Australian National University

ADVANCED 3D AUDIO CAPTURING MICROPHONE ARRAY

TECHNICAL SUMMARY

Invention snapshot

> ANU researchers have developed a new type of microphone array capable of recording 3D spatial sound, using a planar geometry.

> Provides audio signal streams in the form of 'Higher Order Ambisonics' and is compatible with most spatial sound rendering engines, is compact and can be integrated into a number of miniaturized spatial audio applications, and can provide spatial hearing capability to hearing aid devices and autonomous systems.

ANU Technology Transfer Office

Innovation ANU

Background

Spatial 3D audio is not a new concept, however, historically, high quality 3D audio capture and reproduction has been difficult, expensive, and cumbersome, to achieve. Capture and reproduction of 3D audio is becoming increasingly important for many applications including augmented reality (AR), virtual reality (VR), media, digital entertainment, human-machine communications/interfaces, smart homes, hearables, wearables, medical aids, teleconferencing and active noise control within confined spaces[1]. A signal processing technique formerly called spherical harmonic analysis and synthesis (now commonly known as 'Higher Order Ambisonics') is the preferred way of representing and coding spatial 3D audio. The process wherein you capture natural 3D audio and decompose it into 'spherical harmonics' is not straightforward and typically requires a large number of microphones arranged on a symmetric 3D surface. Currently, binaural, hemispherical, or spherical microphone arrays are used to capture 3D sound (Figure 1). Binaural arrays consist of dual microphones that are typically shaped and positioned like human ears. Hemispherical arrays typically involve 2-14 microphone capsules arranged in

Figure 1: Examples of spherical (Eigenmike), hemispherical (AMBEO) and binaural microphone arrays currently used to capture 3D audio.

a tetrahedral pattern, whilst spherical arrays incorporate numerous (e.g. 24-32) pressure microphones which are mounted on the surface of a rigid spherical-shaped baffle (Figure 1). The spherical array geometry has several advantages over the other geometries, e.g. the polar pattern can be directed to any direction in the 3D space without changing the shape of the pattern, plus the spherical array allows full 3D control of the polar pattern and spatial-filtering (such as beamforming) in 3D soundfields. The downside, however, is that this geometry is bulky in nature, which makes this type of array non-portable and inconvenient for practical use. Spherical arrays are widely used for research and experimental purposes, but the size greatly limits the arrays' potential in commercial applications. Researchers from The Australian National University (ANU), in Canberra, Australia, are seeking to address this limitation by developing a microphone array with reduced dimensionality and size that will still offer the same functionality and benefits of the spherical microphone array.

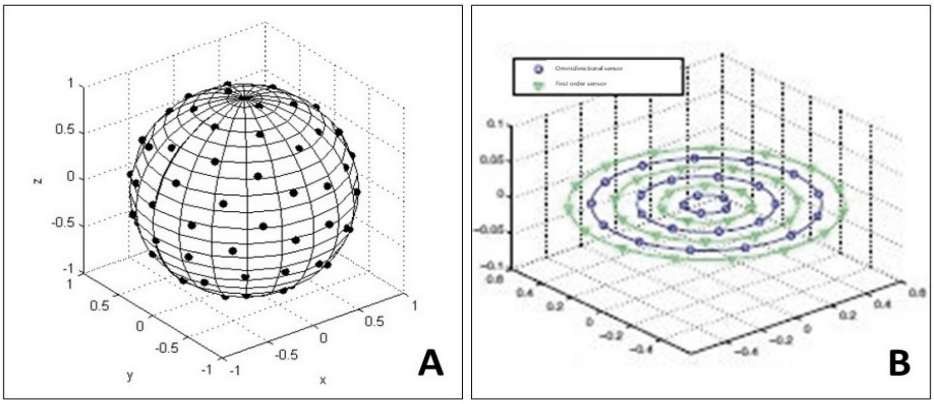

Figure 2: (A) Conventional spherical microphone array (3D configuration); (B) Circular-ring, two sensor planar microphone array (2D configuration)

3D audio engages the listener by offering a spatial bearing that enables them to sense where they are relative to the noises around them. In 3D soundscape, the origins of sounds can perceptibly move about the listener, locating the listener as if they were standing in a real life environment. By inserting 3D audio, new spatial information is introduced, enabling audiences to sense things happening between them, or elsewhere in their environment, completely independent of their eyes. Manipulating this type of audible sensory perception has the potential to completely reshape the entire listening experience.

There are many applications and approaches, but the ultimate goal in the VR, music, film and gaming industries is to deliver a truly immersive experience to the end-user[1]. Companies are working hard to develop the visual technology needed to provide these experiences, but there is still one major piece missing from the puzzle – audio. For instance, a personal and mobile VR experience inherently requires a pair of headphones for audio output, and delivering 3D audio through that 2-channel format (instead of layers of loudspeakers) is no easy task. This is why binaural rendering plays such a critical role in the VR listening experience. This signal processing technology synthesises 3D sound scenes comprised of object, channel and/or ambisonics into 2-channel outputs[1]. What is even more challenging with VR is the audio has to be interactive 3D audio, i.e. it should take into account the actions of the user, not just the sound source's movements in the scene[1]. In less than five-years, 3D spatial audio is expected to revolutionise our standard for multimedia listening[1]. Because of the growing interest for accurately capturing 3D sound, microphone industries will likely go through major changes[1]. The quality of evolving equipment like 'Ambisonics microphones' stills leaves a lot to be desired, highlighting the need for more research and development (R&D). That same equipment is still very expensive, and there are already efforts to disrupt this marketing utilising more efficient technological advancements[1].

Technology description

ANU researchers have developed a novel, compact microphone array arrangement and custom DSP algorithms for enhanced 3D sound capture and delivery. Through the use of directional microphones, the researchers have been able to significantly reduce the size of the array, as compared to conventional spherical 3D arrays (Figure 2A), which have similar functionality.

The new array configuration consists of multiple co-centred circular rings, with two sensor types (directional and omni-directional microphones) placed on a 2D plane (Figure 2B). Vertical alignment along the circular rings as well as perpendicular placement to the omni-directional microphones enables the directional microphones (also known as first order sensors) to measure the pressure gradient and vertical harmonic components of an impinging soundfield. The omni-directional microphones (also known as omni-directional sensors) which are also placed along the circular rings (Figure 2B) are used to measure the sound pressure and horizontal harmonic components of an impinging soundfield. The two sensor types working together in this 2D arrangement, enables the microphone array to extract full 3D soundfield information. This system exploits a special property of the Legendre functions (which represents the soundfield) and uses a combination of omni-directional microphone units to achieve the full functionality of a spherical microphone array in a very compact, planar form factor.

Custom developed algorithms associated with the array produce spatial audio signal streams in the form of 'Higher Order Ambisonics' which is compatible with the latest industrial standard for spatial sound encoding (MPEG-H) as well as the spatial sound format commonly used for YouTube VR contents. Thus, the array is compatible with most spatial sound rendering engines on the market and its compact size makes it incorporable into various consumer electronics. Furthermore, these innovative patent-protected DSP algorithms are used to separate (i.e. decompose) an incoming 3D soundfield into spherical harmonic components. Such decomposition enables the array to perform a plethora of sound refining functions (e.g. sound source location, adaptive beamforming and noise reduction) in order to suppress background noise and enhance the speech signal in order to deliver reliable, clear, crisp audio delivery.

To date, ANU researchers have developed an alpha prototype (Figure 3A). The prototype is of a modular design, wherein individual microphone units can be flexibly attached and removed from a motherboard, which determines the microphone array dimension. This allows convenient prototyping and testing of various array configurations, but its form factor is suboptimal due to the necessity to include connection sockets and plugs to facilitate the modular design. The prototype was used (at WASPAA 17) to demonstrate a binaural reproduction of the array technology. Currently, the team is exploring the use of the array in desktop microphone (Figure 3B) for hearing aid users. The intended end-product is a compact, wearable (or desktop) array that captures the full spatial sound information around the hearing aid user, then a dedicated binaural rendering algorithm processes the captured ambisonic signals to generate two channels of audio output, which are sent to each of the user's hearing aids. The binaural rending algorithm is able to simulate the sound propagation around the user's head, such that the user can benefit from the spatial information of sound just like normal people do. In addition, the intention is to implement advanced array signal processing algorithms for the array (such as beamforming, noise suppression and echo cancellation) to further enhance the hearing aids capability at picking up a desired speech signal, in a noisy background environment.

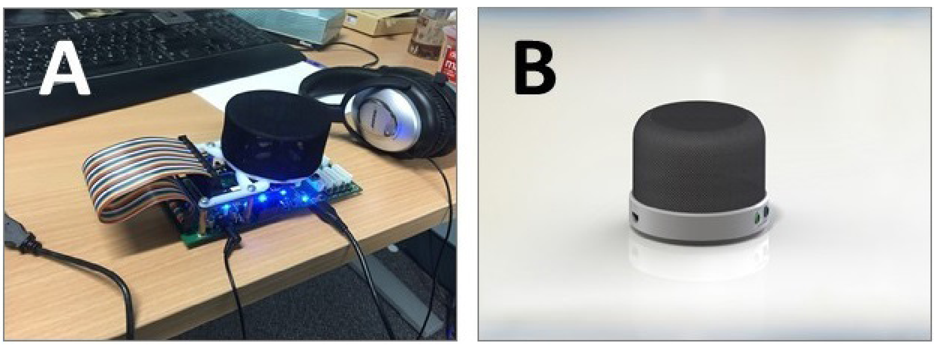

Figure 3: (A) Current alpha prototype of planar microphone array; (B) Proposed beta prototype of planar microphone array

Advantages

- *Compact:*
  - The compact, planar array configuration enables easier integration into consumer electronics, broadening the use of 3D sound recording and analysis.
- *High quality 3D audio:*
  - Removes unwanted background noise (using beamforming) and automatically focuses on the desired sound/speech source (using adaptive steering).
  - Enhanced spatial resolution, broader polar pattern width and advanced DSP algorithms means that the array can provide reliable, clear, crisp audio delivery.
  - Clear voice capture – noise reduction via: (i) spatial filtering, (ii) echo cancellation and noise suppression, (iii) electronic wind-noise reduction, and (iv) multichannel noise reduction processing.
- *Adaptable:*
  - Due to the modular nature of the design, there is the ability to change the quality of the components and thus the functionality of the final product.
- *Customisable:*
  - Array can be tailored to a range of devices and applications.
- *Affordable:*
  - Inexpensive compared to current microphone arrays with 3D recording capability.
- *User-friendly:*
  - Plug and play - using USB or 3.5 mm microphone/audio jack.
  - Portable – compact, flat and lightweight.

Applications and market

The global 3D audio market is expected to reach US$14.5 billion by 2026, expanding at a compound annual growth rate (CAGR) of 16.6% from 2018 to 2026[1]. Increasing adoption of 3D audio in gaming is the major factor driving the 3D audio market. Game developing companies are replacing their traditional 2D audio sound engines with advanced 3D audio sound engines and are incorporating advanced DSP algorithms that can be integrated directly into a game's sound engine[1]. Over the years, cinema screens have transformed from analogue to digital and the current market trend is adoption of 3D audio sounds systems. Furthermore, booming VR and AR and its penetration across museums, exhibitions and personalised consumer devices, is anticipated to create lucrative market opportunity for 3D audio[1]. Technological developments in the music industry and the automotive industry's demand for 3D audio enabled advanced driver assistance systems, are also expected to show growth opportunity for the 3D audio market[1].

A hearable is a wireless device or mechanism that fits into or above the ear, which amplifies or plays audio from a remote source. In other words, a hearable can be defined as a combination of audio devices and wearable technology. With technological advancement, these devices are also being used for various audio processing and hearing aid applications. The overall hearable devices market was valued at US$12.1 billion in 2016 and is expected to reach $23.2 billion by 2023, at a

CAGR of 10.0%[2]. There is an increasing demand for advanced, wireless and smart hearing devices across various industries such as consumer electronics, commercial, sports, gaming, healthcare and others[2]. The major drivers for this market including growing demand for smartphones as a source of entertainment, consumer preference for wearable and portable devices, increasing demand for health monitoring applications and hearing aids, and high investment for R&D for hearables by OEMs[2]. Major upcoming/trending technologies and products in this market include advancements in voice user interface and emerging hearable computers[2]. New platforms gaining traction for human-computer interaction are wearables, hearables, AR, VR, voice UI and others[2].

The hearing aids market is expected to reach US$9.8 billion by 2022 from $7.0 billion in 2017, growing at a CAGR of 7.0%[3]. The major factors responsible for the growth of this market are high prevalence of hearing loss, rising geriatric population, and technological advancements in hearing aids[3]. Technology advanced products that offer better quality of hearing, improved sound processing, and wireless options are witnessing high market demand[3]. Currently, many companies operating in the hearing aids market are investing in R&D to develop advanced hearing aids with improved performance and efficiency. Latest technologies being incorporated include: (i) 'smart hearing aids' that are capable of connecting with other media devices such as TVs, phones and computers[2], and (ii) 'self-learning' and 'self-programming' hearing aid applications that detect the wearer's preferences and adapt automatically[2], plus (iii) specially designed sensors, transmitters and monitors that alert the wearer in critical situations, thereby enhancing the safety of the wearer[2].

Opportunity

The planar microphone array technology is suitable for a range of devices and applications and could benefit a variety of company types – as such it may present multiple potential licensing opportunities.

ANU is interested in identifying industry partners that have a need for this technology, in order to create a number of functional prototypes. ANU is well-placed to work with partners to optimise an array for their specific application.

Patent status

The unique geometry design and accompanying proprietary operating algorithms for the microphone array is owned by ANU and has a priority date of 23/07/2014. National Phase filing is ongoing. ANU is seeking protection in Australia, Canada, China, Europe and Japan. The US patent has been granted (US9949033B2).

Scientific team

Prof Thushara Abhayapala:

Prof Ahbayapala is well-known among industry and academic audio groups. He is credited for proposing 'Higher Order Ambisonics' (now a default in 3D audio standards (MPEG-H)) and for first proposing higher order spherical microphone arrays. Currently, he has projects with SONY and Dolby and is in-discussion with Facebook and Zoox. In the past, he was the Director of the Research School of Engineering at ANU and was the Deputy Dean for the ANU College of Engineering and Computer Science. He is one of the inventors of the underlying technology and brings strong theoretical and experimental expertise to the team, particularly in the fields of spatial soundfield recording, beamforming and array signal processing.

Dr Prasanga Samarasinghe:

Dr Samarasinghe is an expert in audio and acoustics. She was an intern at Dolby while doing her PhD. Currently, she leads an ARC Linkage project with Dolby and was recently hired as a prestigious FERL (Future Engineering Research Leadership Fellow) by the Research School of Engineering at ANU. She has contributed to the development of alpha prototype and is experienced in the areas of spatial audio, array signal processing, virtual acoustics and surround sound systems, sound source localisation, and active noise cancellation.

Mr Fahim Abdullah:

Mr Abdullah is a final year PhD student with prior industry experiences. Recently, he interned at Apple (audio research group) in Silicon Valley and has excellent coding skills in addition to expertise in acoustic signal processing. Fahim is playing a key role in the development and coding of the beamforming and source localisation algorithm in the current prototype.

Contact

Dr Fiona Nelms

Director, Technology Transfer Office

T +61 2 6125 9187
E fiona.nelms@anu.edu.au

www.anu.edu.au/research/innovation

References

[1]Transparency Market Research (2016) 3D audio market – global industry analysis, size, share, growth, trends and forecast 2017-2025. Available at https://www.transparencymarketresearch.com/3d-audio-market.html. Accessed Aug-18

[2]MarketsandMarkets (2015) Hearable devices market by product (headset, headphone and hearing aids), technology (Bluetooth, WiFi), application (consumer, healthcare) and geography – global forecast to 2023, 154 pp. Accessed Aug-18.

[3]MarketsandMarkets (2016) Hearing aids market by product (receiver in ear, behind the ear, in the ear, in the canal hearing aids, cochlear implant, BAHA implant), types of hearing loss (sensorineural, conductive hearing loss) and patient (adult, paediatric) – forecast to 2022, 188 pp. Accessed Aug-18.