Area Sound Enhancement Mic -Extracting Only Target Voice from Crowds and Environmental Sounds-

Takashi Ishiguro Masaru Fujieda
Kazuyoshi Akie

The problem of mics picking up unwanted sounds is more prevalent than ever as teleworking becomes widespread and publically installed systems turn to non-contact interfaces due to the spread of COVID-19.

During a web conference from home, the mic may pick up family voices, pet sounds, disaster warning broadcasts, emergency vehicle sirens, etc., and the sounds transmitted to the other conference participants may contain personal information (**Figure 1**). Even a web conference in the office poses a security risk if confidential conversation is overheard through the mic.

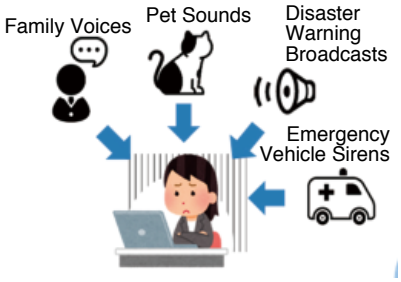

Figure 1. Problems of Home Web Conference

Requirement is growing for non-contact information terminals that allow hands-free communication without the use of a handset (**Figure 2**). However, during hands-free communication in a bustling environment, the mic will pick up unwanted background noise and make it difficult for the operator to hear the voice of the terminal user. Furthermore, if voice recognition is attempted hands-free, the recognition rate is significantly reduced.

Figure 2. Need for Hands-Free Information Terminal

This article introduces an overview of the area sound enhancement mic that solves the above-mentioned problems by enhancing only the sound in front of the mic as shown in **Figure 3,** and OKI's efforts for its commercialization.

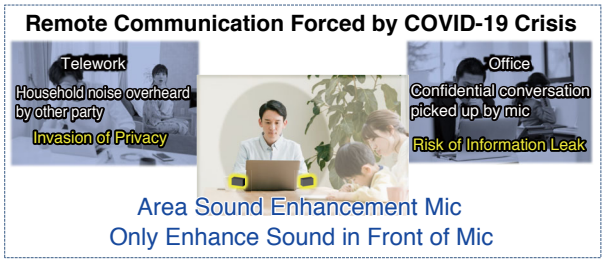

Figure 3. Area Sound Enhancement Mic

Conventional Sound Enhancement Technology

(1) Beamforming

Beamforming is one technology that uses several mics to form a mic array and enhances sound coming from a specific direction while suppressing other sounds[1]. Beamforming uses the time difference between the sounds reaching each mic to enhance and suppress the sounds coming from various directions and forms directivity.

(2) Implementation Example

As an example, a two-channel beamforming technology that forms sharp directivity in the front (and back) with two microphones will be introduced. The block diagram is shown in **Figure 4**. Two-channel beamforming first subtracts the input signal of one mic from that of the other mic. The target sound coming from the front is suppressed since it reaches the two mics at the same time, and the remaining noise from the left and right is extracted. Next, in the frequency domain, spectral subtraction (SS) is performed subtracting the amplitude information of the noise extracted by the subtractor from that of microphone 1's input signal. As a result, a sharp directivity in the front direction can be obtained, and only the target sound will be extracted[2].

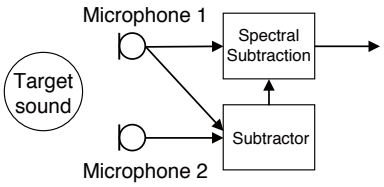

Figure 4. Block Diagram of Two-Channel Beamforming

(3) Beamforming Issues

In beamforming, noise coming from the same direction as the target sound cannot be suppressed. Therefore, when trying to pick up the target sound from a specific area as shown in **Figure 5**, not only the target sound, but noise coming from the same direction is also enhanced.

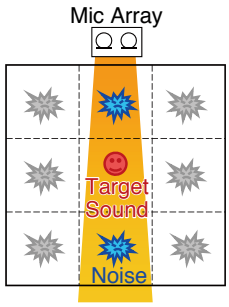

Figure 5. Beamforming Directivity

Area Sound Enhancement Mic

(1) Feature of Area Sound Enhancement Mic

The area sound enhancement mic is OKI's original technology. It utilizes two mic arrays set facing the target area and picks out only the sound from a specific area even if there are numerous noise sources surrounding the area.

(2) Mechanism of Area Sound Enhancement Mic

Using the two-channel beamforming introduced in the previous section, the two mic arrays each form sharp directivity in the front direction. As shown in **Figure 6**, mic array 1 extracts the target sound and noise 1, and mic array 2 extracts the target sound and noise 2. Since human voice is composed mainly of fundamental frequency components and harmonic components existing at frequencies that are integral multiples of the fundamental frequency components, the voice components are almost nonexistent at other frequencies. Therefore, even if the voices of several people are present at the same time, the components rarely overlap at the same frequency, and the target sound and noise components are scattered at different frequencies as shown in the bar graph of **Figure 6**.

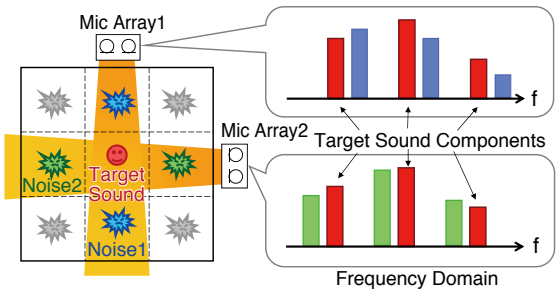

Figure 6. Comparison of Beamforming Output from Each Mic Array

The target sound is extracted by utilizing the fact that the two mic arrays extract the target sound at the same time and with the same magnitude. This flow is shown in **Figure 7.** First, the amplitude information of mic array 2 is subtracted from that of mic array 1 by SS. This suppresses the components of the target sound and allows the components of noise 1 to be extracted. Next, when the amplitude information of extracted noise 1 is subtracted from the amplitude information of mic array 1 by SS, only the target area sound remains. In this way, even if the target sound is surrounded by numerous noises as shown in **Figure 6**, only the target sound can be picked out using the area sound enhancement mic[3].

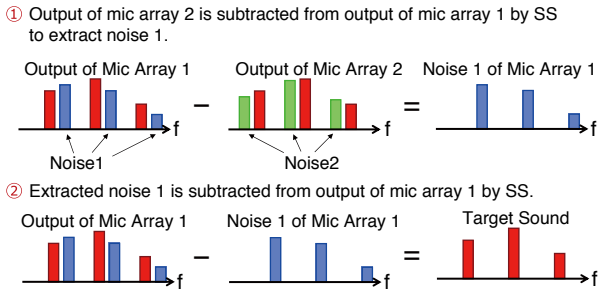

Figure 7. Target Sound Extraction Process

(3) Advantages of Area Sound Enhancement

The advantage of the area sound enhancement mic is that it can pick out sound from a distant area. A conventional mic array technology picks up noise coming from the same direction as the target sound, but the area sound enhancement mic can remove such noise and makes the target sound easier to hear. Furthermore, the area sound enhancement has the advantage of picking up only the voice of a speaker without requiring the speaker to wear a mic.

Usage Scenes

The area sound enhancement mic can be used for web conferencing from the home or office, and used also with information terminals installed in noisy environments. It will contribute to protecting privacy when working from home, protecting confidential information in the office, and enable hands-free calling and non-face-to-face customer service on information terminals (**Figures 8 and 9**).

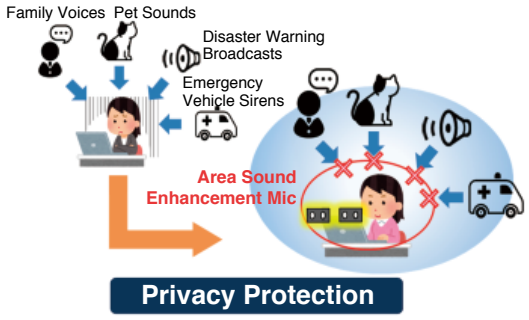

Figure 8. Mic for Teleworking

Figure 9. Mic for Information Terminal

In-House Demonstration Experiment

With the establishment of remote work, the calls in the office, which used to be telephone only, have increased overwhelmingly in both number of calls and that of people making those calls. Along with this trend, the problem of mics picking up conversations involving confidential information and other chats around the office during a web conference has come to surface.

To confirm the effect of area sound enhancement on the above problem, a set of enhancement mics were installed in an online conference booth and trial use was started (**Photo 1**), and more installations are planned in the future.

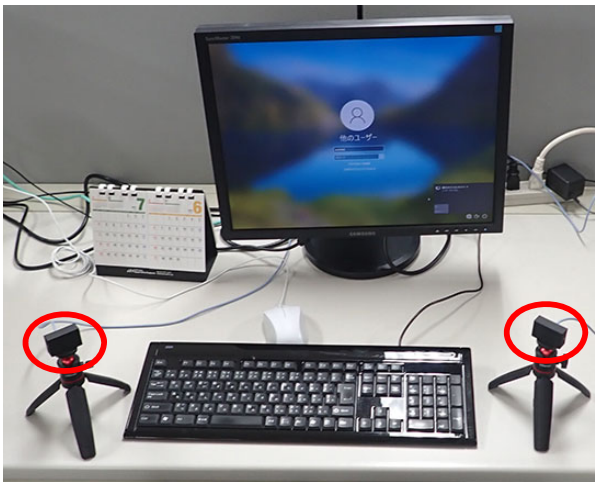

Photo 1. Mics Installed in Online Conference Booth

Demonstration Experiment with Information Terminal

OKI has provided signage display equipped with area sound enhancement mics and "CounterSmart" middleware to control the mics for the remote concierge service at Tokyu Railways' Toyoko Line / Den-en-toshi Line Shibuya Station, which is part of Tokyu Corporation's effort to promote DX (**Photo 2**)[4].

Photo 2. Experiment with Information Terminal at Shibuya Station

Due to the spread of the COVID-19, the number of users who feel uneasy about operating a terminal screen or making face-to-face contact with customer service is

increasing, thus heightening the need for non-contact/non-face-to-face customer services. Furthermore, against the backdrop of labor shortage caused by declining birthrate and aging population, and the progressing work style reform, demand is growing for the realization of remote information services. The “Remote Concierge Service” responds to such requests by enabling operators at Shibuya Station’s Tourist Information Center “WANDER COMPASS SHIBUYA” to remotely provide information (Japanese/English), or by utilizing AI chat to answer inquiries. However, it was necessary to address the issue that the operators had difficulty hearing the users’ voices over the background noise of the crowded train station.

The use of area sound enhancement mic enables a smooth hands-free conversation between the terminal user and the operator at a remote location even in a noisy environment such as a train station. Furthermore, voice recognition can also be used without issue. This has been confirmed through the demonstration experiment.

Efforts for Commercialization

Feeding back the verification results obtained in the demonstration experiment, development is advancing with the goal of releasing products in 2022. Among the efforts is improvement of voice quality such as preventing audio dropout, providing clearer sound, and improving the voice recognition rate.

OKI plans to develop both a consumer product to be sold in home electronics stores and EC sites, and a system product to be installed in equipment such as information terminals. For the consumer product, OKI is co-creating with a manufacturer that has experience developing and selling consumer electronics, and is currently working to develop a user-friendly product for telework use. For systems, specifications are being designed with consideration on simplifying the integration into systems such as guidance terminals, and ease of maintenance/operation.

Conclusion

The overview and commercialization efforts have been introduced for the area sound enhancement mic, which enhances only the sound in front of the mic and prevents the pickup of unwanted noise.

Area sound enhancement mic can be implemented in various devices such as headsets and neckbands, and it can be applied to applications other than Web conferencing and information terminals. OKI will promote proposals to expand product lineup and applications. ◆◆

References

1) Futoshi Asano: Sound Array Signal Processing, Issue 4, 2019, CORONA PUBLISHING CO.,LTD. (in Japanese)
2) Takashi Yazu, Makoto Morito: Sound Processing Technologies for Realistic Sensations in Teleworking, OKI Technical Review, Issue 213, Vol.75 No.2, October 2008 https://www.oki.com/en/otr/2008/n213/pdf/otr-213-R05.pdf
3) Kazuhiro Katagiri, Takashi Yazu: Area Sound Pickup System for Smooth Office Communication, OKI Technical Review, Issue 224, Vol.81 No.2, pp.60-63, October 2014 (in Japanese)
4) OKI Press Release, Verification of non-contact and non-face-to-face customer service under noisy conditions using noise reduction microphones, directional speakers, and touchless sensors, providing “CounterSmart” to the remote concierge service at Tokyu promoted DX Shibuya Station, April 21, 2021 https://www.oki.com/jp/press/2021/04/z21006.html (in Japanese)

Authors

Takashi Ishiguro, Systems Development Department-2, Network Systems Division, Solution Systems Business Group

Masaru Fujieda, Sensing Technologies R&D Department, Innovation Promotion Center

Kazuyoshi Akie, Systems Development Department-2, Network Systems Division, Solution Systems Business Group