
        - Omnidirectional
        - ~60 dB SNR
        - Operating frequency range: 125 Hz-8kHz (avg for CIC hearing aid)
- Digital Signal Processing Algorithms for Signal vs. Background Noise
    - Standby State--no cue from CV Algorithms
        - Stay idle allowing only noise cancellation from digital signal processor when no humans are present allowing background noise to sound natural and non-intrusive
    - A third state?
    - Operating State--cue from CV Algorithms
        - Use beamforming to locate source of sound and amplify it while simultaneously lowering the background noise
- Digital Signal Processor
    - Beamforming (up to 20 dB attenuation)
    - Audio sampling rates 8 kHz to 216 kHz
    - Linear phase FIR filter
    - Noise suppression (up to 20 dB attenuation)
- Power Supply
    - 4 hours of continuous power to entire system
    - Compatible with all hardware components
- *Option* Phone App
    - Connected through bluetooth 4.0 or greater
    - Controls Volume
    - Controls Sensitivity
    - *Option*: Ability to choose individuals to listen to
- System Housing
    - 5 pounds or less
    - Entire system contained
    - *Option:* Wearable system
        - Custom fitted backpack

V. House of Quality

- Engineering Requirements -
- Marketing Requirements -
- ↑ ↑ - Strong Positive Correlation
- ↑ - Positive Correlation
- ↓ - Negative Correlation
- ↓ ↓ - Strong Negative Correlation
- + - Positive Polarity
- - - Negative Polarity

Table 2 -- Engineering-Marketing Trade-off Table

		Efficiency	Output Power	Implementation Time	Weight	Cost	Dimen-sions	SNR	THD
		+	+	-	-	-	-	+	-
High Power	+	↑ ↑	↑	↓ ↓	↓ ↓	↓	↓ ↓	↓ ↓	↓
SNR	+	↑	↓ ↓	↓ ↓	↑	↑ ↑			↑ ↑
Cost	-	↓ ↓	↓	↓ ↓	↓ ↓		↓	↑	↑
Latency	-	↑	↓	↓ ↓		↓ ↓		↑	↑
Portability	+	↓	↓	↓ ↓	↑ ↑	↓	↑ ↑	↑	
User Friendly	+	↑	↑	↓	↑ ↑		↑ ↑	↑ ↑	↑ ↑
Accuracy	+	↓		↓ ↓		↓	↑	↑ ↑	↑ ↑
Resolution	+	↓	↑ ↑	↓		↓		↑ ↑	
Engineering Requirement Targets		>= 70%	<= 30W	<= 8 weeks	<= 5 lbs	<= $1800	<= 93.45 x 67.35 x 42.45 (mm)	< 60 dB	< 1% @ 95 dB SPL

VI. System Housing Mock Up

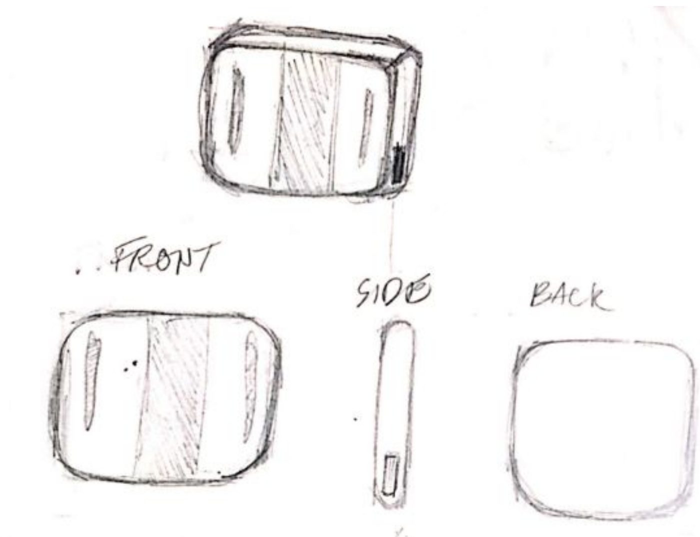


Figure 3. Mock up design of device housing

Dimensions: 93.45 W x 67.35 H x 42.45 D (mm)

Weight: ~138.6 g

*Dimensions and weight have not been finalized.

VII. Estimated Budget and Projected Financing

The goal is for this project to be sponsored by either a company or individual that believes in the successful outcome of this project and supports its purpose. If no sponsorship is found we plan to fund the project ourselves.

Table 3 -- Estimated Budget		
Item	**Price estimate**	**Purchased from**
Microphone array	$50 - $100	Minidsp.com & seedstudio.com
Camera	$150 - $250	Amazon/Best Buy
Tripod	$15 - $20	Amazon
PCB**	$20 - $30	Advanced Circuits/Silver Circuits
Necessary Software	-	Plan to use open source software but may have to purchase
Possible Packaging options (backpack or hat)	$50 - $150	Customized based on technology needs
Processors	$300 - $500	
	Total: (Maximum & factoring in needing multiples of everything for testing)	$1800

** denotes critical parts that will be ordered in excess in case these parts break

VIII. Risk Mitigation Plan

Table 4 -- Risk Mitigation			
Potential Risk	**Severity/ How Likely**	**Mitigation Plan**	**Decision Deadline**
No sponsors are found.	Moderate	Fundraise through GoFundMe & be prepared to self fund if necessary.	March 9
There is no more money to further fund the project than what the budget calls for.	Moderate	Before finalizing the the budget, account for potential parts that might be might break or not work with the project.	February 14
Parts are not received in time.	High	Order same part from a secondary seller and cancel previous purchase.	March 31
Critical parts break in testing.	High	Order more than one part ahead of time to plan as if every part will break twice.	March 31
The project idea to too vast to be completed in the allotted time.	Low	Start with a couple of basic features that make the product fully operational and expand if there is time.	February 14

IX. Milestones

Table 5 -- Senior Design I & II Milestones		
Senior Design I		
Description	**Duration**	**Dates**
Divide and Conquer V1*		January 31
Divide and Conquer V2*		February 14
Decide packaging lead		February 19
Research CV, App, DSP	2 weeks	February 15-March 1
Initial algorithm trades		March 1
Come up with hard specs		March 1
Make decision about features/necessity of app		March 1
Decide what PCB will do		March 1
First round parts procurement	3 weeks	March 1-March 22
Software Development	3 weeks	March 1-March 22
Finalize parts trades	2 weeks	March 22-April 11
Mid-point Draft*		March 28
Second round parts procurement		March 31
Final Draft*		April 11
Final Document*		April 21
Senior Design II		

Description	Duration	Dates
Build Prototype	7 weeks	May 13-June 24
Test, Redesign, Test	3 weeks	June 3-June 24
Plan demonstration	1 week	June 24-July 1
Start talking to potential panel		
Final Prototype	2 weeks	June 24-July 8
Peer Presentation*		???
Final Report*		???
Final Presentation*		???

*Class-related milestones

X. References

1. https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing
2. https://www.mayoclinic.org/diseases-conditions/hearing-loss/in-depth/hearing-aids/art-20044116
3. http://www.ecs.umass.edu/ece/sdp/sdp17/team10/app/res/FPR_report.pdf
4. https://www.hear.com/hearing-loss/sensorineural/
5. https://www.invensense.com/wp-content/uploads/2015/02/AN-1112-v1.1.pdf
6. https://www.digikey.com/product-detail/en/knowles/SPH1668LM4H-1/423-1404-1-ND/5332433
7. Heurig R, Chalupper J. Acceptable Processing Delay in Digital Hearing Aids. *Hearing Review.* 2010;17(1):28-31.
8. https://www.researchgate.net/publication/261211441_A_lip_reading_application_on_MS_Kinect_camera
9. http://www.hearingreview.com/2013/03/designing-hearing-aid-technology-to-support-benefits-in-demanding-situations-part-1/