

A Microphone Array and Voice Algorithm based Smart Hearing Aid

Bharath Sudharsan

Data Science Institute,
National University of Ireland Galway

Manigandan Chockalingam

National University of Ireland, Galway

ABSTRACT

Approximately 6.2% of the world's population (466 million people) suffer from disabling hearing impairment [1]. Hearing impairment impacts negatively on one's education, financial success [2] [3], cognitive development in childhood [4], including increased risk of dementia in older adulthood [5]. Lack of or reduced social interaction due to hearing impairment affects creating or maintaining healthy relationships at home, school and work [5]. Hence, hearing impairment genuinely affects the overall quality of life and wellbeing. The cocktail party effect, which is a healthy hearing individual's ability to understand one voice in a cacophony of other voices or sounds, is an important ability lacking in people with hearing impairment. This inability results in difficulties with simple daily activities such as partaking in group discussions or conversing in noisy restaurants [6]. This smart hearing aid aims to provide much-needed assistance with understanding speech in noisy environments. For example, if a person wants to partake in a group discussion, he/she needs to place the microphone array based unit on a flat surface in front of him/her, such as a table. When conversations take place, the microphone array will capture and process sound from all directions, intelligently prioritise and provide the lead speaker's voice by suppressing unwanted noises, including speeches of other people. This device selects and alternates voices between speakers automatically using voice algorithms. Additionally, the user has the option of further fine-tuning the acoustic parameters as needed through a smartphone interface. This paper describes the development and functions of this new smart hearing aid.

General Terms

Smart Hearing Aid, Advanced Digital Signal Processing, Voice Algorithm

Keywords

Direction of Arrival, Acoustic Echo Cancellation, Single Board Computer, Beam Forming, Noise Suppression

1. INTRODUCTION

The auditory system of people with hearing impairment is unable to isolate speech from noise effectively; thus, requiring an assistive device to do this segregation for the listener. The currently available hearing aids employ many approaches for reducing noise interference and thereby enhancing the desired speech, which is nothing

but different combinations of time and frequency of sound signals. These approaches are good at improving the signal-to-noise ratio and enhance clearer voice hearing in one-to-one communication environments. However, they often fail to deliver or deliver a degraded performance at best in speech-like noisy environments, such as in restaurants. In noisy restaurants where multiple people are speaking simultaneously, there is little to no distinct difference in either the frequency or time of sound source, owing to poor differentiation of noise from the desired speech. This performance of differentiation gets further reduced when the automated detection of speech pause or calculation of modulation index is not accurate. This paper focuses on the design of a circular microphone array-based smart hearing aid, which uses the spatial localisation of voice (both location and direction) along with the time and frequency approaches of sound signal processing. This new smart hearing aid locates and appropriately amplifies a specific voice source amidst a group of voices and noises, using advanced digital signal processing-based voice algorithms. Besides, it provides a user-controlled, real-time Android Graphical User Interface (Android GUI) based tuning to tweak parameters of the voice algorithm running on the microphone-array as the hearing aid user desires based on their listening preferences and environment of the smart hearing aid usage.

1.1 Summary of problem statements:

In 2013, McCormack and Fortnum [7], in their study found that people do not use their prescribed hearing aid due to the continued presence of or inability to control background noise, poor sound or audio quality, lack of volume control, presence of whistling & feedback and inadequate benefits of hearing aid in general. Much of these difficulties continued to be present still due to some of the issues listed below.

- a. Hearing aids currently perform well in close-talking microphone conditions. However, performance degrades when the microphone is far from the voice source. Typically, in a group situation when multiple people are speaking from varied distances due to their mere seating positions from the hearing aid user's microphone, the performance of the currently available hearing aids are not up to par.
- b. Besides, in a noisy place, in an attempt to amplify one person's voice, the currently available hearing aids inadvertently amplify other voice sources and noises.

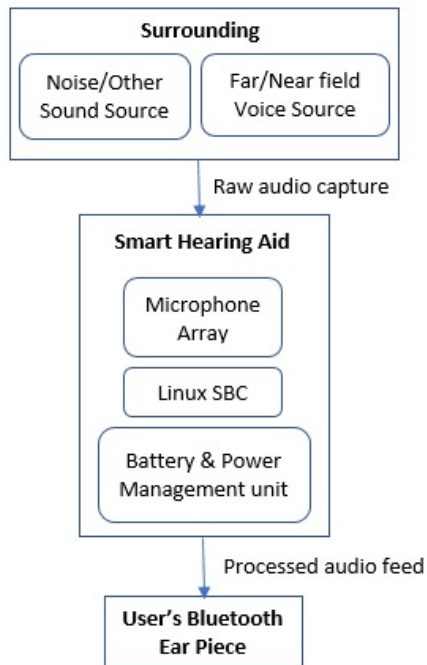


Fig. 1. High-level system diagram of the smart hearing aid prototype

- c. Similarly, in reverberant surroundings, hearing aids tend to amplify late multipath arrivals as well as the direct signal (the actual signal which arrived before the reflected signal)

The hearing aid's inability in locating the voice source, both its distance and its location, is the reason for most of these difficulties listed above.

2. OVERVIEW OF THE SMART HEARING AID

A prototype of a microphone array and voice algorithm based smart hearing aid which is constructed using a Linux Single Board Computer (Linux SBC) interfaced with ReSpeaker V2 [8] and other supporting sub-components as shown in Fig. 1 is developed to overcome the problems mentioned above. As shown in Fig. 2, the user with hearing impairment places the portable hardware on a flat surface surrounding which the speech signals are sourced or picked up from (for example on a group discussion table). Once, people start conversing with the smart hearing aid user, the microphone array processes the sound signals from all directions and produces an output to the user, which is the noise-suppressed speech signal from a specific direction (in other words from a specific person). This specific direction keeps switching automatically, depending on who is talking around the table.

This system segregates speech from noise utilising ultra-compact, omnidirectional digital MEMS microphone array interfaced with a high-performance processor from XMOS to facilitate and run On-Chip advanced Digital Signal Processing (DSP) based Speech algorithms to deliver excellent far-field voice capture, and background noise suppressed audio feed. All these features enhance the listening experience for the user even in the noisiest environments. An Android GUI based acoustic tuning system for the user provides values for parameters of the DSP Speech algorithms running on the

ReSpeaker V2. These values for parameters are tuned according to the user needs, which is based on the environment of smart hearing aid usage. The list of tunable parameters on the Android GUI, their descriptions and their benefits are given in section 2.1. All the parameters use a respective on/off toggle switch to control the feature except the threshold for signal detection in Acoustic Echo Cancellation (AEC). This feature uses a sliding bar to set the desired threshold. In addition to these parameters, the Android GUI has the added option of selecting a specific voice different to the one being amplified and provided to the smart hearing aid user as output based on the automatic speech intensity detection algorithm that usually runs, if needed. This feature is quite useful if the hearing aid user wants to listen to the soft-spoken speaker in the presence of another loud speaking speaker who is speaking simultaneously.

2.1 Tuning parameters list with description and user benefit on activation

- a. **Automatic Gain Control (AGC):** It delivers a near-constant level of speech output irrespective of the voice signal amplitude. Amplifies weak speech and attenuates strong speech, preventing unexpected and undesired extremes of voice amplitudes.
Benefits: The smart hearing aid captures audio signals from locations which are spatially separated. The AGC algorithm running on the microphone array adjusts gain incrementally at a rate corresponding to that of the original input sound signal level to avoid distortion and ensures a constant level of the output signal. Turning on the AGC thus ensures the user neither hears loud sounds too loud nor hears mild sounds too mild.
- b. **Non-stationary noise suppression:** Further suppresses non-stationary noises from the environment, so that the hearing aid user can hear and understand voice much better.
Benefits: Provides boosted reduction of non-stationary noise (sounds such as ambience from a casino, a bunch of cicadas singing, noise when typing on a keyboard, someone eating chips, many motor vehicles passing by, ocean waves) when the user needs to have better access to human voice alone with minimal to no distractions from other meaningful non-stationary sounds.
- c. **Stationary noise suppression:** Further suppresses stationary noises for better hearing of voices.
Benefits: Reduces the stationary noises such as noise from CPU fans, constant buzzing noise from vending machines, HVAC units, etc. further and provides clearer voice output.
- d. **High-pass filter on microphone signals:** Additional filtering of low-frequency sounds that are below 70Hz.
Benefits: Low-frequency sounds, which are mostly ambient noise and rumble from sources such as the humming of electrical mains, air conditioners, furnaces, appliances, etc. (usually have frequencies 50-60 Hz) are actively filtered out to provide better voice hearing, which ranges between 120 and 150Hz, generally.
- e. **Comfort noise insertion:** An advanced noise control mechanism that overcomes the negative effects of total noise removal by allowing comfortable levels of noise.
Benefits: The result of receiving total silence, (complete removal of all background noise due to superior noise reduction) especially for a prolonged period, may have unwanted effects on the listener (user might feel a loss of audio transmission, challenging to understand speech since it might get choppy,



Fig. 2. Smart hearing aid prototype placed on a group discussion table, the dominant pixel from the LED pixel-ring on the device is pointing towards the direction of sound arrival (lead speaker's voice)

jarring due to sudden sound changes). This feature overcomes the unwanted effects mentioned above.

- f. **Transient echo suppression:** Suppresses the transient echoes (short-duration high amplitude sound occurring at the beginning of waveforms of signals) from the hearing aid usage environment

Benefits: Musical sounds, noises or speech and other signals such as glass breaking, newspaper rustling, etc. result in transients causing pre-echo. This transient makes encoding and audio compression difficult. This transient also makes it challenging to process critical speech information and thus negatively impacts the voice legibility. Turning this feature on provides a smooth voice listening.

- g. **The threshold for signal detection in AEC:** AEC aims to remove echoes, reverberation, and unwanted sounds from the captured audio signals by using Digital Signal Processing techniques

Benefits: When the mic array is used in a noisy environment, e.g. malls, tuning based on user input from android GUI will be used to increase the voice-capture accuracy. The background sound/music will be considered as silence and be ignored if the background sound level captured is less than the Threshold AEC level set by the user

3. FUNCTIONING OF THE SMART HEARING AID PROTOTYPE

3.1 Linux SBC utilisation in smart hearing aid

As illustrated in Fig. 3, python script written leveraging external libraries is deployed in the Linux SBC. This script provides commands for tuning parameters of the voice algorithm running on the ReSpeaker V2 based on the user's input through the Android GUI, when available. It then gathers the finally processed audio signal from the microphone array via USB and streams to the user's ear-piece via Bluetooth.

3.2 ReSpeaker V2 utilisation in smart hearing aid

The firmware on the XVF-3000 Chip from XMOS produces 6 Channel mic outputs via USB to the Linux SBC. Channel zero contains audio, which is processed using advanced DSP algorithms. The channels one to four contains raw data from the respective four microphones built-in to the device. Channel five provides raw audio, which is a combination of all raw audio signals from 4 microphones on the ReSpeaker V2.

4. RESULTS

The interaction with the smart hearing aid prototype was performed in different environments as described below

The sitting room at home: As illustrated in Fig. 1, the hearing aid user was seated on a couch conversing with other people seated around and watching television in low volume. The television was four meters away from the prototype and was running in the background during conversation. This room was relatively quiet too.

University cafeteria: At the time of testing, the cafeteria had approximately 10 people seated at different tables. There were multiple stationary (Noise from vending & coffee machines, blowing noise from HVAC's, etc.), non-stationary noise sources (people walking, outdoor noise from windows, etc.) and multiple voice sources. At the table on which the device user was seated, there were five people, all of them having conversations both with the hearing aid user and among themselves. At times two or more people were talking simultaneously (for example, one was engaged in conversation with the hearing aid user, and a couple of others were engaged in conversations among themselves).

Lecture hall: The testing was done in a small lecture hall with a capacity of 40 students. In this simulated lecture session, the hearing aid user was seated in the last row of seats approximately five meters from the lecturer, who was the only source of the voice in a relatively quiet room. The lecturer was constantly on the move all

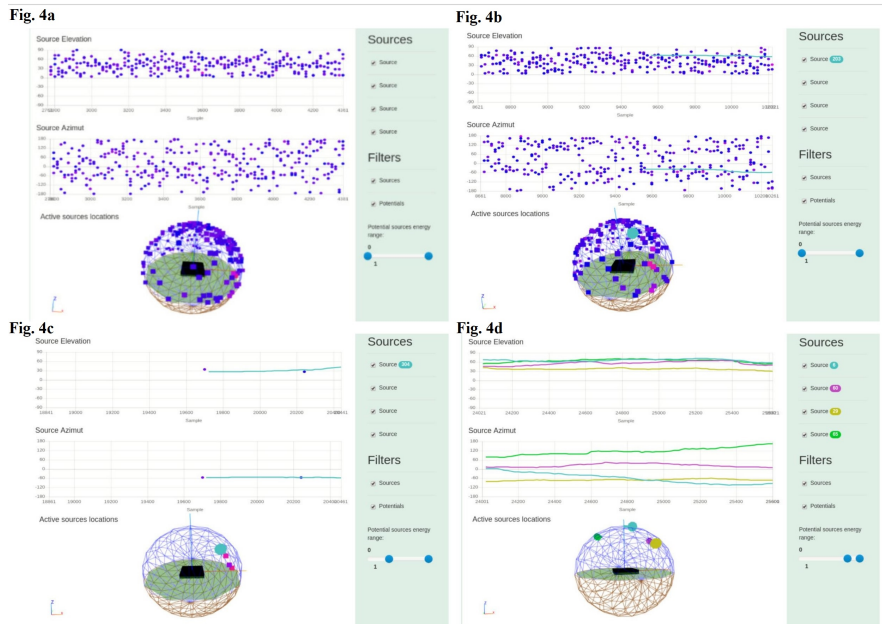


Fig. 4. The representation of sound sources in the university cafeteria testing environment on a unit sphere. Fig. 4a - Raw undifferentiated sounds as captured by the microphones. Fig. 4b - The processed sounds with one identified voice source of interest along with other noises. Fig. 4c - The processed sounds with one identified voice source of interest and other noises suppressed. Fig. 4d - The noise suppressed sounds with four identified voice source of interest, providing the user with Android GUI based select-able option of preferred voice source.

through the session, pacing from one end to the other.

A library named ODAS (Open embeddeD Audition System) [9] was utilized to visualize, localize and track sound to give a better picture of the sound sources in the cafeteria, which are shown in Fig. 4a-d. All the unprocessed sounds consisting of voice sources, along with noise sources captured from the cafeteria, is illustrated on a unit sphere using ODAS studio in Fig. 4a. As seen in this figure, the numerous undifferentiated sound sources are from different

directions and also from different distances in relation to the microphones which are being depicted as blue and purple squares in the unit sphere. Fig. 4b shows the processed sounds which differentiate one voice of interest from other noises, and Fig. 4c shows the same voice source after noise suppression. When there are multiple voices of interest, the device shows all the relevant voices and provides the option of choosing a specific voice by the user as needed, using the Android GUI. This is shown in Fig. 4d, where there are four voices of interests being shown after effective noise suppression.

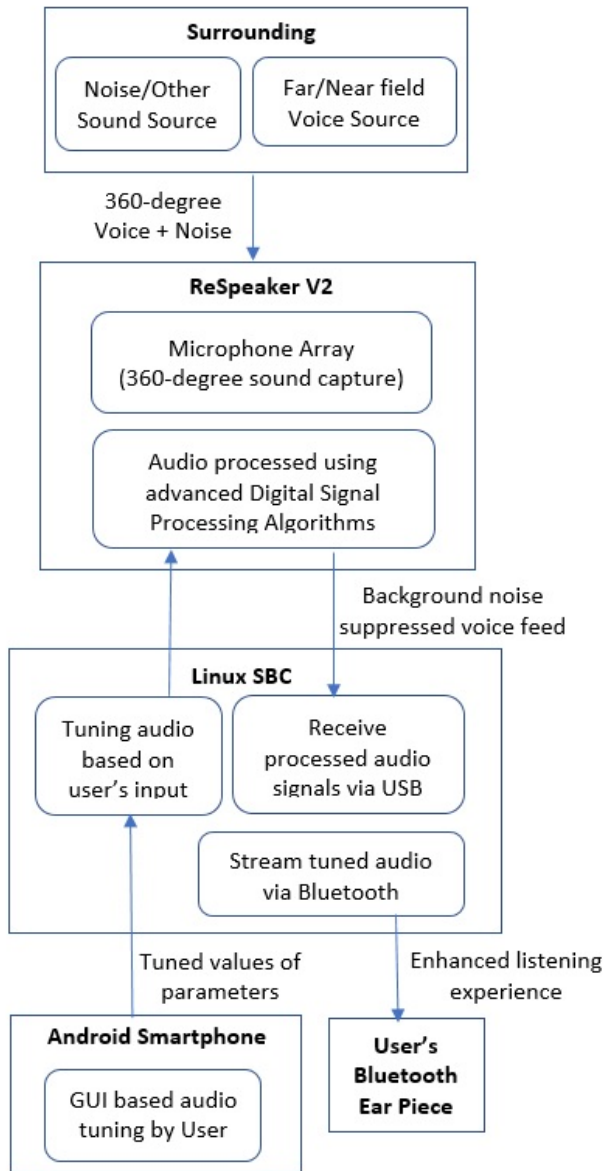


Fig. 3. System flow diagram of the smart hearing aid prototype

sion.

The far-field voice capture and the voice tracking capability (beamforming) of the smart hearing aid was tested in a simulated lecture hall as shown in Fig. 5a and Fig. 5b. The Fig. 5a shows the lecturer's voice source from a distance of approximately five meters from the left end of the lecture hall while Fig. 5b shows the lecturer speaking from the right end of the lecture hall. During both instances it is to be noted that the device's LED pixel lights up (dominant light blue pixel) according to the lecturer's position in relation to the device. The same information is being captured and illustrated in the associated unit sphere graphical representation.

4.1 Advantages of the smart hearing aid:

- Far-field voice capture:** The smart hearing aid has the capability of capturing and processing raw microphone inputs from distances as far as up to five meters. As shown in Fig. 5, lecturer's voice is captured and tracked from last row of seats approximately five meters from the lecturer. This feature makes it easy for the hearing aid user to listen to lectures.
- Lead speaker visual localisation:** There are 12 RGB LED indicators on the Respeaker V2 arranged in the form of a LED pixel-ring. As illustrated in Fig. 1, the visual feedback indicating the direction of speech signal arrival (source) through appropriate lighting up of a specific LED from the pixel-ring on the ReSpeaker V2 makes it easy to locate which speaker's voice is being heard especially when multiple people are talking in a group.
- Beamforming:** All MEMS microphones have an omnidirectional pickup response. It means, their response is the same for sound coming anywhere from around the microphone. Directional response or a beam pattern can be formed by configuring multiple microphones in an array. Thus, enabling us to detect and track the position of the voice across the room. As the speaker interacts with smart hearing aid user while walking around the room, the angle of the microphone beam adjusts automatically to track their voice. Hence, it is effectively possible to point towards the user's direction and suppress noise or reverberation signals from other directions. Beamforming feature of this hearing aid is illustrated in Fig. 5 by capturing and tracking lecturer's voice from the Left end to right end of the lecture hall. This feature makes it easy for the hearing aid user to listen to lectures
- De-reverberation & noise suppression:** In acoustic beamforming, the spatial relationship of the microphone array achieves active microphone noise control and suppression. Since the direction of sound arrival relative to the microphone array is known, the acoustic beamformer passes signals coming from the sound source in a particular direction and filters out signals coming from other directions. This approach is most applicable in situations where one person's voice needs to be heard when multiple people are talking. In few environments, the lead speaker's voice may reverberate (reflect) off hard surfaces around the room, e.g. a window or TV screen (the sitting room at home as shown in Fig. 1). The De-reverberation feature of the advance DSP algorithm on the Mic array removes these reflections and cleans up the voice signal
- Acoustic echo cancellation:** While interacting with audio signals and electronics devices, in some cases, users hear their voice (sometimes with a significant delay) this experience is known as an acoustic echo. Controlling and cancelling acoustic echo is essential for voice-based systems such as ours. For example, if the smart hearing aid user is watching a film on a TV with minimal volume and simultaneously if a person from the same room starts an interaction with the hearing aid user, then the microphones will capture both user's voice and the sound of the film (the acoustic echo). This acoustic echo needs to be cancelled from the voice input so that better listening experience is provided to the user

5. CONCLUSION

The enhanced listening experience is achieved through hardware and software-based techniques used in our system for detecting and differentiating voices from other non-voice sounds. The direction

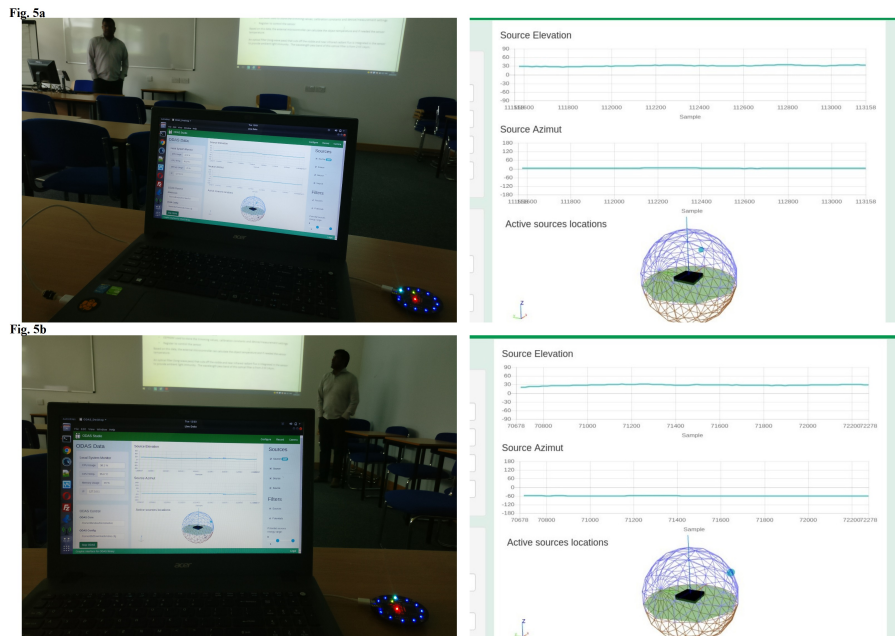


Fig. 5. Illustration of far-field voice capture and the voice tracking capability (beamforming) in a simulated lecture hall. Fig. 5a illustrates the direction of lecturer's voice captured and tracked from the left end of the hall and Fig. 5b illustrates direction of lecturer's voice captured and tracked to the right end of the lecture hall

of arrival feature of microphone array used in this prototype provides information about the location of voice source and automatically redirects and amplifies the microphone data received from that voice source allowing lesser straining by the hearing-impaired client to focus on specific voices. This direction of arrival feature of the microphone array also provides relative distance mapping of voice origin. Thus, the processed voice heard by the hearing-impaired person is not only clear voices of different persons but also, they are heard as an appropriate spatially oriented voice relative to the position of the person in the group to that of the person with hearing impairment.

6. REFERENCES

- [1] "Deafness and hearing loss," 2019.
- [2] P. Home, K. Kvrner, J. R. Casey, R. A. M. J. Damoiseaux, T. M. A. van Dongen, H. Gunasekera, R. G. Jensen, E. Kvestad, P. S. Morris, and H. M. Weinreich, "Panel 1: Epidemiology and diagnosis," *OtolaryngologyHead and Neck Surgery*, vol. 156, pp. S1–S21, 2017.
- [3] D. S. Dalton, K. J. Cruickshanks, B. E. K. Klein, R. Klein, T. L. Wiley, and D. M. Nondahl, "The impact of hearing loss on quality of life in older adults," *The Gerontologist*, vol. 43, pp. 661–8, 2003.
- [4] M. T. Jaiyeola and A. A. Adeyemo, "Quality of life of deaf and hard of hearing students in ibadan metropolis, nigeria," *PLOS ONE*, vol. 13, p. e0190130, 2018.
- [5] M. Cherko, L. Hickson, and M. Bhutta, "Auditory deprivation and health in the elderly," *Maturitas*, vol. 88, pp. 52–57, 2016.
- [6] B. Williger and F. R. Lang, "Hearing aid use in everyday life: Managing contextual variability," *Gerontology*, vol. 61, pp. 158–165, 2014.
- [7] A. McCormack and H. Fortnum, "Why do people fitted with hearing aids not wear them?" *International Journal of Audiology*, vol. 52, pp. 360–368, 2013.
- [8] Bill, "Respeaker mic array v2.0 - seed wiki," 2016.
- [9] introlab, "introlab/odas," 2019.