OverEar - Sound Localization Device  
Guillermo Ao, Hannah Gold, Benjamin Li, Jonathan Ngo

**Abstract​—The purpose of this project is to create a hearing device (OverEar) that will allow people with either hearing loss and/or sound localization problems to focus their concentration on the sound oriented in front of them. This device will filter out sounds of competing orientations as well as extraneous noise. OverEar is a device that makes use of a Biologically Oriented Sound Segregation Algorithm (BOSSA) provided by Kamal Sen and Kenny Chou.**

**Index Terms​— Sound Segregation, Hearing Device, Localization, Algorithm, Biologically-oriented**

—————————— ◆ ——————————

# 1 Need For This Project

Imagine attending an evening gala in a ballroom. You run into a friend at the food line and strike up a conversation. Even though there are people all around you talking at the same volume level, you are still able to identify and understand every word your friend is speaking. This is the Cocktail Party Effect, a phenomenon where one is able to discern the words that someone in front of them is speaking while filtering out any background talk and noise from other attendees at a party [2]. While this is usually a natural function of the brain, not everyone has this ability. Those with conditions such as spatial hearing loss or brain injury are not able to perform the task of isolating sounds. In this case, the condition is known as Cocktail Party Deafness [2]. Moreover, even with the introduction of hearing aids, the Cocktail Party Effect still poses a problem as traditional hearing aids do not perform sound separation, but rather amplify all sounds being received.

Hearing impaired people sometimes experience a lack of ability to localize sound. This inability can influence their ability to communicate, behave socially, and influence quality of life. However, this issue is not limited to only the hearing impaired. Some other people that struggle with the ability to localize sound can be categorized as having Auditory Processing Disorder (APD). This disorder occurs when the brain is unable to process sounds or impaired neural function [1]. Others that have trouble isolating sounds in social environments include those with Autism Spectrum Disorder (ASD). Currently, sensory symptoms have become the emphasis for clinical criteria for ASD diagnosis [3]. It is increasingly popular for scientists to study the influence of senses and perception on children as they develop into adulthood. Thus, a device localizing sound will serve as another avenue to improve the quality of communication for affected individuals.

# 2 Problem Statement And Deliverables

## 2.1 Problem Statement

We would like to create a set of headphones which takes an audio input from a noisy environment consisting of chatter, specifically one where there are multiple talkers, and enables the localization of sounds oriented in the front. The goal of the device would be to discern the localized sound the user hears and amplify it. Amplification occurs via a pre-existing algorithm derived from how animals perceive and segregate sound biologically. One algorithm is provided to us for use by Boston University Professor Kamal Sen and his previous PhD student, Kenny Chou. We have already reached out to them and received approval for this use of their algorithm, and plan to incorporate the algorithm into the headphone device. The device will consist of four main components: headphones for the user to wear, a pair of microphones which can capture sounds (as well as their frequencies, not just the sound level), a pair of speakers which provides sound clarity to the user, and an embedded board which handles the processing of the sound.

We plan to achieve this by first creating a prototype device that connects to a single-board computer (SBC) to process the sound and sends it to a pair of headphones. This prototype will serve as a proof of concept for the design and will provide validity. The second phase of the project will be to create the final self-contained embedded system.

## 2.2 Deliverables

The main deliverables will consist of two core items:

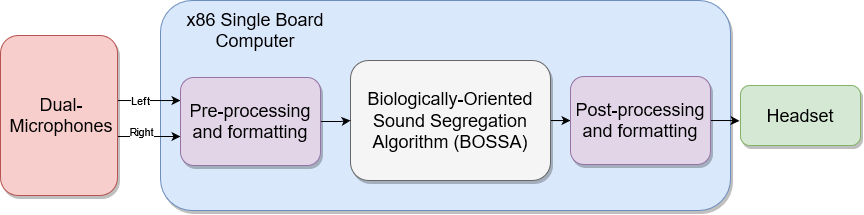
1. A prototype running off a (SBC) that, with some manual input, will be able to perform sound localization in a controlled range of noisy environments, using real-time input and off-the-shelf headphones.
2. A custom made embedded device, containing a PCB with only the necessary circuits, that will be able to perform sound localization in a range of noisy environments without manual input that works in real time with at least an hour of battery life.

We foresee our timeline being the delivery of the prototype device by November 23rd and then the custom made embedded device by ECE day.

To note: since we are using an algorithm already provided to us by Professor Kamal Sen and Kenny Chou, we expect that our software deliverables will include only any pre and post-processing we perform. Any code we develop will be available at <https://github.com/ec463>.

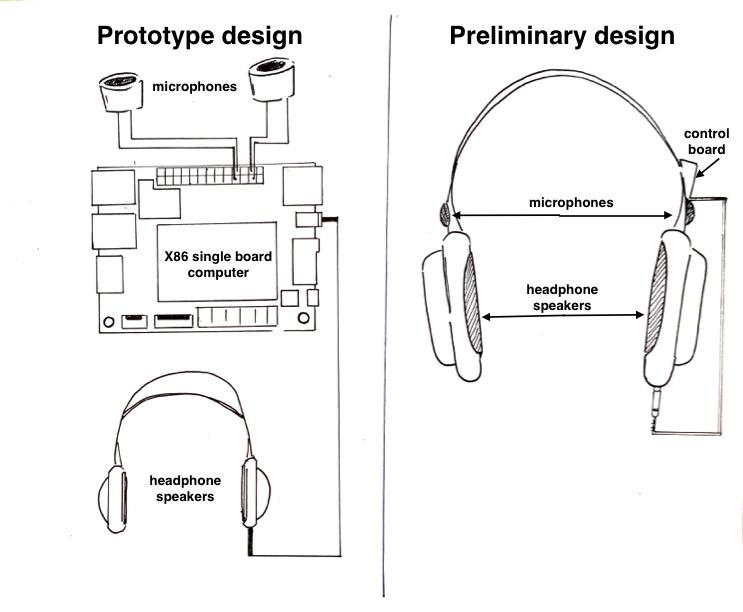
# 3 Visualization

## 3.1 Process

**

*Figure 1: This image shows the process of how the input sound moves through the system. Audio streams in from the left and right microphones of the device. These channels are used as inputs to the BOSSA algorithm. A mask within the algorithm is created and used to filter the sound. The audio is then output via the headphone speakers.*

## 3.2 Prototype



*Figure 2: Here is visualization of the prototype and its use with the LattePanda x86 2GB/32GB V1.0 microcontroller, a single-board computer that runs x86; it will be capable of processing sound in real time.*

**3.3 Software Algorithm**

The algorithm takes in the sound via a real-time stream. Then there is a component called Peripheral Filtering in which the sound is filtered with an equivalent-rectangular-bandwidth (ERB) filterbank. Then, much like our own brains, the sound is segregated into frequencies; the algorithm does so through extracting 64 frequencies within a specified range (the minimum and maximum frequencies centered around 200 Hz and 20000 Hz respectively).

Next in the BOSSA algorithm is the Midbrain model. This model represents the sound segregation that occurs in the subcortical regions of the brain. Sound streams from various spatial directions are encoded by

populations of spatially tuned neurons (STNs). The model normally takes in 5 spatial directions: Θ ∈ {0°, ± 30°, ± 60°}, but in order to be less computationally expensive, the algorithm will use only 3 Θ ∈ {0°, ± 90°}. Each neuron has a certain preference and filters this sound such that a spike train is generated for each neuron. From there, masks filtering certain orientations are created. This process filters the sound and outputs a more localized sound [8].

# 4 Competing Technologies

Solving the Cocktail Party Problem has been a major focus for companies in the hearing space. Many utilize microphone arrays to localize sound, while others take an alternative approach using camera vision. We list some competing technologies below.

## 4.1 Dual-Microphone Hearing Aids

Multiple companies have developed hearing aids that utilize two microphones per hearing aid positioned in the front and back of the device. These products utilize the time difference between noises picked up on each microphone to localize sounds. These devices use processing which does not distinguish between voices but rather reduces overall background noise.

## 4.2 Orcam Hear

Orcam provides a visual solution to the Cocktail Party Problem through a wearable face device that performs sound localization through AI [5]. A user looks at a person-of-interest, and the device uses machine learning to find out who is talking and eliminates all sound except for the target speaker. The Orcam Hear is meant to be used in conjunction with a bluetooth-enabled set of hearing aids as it beams the localized voice to the hearing aid. The price of Orcam starts at $1,500 [6]. This device is expensive and requires a lot of computation to process the audio signals.

## 4.3 Hearing Aid to Solve the 'Cocktail Party' Problem (Patent US20070223710A1)

In this patent, Peter Laurie and Tom Fortescue describe a device that uses 3+ microphones to detect the direction from where sound emanates and processes frequencies to determine which of those sounds come from speech [7]. Ultimately, a direction map is created that can identify where, relative to the listener, a sound is coming from. This patent has been abandoned for unknown reasons.

## 4.4 OverEar

Our device aims to capture the best of both competing products. We plan to stick to a two-microphone setup that will distinguish between voices using Professor Sen’s algorithm. It will also be an embedded system that solely relies on sound as input. Processing will be done on the device entirely, and it will not interface with the cloud.

# 5 Engineering Requirements

Our final deliverable will follow these engineering specifications:

1. Performs processing without manual input
2. Works in real time using data input from two microphones
3. Operation with at least an hour of battery life
4. Headwear will be less than 1 pound
5. PCB has been designed and manufactured to contain only the necessary circuits
6. Device cost will be < $1000
7. Device will not need to be connected to the Internet
8. Device fits into 10 cm x 10 cm space

In order to validate that our device works, we will be performing a series of experiments to test both objective and subjective validity. Professor Sen has connected us with one of the research managers who will help us design an experiment with people who experience the Cocktail Party Problem. One fellow student who has had difficulty with localizing sound has reached out to us, and we will include them in our experiment. For a more subjective test, we will be passing the processed sound to Google’s Speech to Text API. If words are distinguishable, we can then say that our algorithm works. From these two validations, we can check that the OverEar device solves the cocktail party problem.

# 6 Conclusion

OverEar addresses a weak point in the hearing aid industry by targeting Cocktail Party Deafness [2]. The device is meant to improve the quality of communication of people with hearing conditions by using a biologically-oriented sound segregation algorithm to isolate sound rather than amplifying all sounds.

The device is both portable and comfortable, utilizing dual microphones to replicate the biological systems of humans. This configuration has the added benefit of avoiding existing patents [7]. In addition, our device is made at a lower price point relative to competing products that use much more computationally expensive operations.

**Acknowledgment**

The authors wish to thank Kamal Sen and Kenny Chou for their advice and algorithm contribution.

**References**

1. “Auditory processing disorders,” *Hearing Health Foundation*. [Online]. Available: https://hearinghealthfoundation.org/auditory-processing-disorder-about. [Accessed: 25-Sep-2021].
2. Bronkhorst, Adelbert. (2000). The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. Acta Acustica united with Acustica. 86. 117-128.
3. J. C. Skewes and L. Gebauer, "Brief Report: Suboptimal Auditory Localization in Autism
4. Spectrum Disorder--Support for the Bayesian Account of Sensory Symptoms," *J. Autism Dev. Disord.,* vol. 46, *(7),* pp. 2539-2547, 2016. Available: https://ezproxy.bu.edu/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fscholarly-journals%2Fbrief-report-suboptimal-auditory-localization%2Fdocview%2F1826536013%2Fse-2%3Faccountid%3D9676. DOI: <http://dx.doi.org.ezproxy.bu.edu/10.1007/s10803-016-2774-9>.
5. Orcam, “Hearing Amplification Devices Market Is Disrupted By Orcam,” *StackPath*. [Online]. Available: https://www.orcam.com/en/blog/hearing-amplification-devices-market-is-disrupted-by-orcam/. [Accessed: 17-Oct-2021].
6. A. Al-Heeti, “Orcam Hear will amplify, beam other voices to your hearing aids,” *CNET*, 10-Jan-2020. [Online]. Available: https://www.cnet.com/health/personal-care/orcam-hear-will-amplify-beam-other-voices-to-your-hearing-aids/. [Accessed: 17-Oct-2021].
7. P. Laurie, “US20070223710A1 - hearing aid to solve the 'cocktail party' problem,” *Google Patents*. [Online]. Available: https://patents.google.com/patent/US20070223710A1/en?q=cocktail%2Bparty&oq=cocktail%2Bparty. [Accessed: 17-Oct-2021].
8. K. F. Chou, V. Best, H. S. Colburn, and K. Sen, “A biologically oriented algorithm for spatial sound segregation,” *bioRxiv*, Nov. 2020. [Online]. Available: <https://www.biorxiv.org/content/10.1101/2020.11.04.368548v1.article-info>. [Accessed: 21-Oct-2021].