

Poster: Networked Acoustics Around Human Ears

Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, Romit Roy Choudhury

University of Illinois at Urbana-Champaign

{ssh19, nroy8, jguan8, haitham, croy}@illinois.edu

ABSTRACT

Ear devices, such as noise-canceling headphones and hearing aids, have dramatically changed the way we listen to the outside world. We re-envision this area by combining wireless communication with acoustics. The core idea is to scatter IoT devices in the environment that listen to ambient sound and forward it over their wireless radio. Since wireless signals travel much faster than sound, the ear-device receives the sound much earlier than its actual arrival. This “glimpse” into the future allows sufficient time for acoustic digital processing, serving as a valuable opportunity for various signal processing and machine learning applications. We believe this will enable a digital app store around human ears.

CCS CONCEPTS

• **Networks** → *Sensor networks*; • **Human-centered computing** → *Ubiquitous and mobile devices*;

KEYWORDS

Acoustics, Wireless, Internet of Things, Wearables, Earphone

ACM Reference Format:

Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, Romit Roy Choudhury. 2018. Poster: Networked Acoustics Around Human Ears. In *The 24th Annual International Conference on Mobile Computing and Networking (MobiCom '18)*, October 29–November 2, 2018, New Delhi, India. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3241539.3267760>

1 INTRODUCTION

Consider an office environment where Alice gets disturbed in her cubicle by loud corridor conversations (Figure 1). Imagine a small IoT device – equipped with a microphone and wireless radio – mounted on the door in Alice’s office. The IoT device listens to the ambient sounds (via the microphone)

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MobiCom '18, October 29–November 2, 2018, New Delhi, India

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5903-0/18/10.

<https://doi.org/10.1145/3241539.3267760>

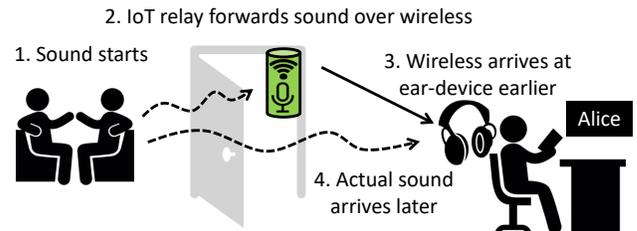


Figure 1: An intuitive example of networked noise cancellation.

and broadcasts the exact sound waveform over the wireless radio. Now, given that wireless signals travel much faster than sound, Alice’s noise-canceling headphone receives the wireless signal first, extracts the sound waveform from it, and gains a “*future lookahead*” into the actual sound that will arrive later. When the actual sound arrives, the headphone is already aware of the signal, and has had sufficient time to produce the appropriate “anti-noise” signal, canceling the noise at Alice’s ear. Viewed differently, the wireless network serves as a “fast network backbone” that delivers the sound information before it actually arrives at the receiving device, laying the foundations for *networked* noise cancellation.

By contrast, in today’s noise-canceling headphones (from Bose, Sony, etc.), the sound signal arrives at the microphone (located at the headphone) and the human ears almost simultaneously. As a result, the headphone hardly has time – less than $30\mu\text{s}$ [2] – to process the sound and produce the appropriate anti-noise signal. As shown later, the tight time budget limits the headphone’s ability to estimate the channel and produce the anti-noise in time, leading to unsatisfactory performance in noise cancellation. Our system provides 100X longer time budget instead, leading to more accurate anti-noise signal and better cancellation quality.

Noise cancellation is only one of the many applications that benefit from this *networked* architecture. Now that sound is available to the ear-device earlier than actual arrival, complex digital processing (as opposed to analog) and machine learning algorithms can now have sufficient time to execute, enabling an acoustic digital app store around human ears.

This poster summarizes and extends our recent work in ACM SIGCOMM’18 [5]. We start with a glimpse of today’s ear-device processing pipeline; then introduce our networked architecture; and finally zoom into a few applications.

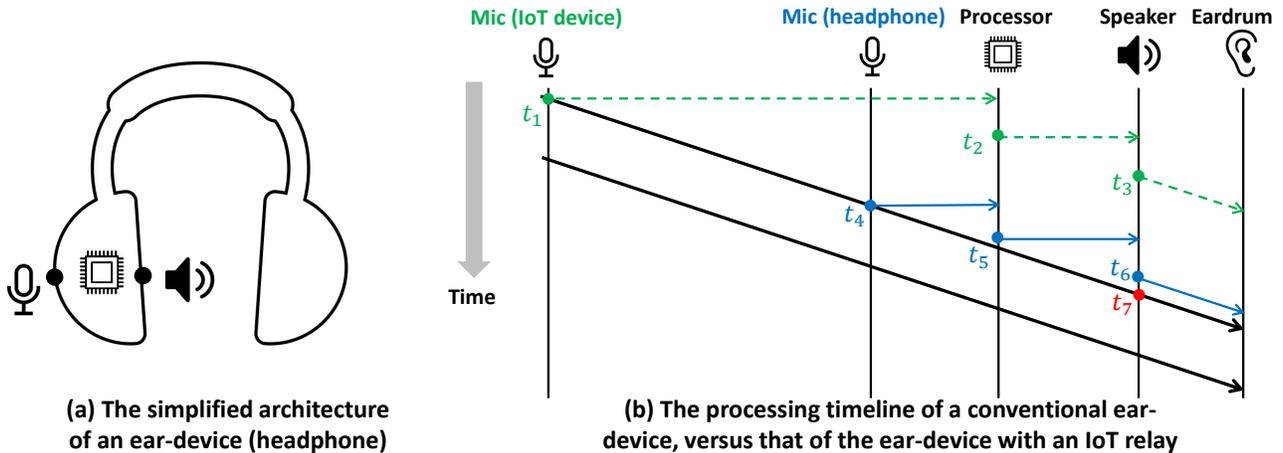


Figure 2: (a) An ear-device processes the sound from its microphone and sends its output to the speaker. (b) A microphone on the IoT device enables the ear-device to hear the sound way earlier.

2 TODAY'S EAR-DEVICE PIPELINE

Today's ear devices follow a similar architecture, as shown in Figure 2(a) – it has one or more microphones on the outer edge, a processor (usually a DSP) inside the device, and a speaker that is located closer to human eardrum. The processor's job is to take the microphone's signal as input, apply appropriate signal processing techniques, and finally output the sound through the speaker, so that the human will hear a manipulated version of the outside sound.

Figure 2(b) shows the timeline of this operation, in which the time flows in the vertical direction. For today's ear devices (blue solid lines), the outside sound first arrives at the microphone at time t_4 . The microphone sends the electrical signal to the DSP, which performs the following three steps: (1) sampling the signal, incurring an ADC (analog-to-digital conversion) delay; (2) computing the appropriate output signal, incurring a DSP computational delay; (3) converting the output signal to analog and sending it to the speaker (at time t_5), incurring a DAC (digital-to-analog conversion) delay. Finally, the speaker plays the sound at time t_6 after a speaker playback delay. In sum, the overall delay ($t_6 - t_4$) is:

$$\text{Overall Delay} = \text{Delay at \{ADC + DSP + DAC + Speaker\}}$$

However, at time t_7 , the actual sound (in the air) has already arrived at the speaker's position. The ideal case is to have t_6 no later than t_7 – a challenging task for the headphones. This is because ($t_7 - t_4$) is only tens of microseconds, due to the ear-device's small form factor. This is the reason why noise-canceling headphones all use dedicated audio hardware and DSPs, trying to minimize the overall delay.

3 OUR DESIGN

Our design relaxes this constraint. Figure 2(b) also shows the pipeline of our design (green dashed lines), where we

connect the ear-device with IoT relay devices wirelessly. We are now able to obtain the sound sample as early as time t_1 (as opposed to t_4), way before its actual arrival. Imagine the IoT relay is pushed 1 m away – we will have $t_7 - t_1 \approx 3$ ms, which is easily 100X larger than that in conventional design. This offers a much-needed time cushion for digital computation.

Figure 3 shows the relay hardware, which is designed in analog to avoid delays from digitization and processing. The relay uses frequency modulation (FM) in the 900 MHz ISM band. When connected with multiple IoT relays in different directions, the DSP in the ear-device uses the GCC-PHAT cross-correlation algorithm to determine which relay to use (i.e., which relay is closest to the sound source and offers maximum “glimpse” into the future).

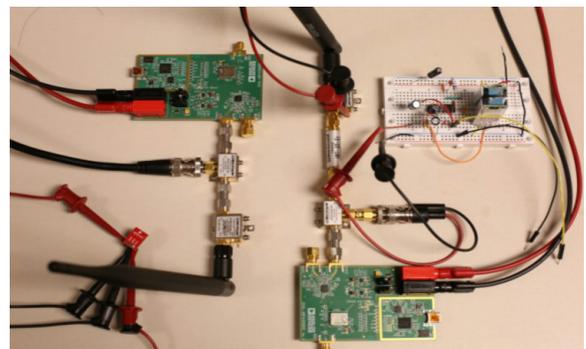


Figure 3: The analog relay hardware (TX and RX) to piggyback sound on RF.

4 EXAMPLE APPLICATIONS

Below we show a few example applications that can be enabled or improved, with our proposed networked architecture.

■ Noise Cancellation

Noise cancellation benefits from this networked architecture, for at least the following two reasons:

- (1) Timing: The DSP processor will now have sufficient time to complete the computation and play the anti-noise in time (before t_7). This is especially important for higher frequencies, in which any small time delay will incur a large phase mismatch between noise and anti-noise.
- (2) Signal Processing: Generating the anti-noise is much more than simply flipping the microphone signal to get its opposite; the headphone needs to estimate the channel from the sound source to the speaker, as well as the *inverse* of the channel from the sound source to its microphone, in order to correctly cancel the noise. This *inverse* operation calls for non-causal filtering, where a much longer “glimpse” into the future sound significantly improves the anti-noise computation, improving the core of noise cancellation.

We design algorithms which utilize these benefits and run on a general-purpose DSP board (detailed in [5]). Figure 4 shows the performance of our system, compared to the Bose QC35 noise-canceling headphone. For our system, we connect our DSP board with the Bose headphone while turning its power off, so that we use Bose’s hardware but run our algorithms. On average, we outperform Bose’s performance by 8.9 dB.

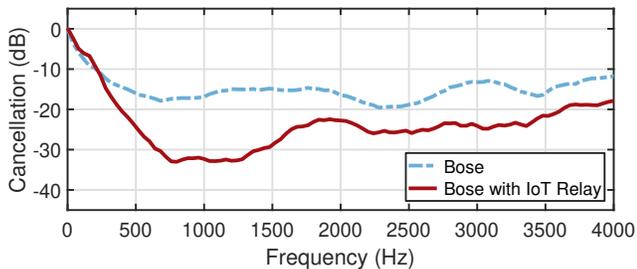


Figure 4: Bose’s noise canceling performance improves with an IoT relay.

■ Speech Enhancement

It is particularly challenging for people with hearing loss to understand conversations in noisy environments, such as in a restaurant, even with hearing aids [1]. By putting an IoT relay on the table, the speech enhancement will benefit because of:

- (1) Minimized Delay: The delay of the output will be minimized, making the sound more natural for users. This is because the user hears the sound twice – first from the air, and then from the hearing aid. Studies have shown that a few milliseconds of delay that exists in digital hearing aids can create noticeable distortion to the sound [3]. Moreover, many sophisticated speech enhancement algorithms failed

to run on today’s hearing aids simply because of long processing time [4]. The networked architecture offers more time budget, making these algorithms affordable.

- (2) Microphone Array Processing: Although microphone arrays are powerful in noisy and reverberant environments (that are especially challenging for hearing aids), they are too large to be equipped on hearing aids which lack spatial diversity. They can now, however, be equipped on a separate IoT device, and deliver high-quality speech signals to the user wirelessly.

■ Digital Equalizer (EQ)

Finally, as one example of acoustic augmented reality, real-time user-configurable equalization for outside sound should now be possible. Imagine a guitarist playing in a concert may choose to reduce the low-frequency sound from the drummer; the elder with age-related hearing loss can choose to enhance the high-frequency sound in day-to-day circumstances; or a normal user wants to add the bass sound when listening to live music. With the sound being delivered to the ear-device much earlier, we envision all these can be configured on-the-fly, simply using a phone app.

As a summary, the wireless signals are playing the role of a control plane, while the acoustic signals are the data plane (Figure 5), essentially because wireless signals travel much faster than acoustics. This section only lists a few applications, but we envision many others should be possible.

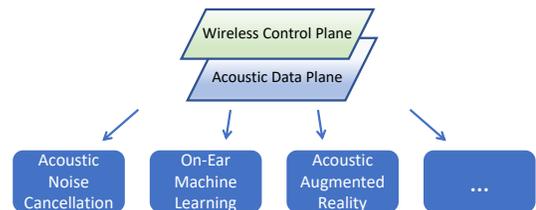


Figure 5: Wireless signals are playing the role of a control plane, and the acoustic signals are the data plane.

REFERENCES

- [1] 2018. Hearables Challenge. (2018). Retrieved Aug 2, 2018 from <https://www.challenge.gov/challenge/hearables-challenge/>
- [2] 2018. Here Active Listening. (2018). Retrieved January 30, 2018 from <https://www.kickstarter.com/projects/dopplerlabs/here-active-listening-change-the-way-you-hear-the/description>
- [3] Jeremy Agnew and Jeffrey M Thornton. 2000. Just noticeable and objectionable group delays in digital hearing aids. *Journal of the American Academy of Audiology* 11, 6 (2000), 330–336.
- [4] R Herbig and J Chalupper. 2010. Acceptable processing delay in digital hearing aids. *Hearing Review* 17, 1 (2010), 28–31.
- [5] Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, and Romit Roy Choudhury. 2018. MUTE: Bringing IoT to Noise Cancellation. In *ACM SIGCOMM*. 282–296.