

NOISE MASKING FOR SOLEMNITY AT THE HIROSHIMA PEACE MEMORIAL CEREMONY

SHINTARO SATO, WAKANA HISAGI AND SHUNSUKE ISHIMITSU

Graduate School of Information Sciences

Hiroshima City University

3-4-1, Ozuka-Higashi, Asa-Minami-Ku, Hiroshima 731-3194, Japan

sato.shint@mazda.co.jp; hisagi@sd.info.hiroshima-cu.ac.jp; ishimitu@hiroshima-cu.ac.jp

Received June 2022; revised October 2022

ABSTRACT. *On August 6 every year, the Hiroshima Peace Memorial Ceremony is commemorated in Hiroshima, Japan, to comfort the spirits of the atomic bomb victims and pray for the realization of permanent world peace. However, because various political marches are conducted in and around the venue, demonstrators' voices disturb the ceremony. Hiroshima City has repeatedly requested that demonstrators lower their volume and change the demonstration route. However, there has been no improvement in this regard. Since 2015, there has been a considerable decrease in demonstrators' disruptions owing to crowds and tents set up as sound barriers. Since 2020, the number of attendees has declined by 90% owing to the COVID-19 pandemic; thus, demonstrators' voices are easily heard in the venue. Previous studies applied active noise control, including high frequencies, to solving this problem. Implementing ANC is difficult because controlling high-frequency sounds in wide-open spaces is complex. In this study, we attempted to reduce demonstrators' noise by masking it with noise acceptable to the ceremony attendees in order to keep solemnity. Brown noise was found to be effective in the selection of masking noise. The experimental results confirmed the effective masking of male demonstrators. Therefore, we introduced adaptive masker control to mask all demonstrators and confirmed its effectiveness.*

Keywords: Masking, Public ceremony, Adaptive control, Annoyance, Noise

1. **Introduction.** Hiroshima, Japan, was the world's first city to experience an atomic bombing. The Hiroshima Peace Memorial Ceremony is held on August 6 each year to commemorate the event, comfort the spirits of the victims, and pray for permanent world peace. Therefore, this event should be conducted quietly and solemnly. However, political marches are conducted simultaneously in and around the venue, and demonstrators' voices and megaphone emissions disturb the ceremony's progress. Hiroshima City has repeatedly requested that demonstrators reduce the intensity of their megaphones and change the demonstration route; however, such requests have been ignored. In 2015, Tanaka and Ishimitsu attempted to solve this problem through a noise reduction approach using active noise control (ANC) [1]. The effectiveness of this method could not be confirmed as the sound pressure level of the demonstrators' voices was lower than the ambient noise that year. As the attendees were eager to meet the prime minister Shinzo Abe, they served as an acoustic absorbent material. Therefore, they reduced the noise. Tents have been set up as sound barriers, significantly reducing the noises disturbing the ceremony attendees. Moreover, ANC frequently adapts to low-frequency noise. In a previous study, ANC was applied with a field-programmable gate array (FPGA) to mask machinery noise, including high frequencies [2]. However, implementing ANC to mask

demonstration sounds during such events is problematic because it is impossible to control high-frequency sounds in wide-open spaces [3]. Owing to the COVID-19 pandemic in 2020, the number of attendees declined by 90%; therefore, the venue was less crowded, and demonstrators' voices were easily heard. This study investigated the possibility of reducing demonstrator noise through the masking principle [4]. Previous studies confirmed that masking is effective in protecting speech privacy. For example, conversations conducted at dispensing pharmacies are likely to contain patients' personal information; therefore, Tamesue et al. considered masking to protect speech privacy [5]. Zhou introduced the DNN method to avoid noises in speech signals, and the method could recover the speech components and spectrum structure polluted by noises [6]. Based on this information, masking may effectively reduce the noise from demonstrators. This masking was the first study to attempt to mask demonstrator noise at the Hiroshima Peace Memorial Ceremony. In the previous studies [7,8], the characteristics of noise that are not annoying as mask noise were considered, and the volume was manually adjusted during control. However, a manual adjustment of each difference in voice is labor intensive. Therefore, adaptive control of mask noise was also examined in this case. If the noise is successfully masked, the serenity of the ceremony can be preserved. The remainder of the paper is organized as follows. Section 2 explains the basis of masking, and Section 3 discusses the methods used. Then Section 4 explains the masking experiments performed at the ceremony and the results obtained. Section 5 explains adaptive masking. Finally, Section 6 presents the conclusions of this study along with the scope of future research.

2. Masking. *Masking* is a phenomenon in which the auditory threshold (minimum audible value) of a particular sound increases because of the presence of another sound [8]. The sound to be masked is referred to as a maskee, whereas that used for masking is called a masker.

Masking has been defined (ANSI, 199+a) as

1) the process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound;

2) the amount by which the threshold of audibility of a sound is raised by the presence of another (masking) sound, expressed in decibels.

It has been known for many years that a signal is most easily masked by a sound having frequency components close to, or the same as, those of the signal [9,10]. This led to the idea that our ability to separate the components of a complex sound depends, at least in part, on the frequency analysis that takes place on the basilar membrane. It also led to the idea that masking reflects the limits of frequency selectivity. If the selectivity of the ear is insufficient to separate the signal and the masker, then masking occurs. Thus, masking can be used to quantify frequency selectivity [11].

Simultaneous masking is usually explained in terms of two underlying processes. Masking may correspond to a "swamping" of the neural activity of the signal, a suppression of that activity by the masker, or a mixture of the two. If the masker is amplitude modulated so that the modulation is coherent or correlated in different frequency bands, a reduction in the signal threshold, known as comodulation masking release, can occur. Suppose the magnitude of the output from any single auditory filter is made unreliable due to the presence of the signal by randomizing the overall sound level of each stimulus. In that case, subjects can still detect the signal by comparing the outputs of different filters, a process called profile analysis.

3. Experimental Method. Here, the maskee is the demonstrators' voices amplified using megaphones. The masker is a sound that is not considered disturbing, and it can be

played at a volume that does not bother the attendees. Therefore, we selected colored noise, that is, white, pink, or Brownian, as possible candidates for the masker [7]. The most effective masker was determined by sequentially playing each colored noise simultaneously with the audio data of the ceremony recorded in the previous year. The result of this observation was that Brownian noise was the best candidate for the masker with a jury test as the preliminary experiment. The frequency characteristics of each colored noise are shown in Figure 1. As shown in the figure, as the frequency increases, the power density of the Brownian noise decreases by 6 dB per octave. Low sounds are effective in masking high sounds. Therefore, compared with white and pink noises, Brownian noise can be expected to be less noticeable in the treble, even with increasing volume.

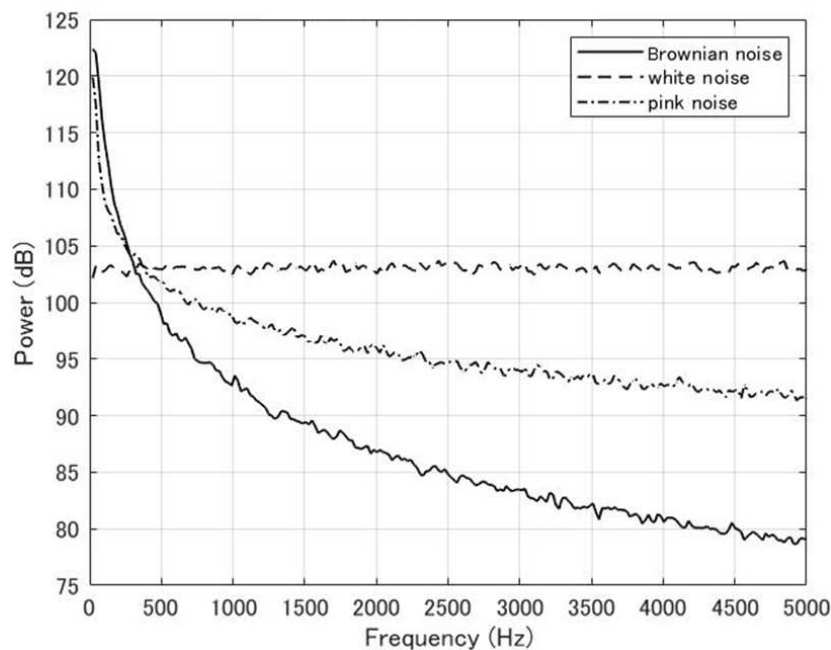


FIGURE 1. Comparison of the frequency characteristics of white, pink, and Brownian noise

4. Masking Experiment. An on-site masking experiment was conducted at the venue during the Hiroshima Peace Memorial Ceremony. Figure 2 shows the experimental field map, with the area outlined in red indicating the Hiroshima Peace Memorial Park. The ceremony and masking venue are conducted inside the park. The demonstrators come from the Atomic Bomb Dome yelling in unison, with their voices intensifying across the river as they get closer. An experimental tent was set up behind the attendees' tent, where we investigated the masking of demonstration sound by adjusting the masker volume. During the ceremony, the environmental sound was recorded binaurally using binaural microphones with and without masking. The wiring diagram of the equipment used in the masking experiment is depicted in Figure 3. The masker was played through masker speakers via a PC. The sound pressure level of the masker was adjusted using a mixing console. The first half of the ceremony was quiet because there were no demonstrators. However, as the prime minister's greetings began, the demonstrators' voices and the sound of megaphones gradually became audible. The demonstrators began marching from a place approximately 200 m east of the ceremony venue. The differences in frequency characteristics between the voices of male and female demonstrators are shown at a sampling frequency of 65536 Hz in Figure 4.

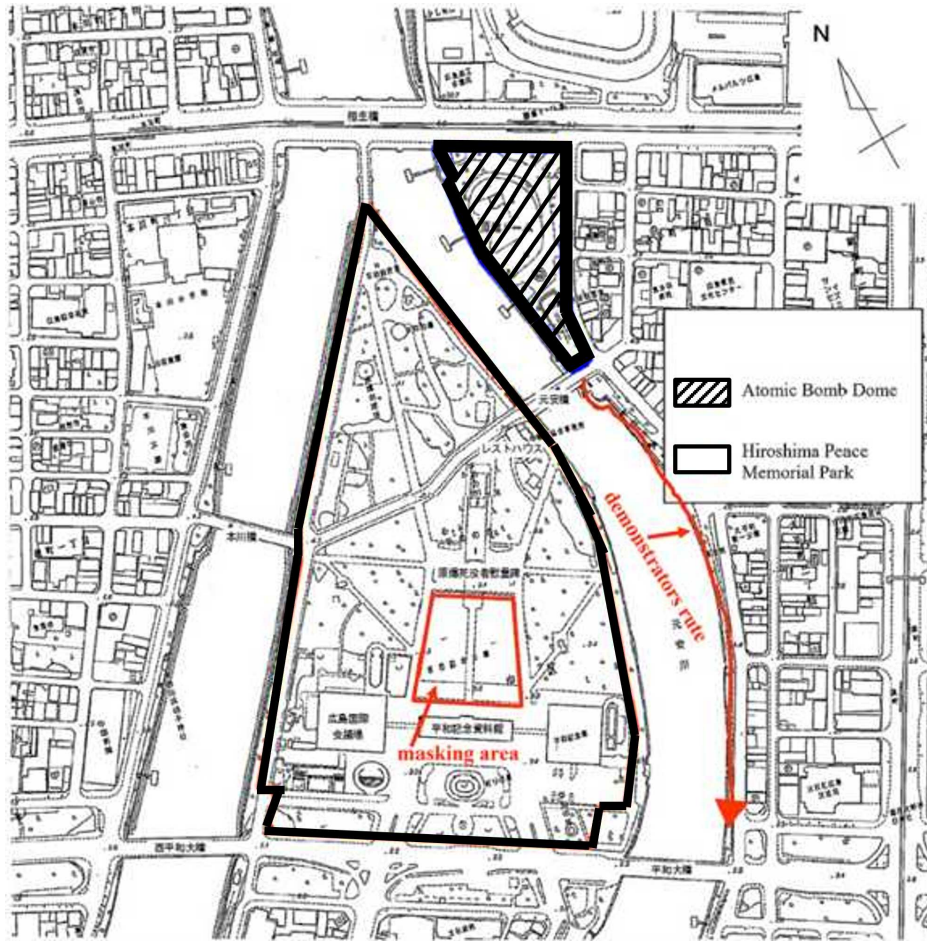


FIGURE 2. Experiment field map

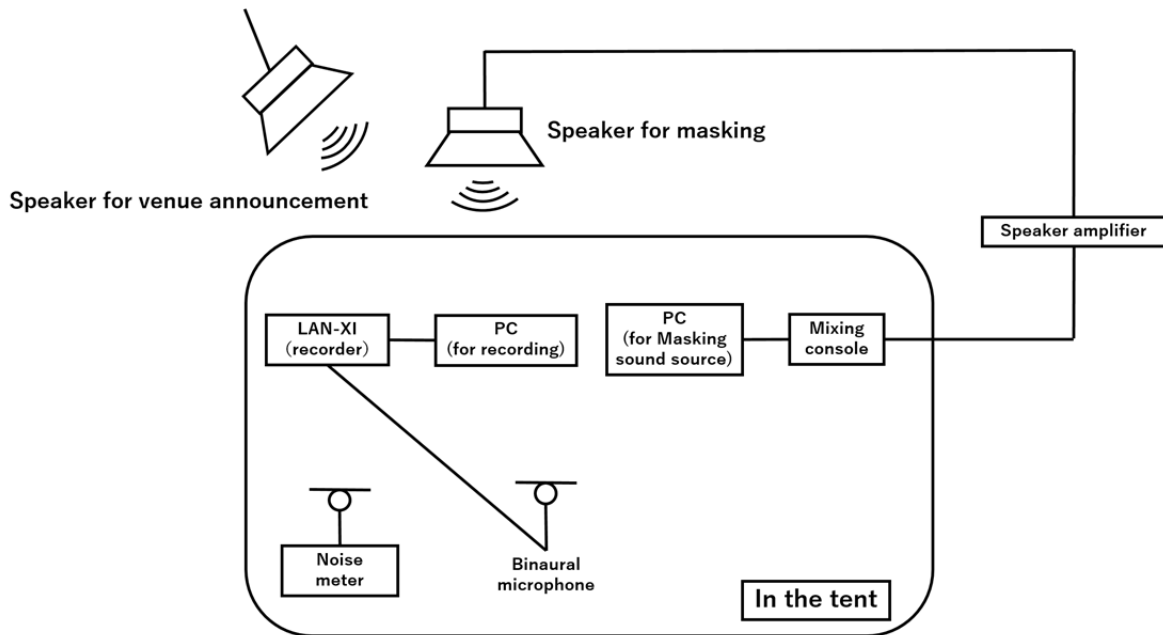


FIGURE 3. Schematic sketch of the equipment

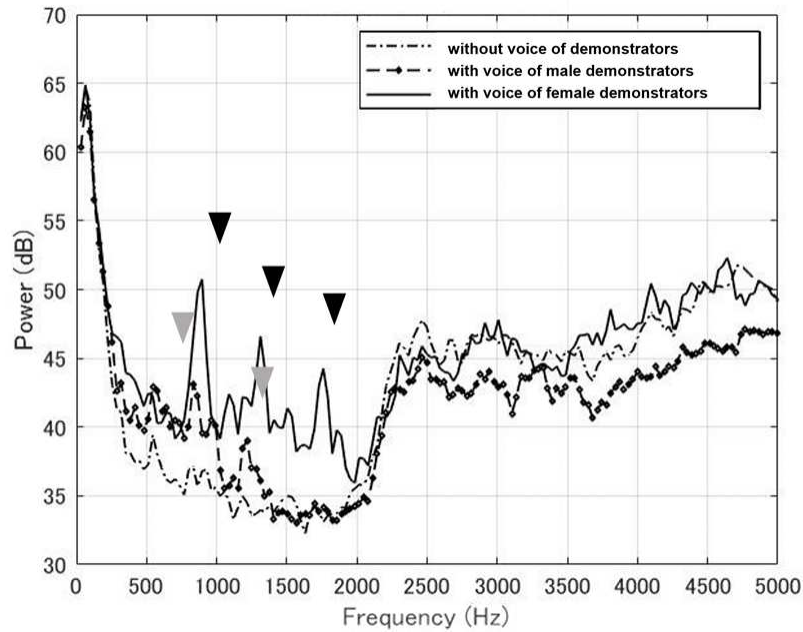


FIGURE 4. Difference between the frequency characteristics of male and female demonstrators' voices

As shown in Figure 4, there were prominent peaks at 640 and 1216 Hz for male voices (▼) and 896, 1312, and 1760 Hz for female voices (▼). This result reveals that examining the volume and frequency characteristics of the masker is crucial for masking both male and female voices.

The women used megaphones during the demonstrations this year; thus, the volume of female voices was louder, and the peak of the fundamental frequency of female voices appeared to be higher.

The frequency characteristics of the male voices were masked using Brownian noise with an agreeable volume (Figure 5). When the voices of male demonstrators were masked, the

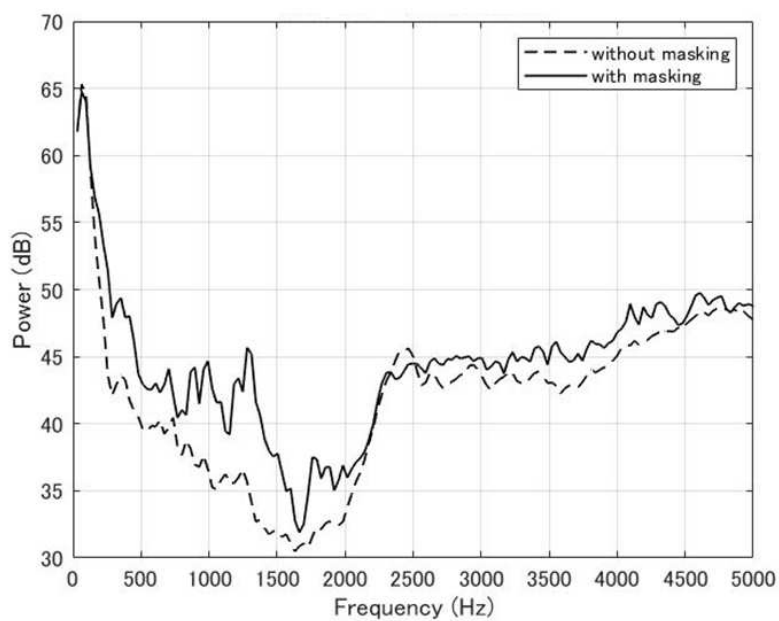


FIGURE 5. Frequency characteristics when masking out the voice of male demonstrators

power in the frequency range of 100 Hz to 2 kHz was more robust than the original data. During the experiment, the sound pressure level was monitored with a noise meter and was observed to be 60-65 dB without the masker and 66-70 dB with the masker. The Brownian noise was not perceived as noisy. Therefore, we consider that masking the voices of male demonstrators was successful [7].

However, the voices of the female demonstrators were not successfully masked. As mentioned previously, women were using megaphones; hence, their voices were more audible than the male voices. It is necessary to consider the masker type and its corresponding volume in further detail to efficiently mask both male and female voices. When the Brownian noise was played at a volume capable of masking the female voices, it became irritating. Meanwhile, when pink noise was played to mask out the female voices, its treble range became conspicuous, making it difficult to play during the ceremony. The reason was that the pink noise had a flatter sidelobe and seemed more suitable for masking female voices. However, it was still annoying. The same was true for white noise.

5. Adaptive Masking. During the experiment on the day of the ceremony, the volume of the masker was adjusted manually. However, the noise generated by the demonstrators was unsteady; therefore, the sound pressure level and frequency characteristics were constantly changing. An adaptive masking algorithm was developed to control the masker level. This algorithm extracts the audio band of the demonstrators' voices, calculates the signal-to-noise (SN) ratio with the background noise, and controls the volume of the masker. A flowchart of the algorithm is shown in Figure 6. In this algorithm, first, after sound is captured, the speech and non-voice portions are discriminated by linear prediction error, and the frequency band of the speech portion is identified. Next, the SNR between the speech and non-voice portions is calculated, and if the SNR is greater

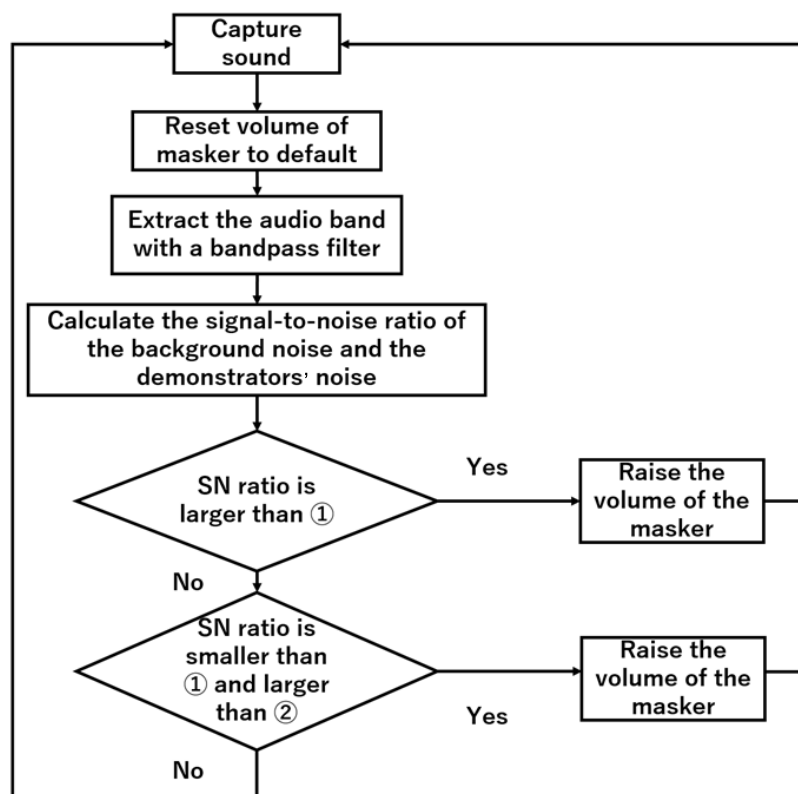


FIGURE 6. Flowchart of adaptive masking algorithm

than the value set in the first step ①, the volume of the masker is increased, and if it is less than the value set in the second step ②, the volume of the masker is increased in slightly smaller increments. The algorithm was adopted for a recorded data of loud male demonstrators near the Hiroshima Peace Memorial Ceremony. The time-varying SN ratio between masker and maskee is shown in Figure 7, and the time-varying masker is shown in Figure 8. Here, the demonstrators' voices were the signal (S), and the background noise was the noise (N). A comparison of Figures 7 and 8 indicates that adaptive masking may be feasible because the time-varying masker is synchronized with the SN ratio.

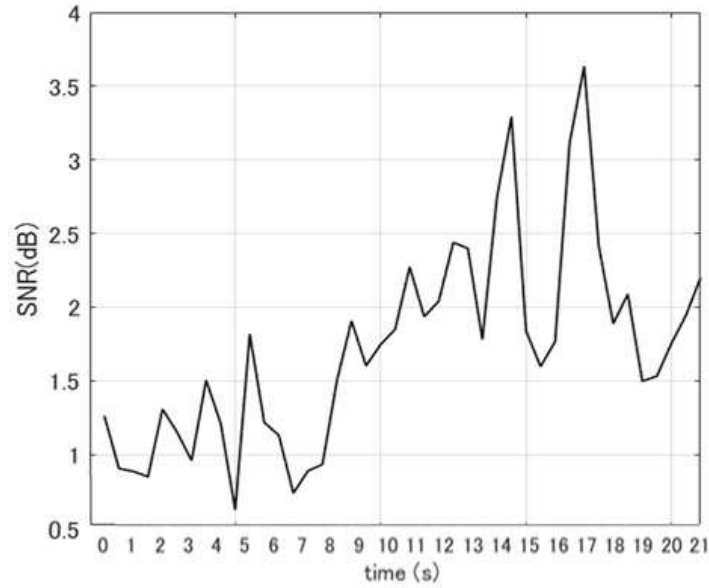


FIGURE 7. Time variation of SNR between masker and maskee

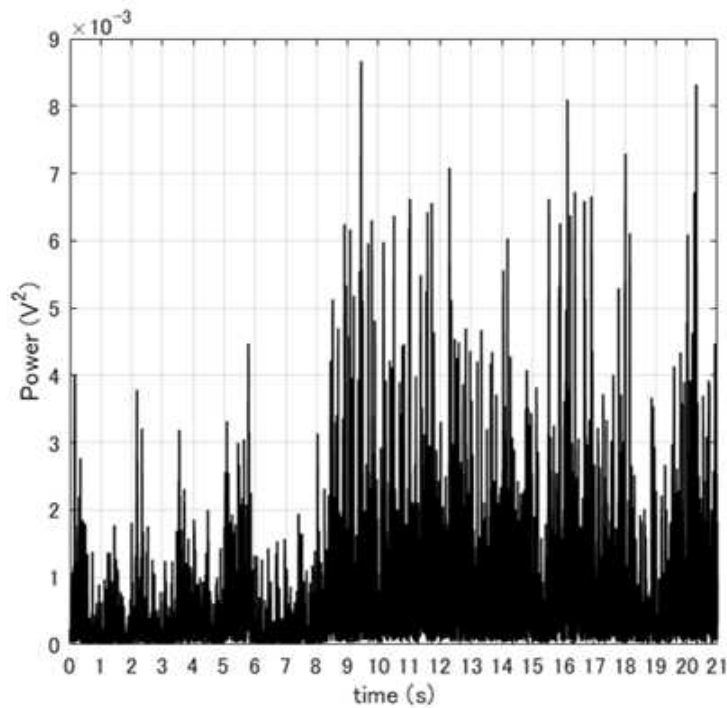


FIGURE 8. Time variation of amplitude of masker

6. **Conclusion.** In this study, we investigated the possibility of masking the noise caused by the demonstrators' voices and megaphone sounds during the Hiroshima Peace Memorial Ceremony. The outcomes and findings can be described as follows.

- Masking is considered a practical approach to making demonstrators' voices and megaphones inaudible and ensuring silence at the Hiroshima Peace Memorial Ceremony.
- The voices of the male demonstrators were masked successfully using Brownian noise, and they were not perceived as noisy.
- The voices of female demonstrators using megaphones were louder than male voices; thus, using Brownian noise to mask them was ineffective.

Based on the above, we incorporated an adaptive volume control system. Since the volume control worked well, it is believed that effective use of this system will provide a quiet environment for people praying and guests at the Hiroshima Peace Memorial Ceremony.

In the future, we plan to investigate different masker types and volumes to successfully mask both male and female voices to maintain a sense of calm during the Hiroshima Peace Memorial Ceremony and other such occasions.

REFERENCES

- [1] D. Tanaka and S. Ishimitsu, Noise reduction for the Hiroshima Peace Memorial Ceremony, *Proc. of the 23rd International Congress on Sound & Vibration*, Athens, Greece, 2016.
- [2] S. Sato, S. Ishimitsu, T. Kimura and K. Kurokawa, Study of broadband active noise control, *ICIC Express Letters*, vol.14, no.8, pp.807-812, 2020.
- [3] P. A. Nelson and S. J. Elliott, *Active Control of Sound*, Academic Press, London, 1992.
- [4] G. A. Miller, The masking of speech, *Psychological Bulletin*, vol.44, no.2, pp.105-129, 1947.
- [5] T. Tamesue, S. Yamaguchi and T. Saeki, Study on achieving speech privacy using masking noise, *Journal of Sound and Vibration*, vol.297, no.6, pp.1088-1096, 2006.
- [6] Q. Zhou, Research on English speech enhancement algorithm based on improved spectral subtraction and deep neural network, *International Journal of Innovative Computing, Information and Control*, vol.16, no.5, pp.1711-1723, 2020.
- [7] S. Sato and S. Ishimitsu, Basic examination of noise masking in the Hiroshima Peace Memorial Ceremony, *ICIC Express Letters, Part B: Applications*, vol.12, no.3, pp.263-268, 2021.
- [8] S. Sato, W. Hisagi and S. Ishimitsu, Sound barrier at Hiroshima Peace Memorial Ceremony, *Proc. of the 27th International Congress of Sound and Vibration*, 2021.
- [9] R. D. Luce, *Sound & Hearing*, Psychology Press, London, 1993.
- [10] A. M. Mayer, Researches in acoustics, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol.37, no.226, pp.259-288, DOI: 10.1080/14786449408620544, 1984.
- [11] R. L. Wegel and C. E. Lane, The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear, *Physical Review Journals Archive*, vol.23, pp.266-285, 1924.

Author Biography



Shintaro Sato received M.E. at Hiroshima City University in 2021. He has been employed by Mazda Motor Corporation since 2021 and is working on developing car sound devices. His main area of interest is sound quality control on automotive audio.



Wakana Hisagi is currently enrolled as a second-year M.E. student at Hiroshima City University and a double degree student with Hochschule Hannover. She specializes in adaptive control of sound.



Shunsuke Ishimitsu is a professor in the Department of System Engineering of the Graduate School of Information Sciences, Hiroshima City University, Japan. He received M.E. and Ph.D. from Toyohashi University of Technology in 1990 and 1995, respectively. He was an employee at Pioneer Corporation from 1990 to 1997 and an associate professor at Oshima National College of Technology from 1997 to 2005. He was also a visiting researcher at ISVR, the University of Southampton, from 2001 to 2002. Then, he was an associate professor at the University of Hyogo from 2005 to 2007. Then he has been served as a professor at Hiroshima City University and Dean of the Graduate School of Information Sciences, Hiroshima City University, Japan. His main area of interest is sound quality evaluation, psychoacoustics, active noise control, and signal processing. He has published over 100 papers in refereed journals and 500 conference papers and is author of *Sound Design Theory* (2022), co-author of *Basis of Ergonomics* (with H. Sato 2018) and *Production System Engineering* (with Y. Shimizu et al. 2001).