2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Car Running Noise Detection System Using Frequency Change for Deaf and Hard-of-Hearing People

Akemi Matsuo[1] and Taku Itami[2]

Abstract— **Deaf and hard-of-hearing (DHH) people have difficulty obtaining information from hearing. Therefore, information from other senses, such as sight, plays an important role. However, in everyday life, DHH people face life-threatening challenges, such as the inability to notice approaching hazardous sounds from blind spots or from behind, and the time it takes them to avoid danger. In this paper, we focus on one of the hazardous sounds, the driving noise of a car, and propose a system to detect the approach and passage of a car to the device. The effectiveness of this system is demonstrated by detecting in real time the approach or passage of a car toward the wearer based on the time variation of the frequency received by the microphone installed on the smart glasses, and displaying the results on the lens. We discuss the proposed method and our validation of determining the approach and passage of cars.**

I. INTRODUCTION

The number of deaf and hard-of-hearing (DHH) people in Japan is estimated to be about 20 million, or about 15% of the Japanese population. In addition, since DHH tends to increase with age, it is estimated that this ratio will further increase in the future[1].

In everyday life, auditory cues provide information about events outside people's field of vision and help detect potential hazards. Especially in road traffic, it is very important to extract meaningful sounds from various noises and to recognize what the sound source is in order to decide what to do next. However, because DHH people have difficulty discriminating sound sources, it is difficult for them to notice the sound of a car running or warning sound at intersections and other busy traffic situations, making them more likely to be involved in an accident or traffic jam than people with normal hearing. In fact, it has been shown that people with hearing loss are 70% more likely to be involved in traffic accidents than people with normal hearing, putting them at risk for life-threatening injuries. But today's society is not necessarily barrier-free for such people[2][3][4][5].

In addition, the sirens of ambulances and police cars are representative of hazardous sounds in road traffic, and these sirens have been the subject of many studies on hazardous sound detection[6][7][8][9]. These emergency vehicles are characterized by warning lights mounted on the top of the vehicle in addition to sirens. The visual effect of warning lights allows people to identify the presence of emergency vehicles, and they are particularly effective at night. However, in bright daylight, their light may be blurred, difficult to see due to low contrast against the brightness of the day, or hidden by other cars. If the approach of emergency vehicles is difficult to obtain both audibly and visually, those vehicles can be more dangerous to DHH people than ordinary cars[10].

For the above reasons, it is crucial to alert DHH people of the danger and encourage them to avoid the danger as soon as a car approaches.

Many companies are now developing devices and applications for DHH people. Among these, smartphones[11][12][13], head-mounted displays[14][15][16], and smartwatches[17][18] have received particular attention with regard to the development of wearable sound recognition, for reasons included improved privacy, social acceptability, and integrated support for both visual and haptic feedback.

”VisAural”, developed by Benjamin M Gorman, is a system that converts sound information into visual cues. Sound direction is detected by a beamforming method using an array of microphones on the head mount, and LEDs are placed around the periphery of the user's field of vision to guide the user to the sound source. This allows DHH people to visually perceive the direction of sound[14].

”Sound Watch”, developed by Dhruv Jain, is an application for smartwatches that analyzes ambient sounds with a deep-CNN-based sound classifier to provide real-time sound source recognition assistance to people with hearing loss. By providing visual and tactile feedback, the system is capable of detecting and transmitting a variety of sounds, including sirens and other warning sounds, as well as birdsong, knocks on doors, and so on[17].

With these developments, there has been much research in Japan and abroad on real-time car proximity detection in traffic environments for DHH people. Three examples of such methods are deep learning-based, Doppler effect, and frequency response-based methods.

The deep learning base can be further classified by its learning content, which can be divided into three patterns: voice data, spectrogram data of voice, and image data from cameras. In any pattern, each data is learned in advance, and the approach of a car is detected by matching the actual measurement with the learned data[17][19]. This method obtains high performance by training with a large amount of data. However, it is not practical for wearable devices at this time because it requires a large amount of data and a computer with high computing power for learning.

The Doppler effect, “the closer the sound source is, the higher the audible sound is, and the further away the

[1]Graduate School of Science and Technology, Aoyama Gakuin University, Kanagawa, Japan a.matsuo.sce@gmail.com
[2]Associate Professor Department of Electronics and Bioinfomatics School of Science and Technology, Meiji University, Kanagawa, Japan itami@meiji.ac.jp

sound source is, the lower the audible sound is," detects the approach of a vehicle by detecting slight changes in frequency values. This method is effective and accurate for ambulance and police car sirens, whose frequency values are defined by law[20]. However, detection of this frequency change requires a high-performance PC for verification of the detection algorithm and a microcontroller specialized for siren sound detection, making practical application difficult from a cost standpoint. It can also be said that the Doppler effect is very difficult to detect for the driving noise of cars, which is the target of detection in this study, because the frequency value is not defined to a specific value depending on the road condition and the type of car.

For these reasons, in this study, we propose a system that uses the frequency characteristics of cars in road traffic to detect the approach and passage of cars in real time and display the output on the lenses of smart glasses. By using smart glasses, DHH people can visually see the detection results. As shown in Fig.1, this method uses only the recording analysis of a single microphone installed on the smart glasses for detection, making it possible to reduce weight and cost. The effectiveness of this system is demonstrated by detecting in real time the approach or passage of a car toward the wearer based on the time variation of the frequency received by the microphone installed on the smart glasses, and displaying the results on the lens.

Fig. 1. Smart glasses that perform sound analysis in real time and output the results on the lens to enable visual recognition of detection results.

II. PROPOSED METHOD

According to the recent data on car driving noise on 20 public roads in Japan, it is clear that the sound power level of car driving noise increases in the frequency range of 500 Hz to 1600 Hz[21]. In an outdoor sound environment on a sunny day, we actually recorded audio in a noiseless condition with no car traffic, and three patterns of audio when driving a car (α:uphill, β:flat road, γ:downhill), respectively, and analyzed the recorded results using a spectrum analyzer with frequency on the horizontal axis and decibel value on the vertical axis. Figures 2, 3, 4, and 5 show the respective results. These actual measurement results also confirm that, compared to noise, the running sound of a car contains more frequency components around 1000 Hz. In other words, when a car approaches from a noise-only condition, the

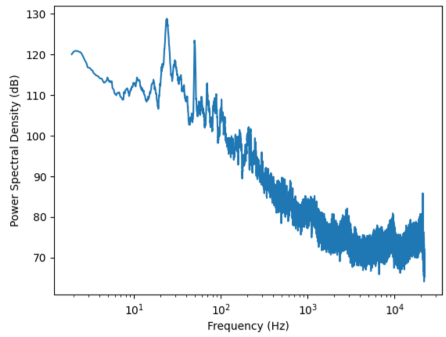

Fig. 2. Frequency response when there is no car driving and only noise

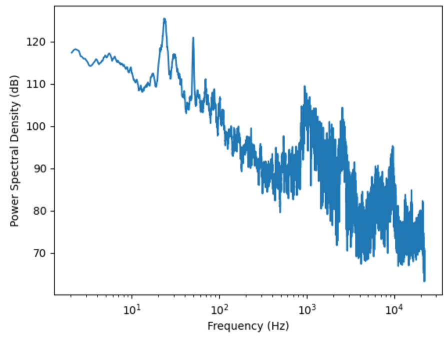

Fig. 3. Frequency response for driving noise (α:uphill)

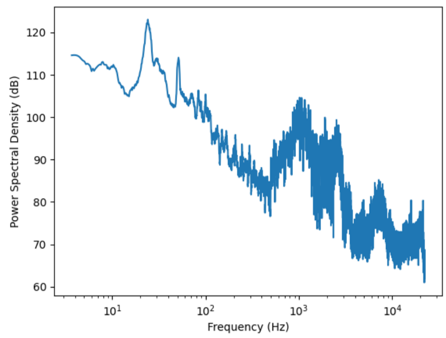

Fig. 4. Frequency response for driving noise (β:flat road)

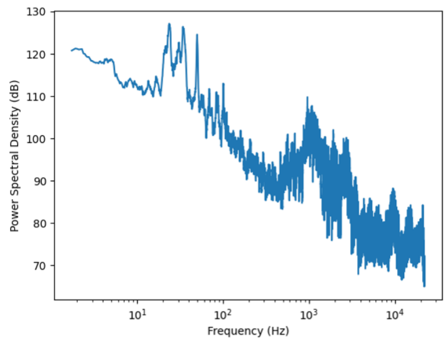

Fig. 5. Frequency response for driving noise (γ:downhill)

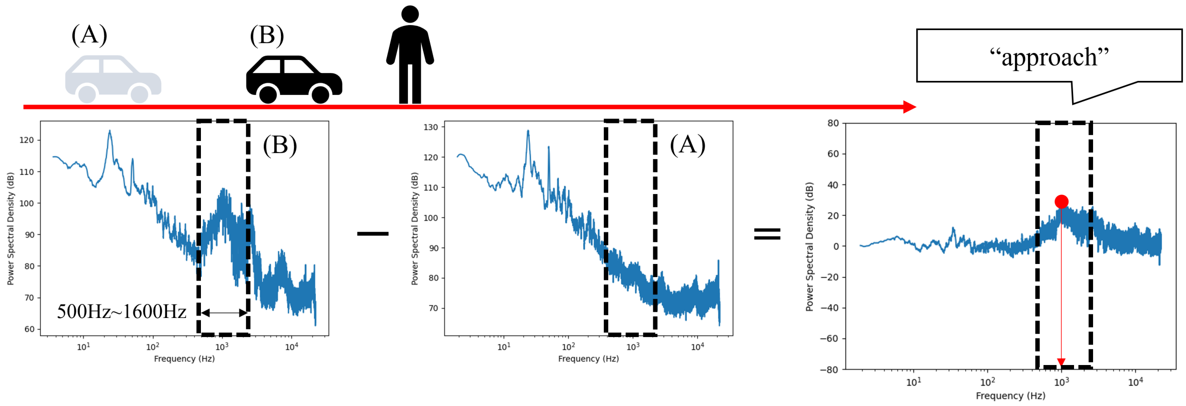

Fig. 6. Principle of car approaching detection. Chronologically, subtract the previous sound (A) from the current sound (B). If the resulting frequency with the highest decibel value is between 500 Hz and 1600 Hz, the system determines that a car is approaching and displays the word " approaching " on the lens of the smart glasses.

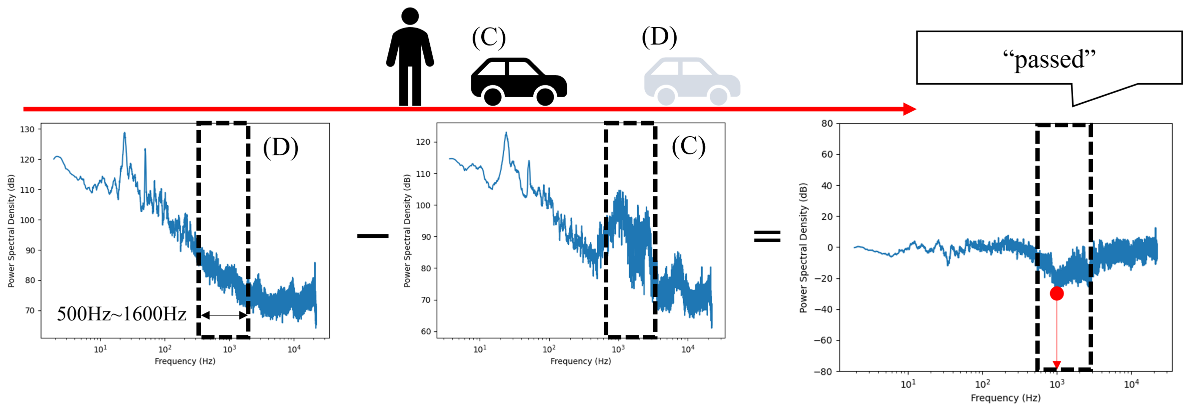

Fig. 7. Principles of car passage detection. Chronologically, subtract the previous sound (C) from the current sound (D). If the resulting frequency with the lowest decibel value is between 500 Hz and 1600 Hz, the system determines that a car is passing and displays the word " passed " on the lens of the smart glasses.

sound level of 500 Hz to 1600 Hz, which is the frequency characteristic of car running sound, will gradually increase, and conversely, when a car passes and returns to noise only, the sound level of 500 Hz to 1600 Hz will gradually decrease. In this method, the approach and passage of a car is determined by observing the frequency characteristics of this car running sound in time series.

First, consider the situation shown in Fig. 6 regarding the detection of approaching a car. (A) is a state in which the device wearer is at a distance from the car and only noise is recorded by the microphone, and (B) is a state in which more time has passed than in (A), the car is closer to the device wearer, and driving noise is included in the recording. At this time, subtract the frequency response of the recording in (A) from the frequency response of the recording in (B), then the pre-approach audio component can removed from the audio of the approaching car. If the frequency with the highest decibel value in the resulting subtracted sound, i.e., the frequency component most abundant in the sound, is within the frequency range of 500 Hz to 1600 Hz, which is the frequency range of car running sound, the system judges that the car running sound is gradually becoming louder and detects an approaching car.

Next, consider the situation shown in Fig. 7 for detecting the passage of a car. (C) is a state in which a car is passing close to the device wearer and the recording includes driving noise, and (D) is a state in which more time has passed than in (C), the device wearer and the car are at a distance, and only noise is recorded by the microphone. At this point, as with the approach, subtract the frequency response of the recording in (C) from the frequency response of the recording in (D), then the component of the sound at the time of passage can removed from the sound after the car passes. If the frequency with the lowest decibel value in the resulting subtracted sound, i.e., the fewest component in the sound, is within the frequency range of 500 Hz to 1600 Hz, which is the frequency range of car running sound, the system judges that the car running sound is gradually becoming quieter and the passing of a car is detected.

III. Verification of effectiveness

A. experimental environment

This method assumes that the sound does not include human speech or the sound of a bicycle bell, since it assumes an outdoor situation in which a car is driving from a noise-only state. Therefore, this experiment was conducted in a quiet room, using pre-recorded sound of a car driving as sound sources. The background noise is 10.5[dB].

B. experimental equipment

As shown in Fig. 1, a microphone(Comica Technology Ltd. CVM-V02C labelled microphones) was placed at one end of the smart glasses(Nreal Air NR-7100RGL, Nreal Corporation Japan). The experimental environment is shown in Fig. 8, where a sound source was placed 0.5 m away from the front of the smart glasses to capture data. For the sound source, we used audio recordings of three patterns of car driving sounds in a real environment, α:uphill driving, β:driving on a flat road, and γ:driving downhill.

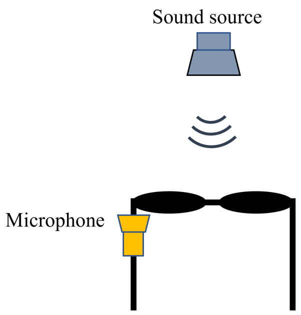

Fig. 8. The experimental environment for judging the approach and passage of a car is shown below. Three patterns of car running sound were played from the front of the smart glasses, and the frequency changes were analyzed by the program to confirm whether approaching and passing cars could be detected correctly.

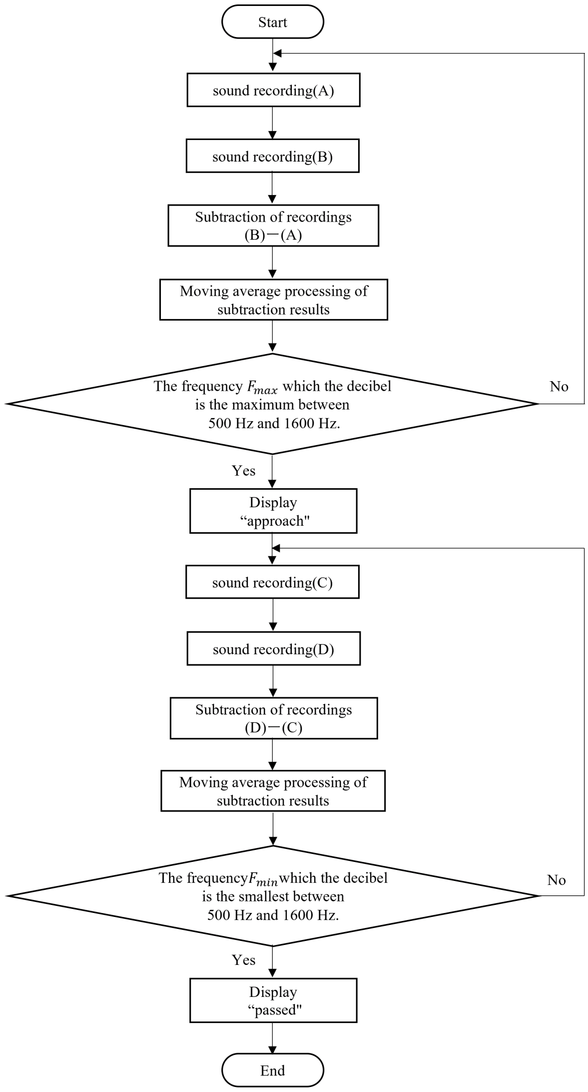

Fig. 9. Data processing

C. data processing

Figure 9 shows a flowchart of data processing. All data processing was performed in a computer (LAVIE DESKTOP-NUK21OR 8.00GB 64bit). Since this method uses the frequency change of sound by time series, and it is clear from previous study[22] that the change of sound can be sufficiently observed in 0.3 seconds, we record for 0.3 seconds, which is the equivalent of 100 samples.

At the start of program execution, two 0.3 seconds recordings are made in succession, and the frequency characteristics of the horizontal axis frequency and vertical axis decibel value are observed for each. About those frequency characteristics, subtract the previous recording from the later recording. And, for the results, every 20 samples moving average processing is performed to make the movement of the characteristic waveforms clearer. In the characteristics after this processing, if the frequency "$Fmax$" which the decibel is the maximum is between 500 Hz and 1600 Hz, it is judged that the driving noise component in the sound is getting louder, i.e., the car is approaching, and the word "approach" is displayed on the lens of the smart glasses. If the approach of a car is not detected at this point, two recordings are made again, and the system returns to the detection process.

After detecting the proximity, make two consecutive recordings of 0.3 seconds again, and observe the frequency characteristics of the horizontal axis frequency and vertical axis decibel value for each. As before, subtract the previous recording from the later recording, and perform a moving average process on the results every 20 samples. In the characteristics after this processing, if the frequency "$Fmin$" which the decibel is the smallest is between 500 Hz and 1600 Hz, it is judged that the driving noise component in the sound is getting smaller, i.e., the car has passed and moved away, and the word "passed" is displayed on the lens of the smart glasses. As before, if the passing of a car is not detected at this point, two recordings are made again and then return to the detection process.

By repeating this process, the system detects the approach and passage of a car's running sound in real time. In this experiment, detection accuracy was verified by performing three trials of this process for each sound source.

IV. RESULTS

Table.1 shows the results of the approach and passage detection for each of the three patterns of car running sound. From left to right, the running sound used as the sound source, the results of detection accuracy for approaching, the results of detection accuracy for passing, and the results of overall detection accuracy. Figure 10 shows the percentage of overall car driving noise detection results. The results in Table.1 confirm that the detection accuracy is about 89% for approach, about 67% for passage, and about 78% overall. For each pattern of the car's running sound, detection accuracy was better on flat roads than on hills. But in the results of correct detection, the "approaching" and "passed" results were displayed almost in real time at the beginning and end of the car running sound in the audio. And, as shown in Fig. 10, the car's driving noise was accurately detected with an accuracy of about 80%. These confirm that this method is capable of detecting approaching and passing cars.

TABLE I

RESULTS OF DETECTING THE APPROACH AND PASSAGE OF A CAR RUNNING SOUND. THIS TABLE SHOWS THE THREE PATTERNS OF CAR RUNNING SOUND, THE ACCURACY OF APPROACH DETECTION, THE ACCURACY OF PASSING DETECTION, AND THEIR OVERALL DETECTION ACCURACY FOR EACH RUNNING SOUND.

	detection accuracy		
	approach	passed	overall
α : driving uphill	100%	33%	67%
β : driving on a flat road	100%	100%	100%
γ : driving downhill	67%	67%	67%
Total	89%	67%	78%

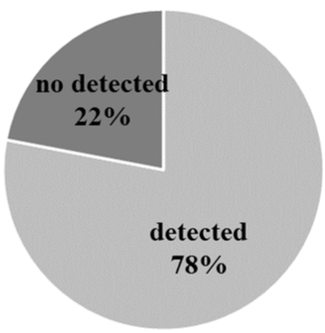

Fig. 10. The percentage of overall detection results. Approximately 80% accuracy was confirmed for the detection accuracy of car runs using this method.

V. DISCUSSION

A possible reason for the lower detection accuracy on uphill and downhill than on flat roads is that the frequency change in the two recordings may have been unexpected. In this method, the recording time is set to 0.3 seconds in order to capture sufficient changes in the sound. However, if the frequency change in that 0.3 second period is small, or if the change in other frequency values is more significant due to sudden noise, the frequency characteristics that appear when subtracting the two are very different from the frequency characteristics assumed by this method, and thus the method will not correctly detect the change.

In addition, this method detects approaching and passing cars using frequency changes that occur before and after 0.3 seconds, but the results of these frequency changes vary depending on the motion of the car traveling at that time. For example, if a car is traveling slowly at 10 km/h on a flat road, and if it is traveling at a high speed of 100 km/h on the same road, it can be inferred that the latter case has a more significant frequency change within the recording time. It cannot be assured that the three running sounds used in this experiment were at the same running speed, which may have caused differences in detection results. In this study, three patterns of car running sound were used: uphill, flat road, and downhill. To compare the detection accuracy in each pattern more accurately, it is necessary to consider differences in running speed as well.

The driving noise of a car is composed of engine noise, exhaust noise, and tire noise, and the contribution of these generated noises varies greatly depending on the road conditions and driving conditions. Therefore, it cannot be said that the frequency response will be the same on all roads. In order to develop a more accurate method, it is necessary to analyze the components of driving noise in more detail and classify and pattern the frequency characteristics according to road conditions and weather[23]. And, there are many car, pedestrians, bicycles, etc. in actual road traffic. In order for this method to be effective in such an environment, noise processing during recording and extraction of specific sounds are necessary.

Because this method uses subtraction between recordings, it takes two recordings of 0.3 seconds, or 0.6 seconds of recording time. Considering that the speed limit on ordinary roads in Japan is roughly 60 km/h, this means that the car is moving 10 m during the recording time. Including the time for the device wearer to recognize the approach and move to a safe location, it is assumed that the car is sufficiently close to the wearer. The recording time needs to be considered in the future, both for the safety of DHH people and for the real-time performance of the analysis.

Although this method examined the detection of approaching and passing car running sound, it is believed that simultaneous sound source localization would be effective to minimize the time required for DHH people to recognize the detection results and find the sound source. Various methods exist for sound source localization, including beamforming methods[24][25], machine learning[26][27][28], and stereo audio analysis[29]. Since these methods of sound source localization are recognized for their real-time performance and sufficient measurement accuracy, it is expected that the fusion of this method and sound source localization will promote faster hazard avoidance for DHH people.

VI. CONCLUSIONS

In this paper, we proposed a system that detects the approach and passage of a car's running sound to the device, with the aim of facilitating hazard avoidance for DHH people, and experiments were conducted using smartglasses to verify the effectiveness of this system. The experimental results showed that the system was able to detect the approach and passage of a car with a probability of about 80%, and it can be said that this method can be used to realize a system that assists DHH people in responding quickly to hazardous sounds in real time.

Future tasks include more detailed analysis of the sound components that make up the sound of a car driving, so that it is not affected by road conditions; improving real-time performance by doing so; and integrating sound source localization to complete the system to a practical level.

REFERENCES

[1] Ashita SARUDAKE, Hiroshi NUNOKAWA and Kenzo ITOH, "Life Sound Support System Using Smartphone for Person with Hearing Loss," International Journal of Affective Engineering, vol. 15, No.1, pp. 97-105(2016).

[2] Yasuhiko Tsuge and Noboru Ohnishi, "Discrimination of Alarm Signal for Deaf People," The Institute of Image Information and Television Engineers, IDY99-66/HIR99- 1/MIP99-20 AIT99-20/NIM99-20/VIS99-20.

[3] Yuki KANDA, Katsuyuki ONISHI, Makoto ISOGAWAQ, Yoshihisa UCHIDA "Proposal of sound source position estimation method by portable sound direction estimation device" No. 18-2 Proceedings of the 2018 JSME Conference on Robotics and Mechatronics, Kitakyushu, Japan, June 2-5, 2018

[4] T Suneel, A S Vishnu Mahesh, B Akhil Kumar, K Blessy Babu, B Gayatri, "Horn Detection System for Four-Wheeler Using Arduino", 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 22-23 February 2024

[5] Matthias Mielke, Rainer Brueck, "Design and evaluation of a smartphone application for non-speech sound awareness for people with hearing loss", 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 25-29 August 2015

[6] Yuvarani P, Suvetha C, Shree Subha M, Subashri V, Leena G, "Automatic Recognition of Ambulance Siren by Traffic Signal", Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 2024

[7] Lorant Andras Szolga, Alexandru Nicolae Strugaru, "Emergency Siren Detection System for Deaf People", International Conference on e-Health and Bioengineering (EHB), 2021

[8] Dharma Rane, Pushkar Shirodkar, Trilochan Panigrahi, S. Mini, "Detection of Ambulance Siren in Traffic", International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), 2019

[9] Sagar Bapodara, Shyam Mesvani, Manish Chaturvedi, Pruthvish Rajput, "Traffic Congestion and Emergency Vehicle Responsive Traffic Signal Control in Resource Constrained Environment", 11th International Symposium on Electronic Systems Devices and Computing (ESDC), 2023

[10] Josef Palecek, Martin Cerny, "Emergency horn detection using embedded systems", IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMI), 2016

[11] Danielle Bragg, Nicholas Huynh, and Richard E. Ladner, " A Personalizable Mobile Sound Detector App Design for Deaf and Hard-of-Hearing User ", In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility, 3–13.

[12] Matthias Mielke and Rainer Brueck, "Design and evaluation of a smartphone application for non-speech sound awareness for people with hearing loss ", In Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, 5008–5011.

[13] Liu Sicong, Zhou Zimu, Du Junzhao, Shangguan Longfei, Jun Han, and Xin Wang. "UbiEar: Bringing Location-independent Sound Awareness to the Hard-of-hearing People with Smartphones ", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, 2: 17

[14] Benjamin M Gorman. 2014. "VisAural: a wearable sound-localisation device for people with impaired hearing ". In Proceedings of the 16th international ACM SIGACCESS conference on Computers and accessibility, 2014

[15] Eric G Hintz, Michael D Jones, M Jeannette Lawler, Nathan Bench, and Fred Mangrubang, " Adoption of ASL classifiers as delivered by head-mounted displays in a planetarium show", Journal of Astronomy and Earth Sciences Education (JAESE) 2, 1: 1–16, 2015

[16] Dhruv Jain, Leah Findlater, Christian Volger, Dmitry Zotkin, Ramani Duraiswami, and Jon Froehlich. " Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing". In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 241–250, 2015

[17] SoundWatch: Exploring Smartwatch-based Deep Learning Approaches to Support Sound Awareness for Deaf and Hard of Hearing Users Dhruv Jain1 , Hung Ngo1 , Pratyush Patel1 , Steven Goodman2 , Leah Findlater2 , Jon Froehlich1 1Computer Science and Engineering, 2Human Centered Design and Engineering University of Washington, Seattle, WA, USA, ASSETS 2020

[18] Steven Goodman, Susanne Kirchner, Rose Guttman, Dhruv Jain, Jon Froehlich, and Leah Findlater." Evaluating Smartwatch-based Sound Feedback for Deaf and Hard-of-hearing Users Across Contexts ". In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1– 13.

[19] Van-Thuan tran, Wei-Ho Tsai「Audio-Vision Emargency Vehicle Detection」IEEE SENSORS JORNAL, VOL.21, NO.24, DECEMBER 15, 2021

[20] Takuya Miyazakia , Manabu Shimakawab , Yuhki Kitazono, " Research of Ambulance Siren Detector using dsPIC " Journal of the Institute of Industrial Applications Engineers, Vol. 2, No. 1, pp. 11-15, Mar. 2014

[21] Miki Yonemura, Hyojin Lee and Shinichi Sakamoto, " Sound power level and frequency characteristics of vehicles running on general roads measured at 20 sites nationwide," The Journal of the Acoustical Society of Japan, 2019

[22] Akemi Matsuo, Taku Itami, Jun Yoneyama, "Sound Source Proximity Detection System for Deaf and Hard-of-Hearing People Using Smartglasses Equipped with Microphone", AROB 2024

[23] Prediction o road traffic noise taking account of transient running conditios of vehicles-part 1. Relationships between running cinditions and noise radiation Yasuo Oshino, Keisuke Tsukui, Hideki Tachibana, Acoustical Society of Japan 1994, pp205-214

[24] Yasuhiro Kagimoto, Katsutoshi Itoyama, Kenji Nishida, Kazuhiro Nakadai "Evaluation of spatial source separation using NMF with multiple micro phone arrays under reverberation" SIG-Challenge 058-05

[25] ASRIT PALANKI, " Simulation of microphone inaccuracies and robustness analysis of Beamformers inside a reverberant environment ", 2012

[26] K. Nakadai, S. Masaki, R. Kojima, O. Sugiyama, K. Itoyama, and K. Nishida, " Sound source localization based on von-mises-bernoulli deep neural network, " in 2020 IEEE/SICE International Symposium on System Integration (SII), 2020, pp. 658–663

[27] N. Yalta, K. Nakadai, , and T. Ogata, "Sound source localization using deep learning models, "Journal of Robotics and Mechatronics, vol. 29, no. 1, pp. 37–48, 2017

[28] Y. Sudo, K. Itoyama, K. Nishida, and K. Nakadai, " Sound event aware environmental sound segmenta tion with mask u-net, "Advanced Robotics, vol. 34, no. 20, pp. 1280–1290, 2020

[29] Akemi Matsuo, Taku Itami, Jun Yoneyama, "360○ Sound Localization Support System for Deaf and Hard-of Hearing People Using Smartglasses Equipped with Two Microphone", SII 2024