Anomaly Detection Technique in Sound to Detect Faulty Equipment

Hisashi Uematsu, Yuma Koizumi, Shoichiro Saito, Akira Nakagawa, and Noboru Harada

Abstract

The Internet of Things has become an increasingly active research field in recent years, and it is useful for collecting information from diverse sensors that can be analyzed to detect the operating status and anomalous behavior of equipment. We introduce here an anomaly detection technique in sound that can be used to detect anomalies in equipment by analyzing sounds picked up by microphones, even in environments where special sensors cannot be installed.

Keywords: anomaly detection in sound, deep learning, noise reduction, acoustic features

1. Introduction

Suppose you use a washing machine every day, but it starts making an unusual rattling noise and then stops working. Or suppose your refrigerator starts making a strange groaning sound and then breaks down after a few weeks. Most of us have probably experienced situations such as these at some point. Maintenance procedures are often triggered when equipment starts making unusual sounds. This applies not only to household appliances but also to commercial devices such as manufacturing equipment and building air conditioning systems. In recent years, services have started to emerge whereby various sensors are used to monitor the functioning of equipment and to detect any anomalies instead of relying on workers to perform this task. This article introduces a system that can automatically determine if equipment is operating normally or has an anomaly based on the sounds it makes.

2. Difficulties in automatically detecting anomalies in sounds

It can be hard to detect anomalies from the sounds made by equipment because it is difficult to collect a large volume of anomalous sounds (i.e., sounds made when equipment is operating abnormally). Remark-able progress has recently been made with machine learning techniques such as deep learning. These techniques make it possible to learn discrimination rules as to whether the operating state is normal or anomalous by training deep neural networks (DNNs) using a huge amount of training sound data obtained from equipment operating in normal and anomalous states. However, the frequency of equipment failure in real environments is very low, and the number of ways in which equipment can fail is also very large. Thus, it is not feasible to collect a sufficient amount of training sound data corresponding to anomalous operating states. It has therefore been difficult to apply these approaches to anomaly detection in sound.

Moreover, if it does become possible to detect anomalous states based on the sounds made by equipment, then this technology is likely to be introduced in factories, where there are many other types of equipment operating alongside the equipment being tested for anomalies. All the sounds emitted by the additional equipment will constitute background noise that interferes with the collection of sounds from the target equipment. In such environments, the background noise can drown out the sounds made by the operation of the target equipment. Therefore, it is essential to reduce this background noise in order to use sound as a means of detecting equipment operating

anomalously.

In this article, we present an overview of how we overcame the two above issues, and we introduce some practical examples of anomaly detection in sound in factories and other noisy environments.

3. Anomaly detection technique in sound using only normal operating sounds as training data

It is difficult to apply conventional machine learning approaches to detect anomalies using sound, as mentioned above. Furthermore, depending on factors such as the environment where the target equipment is installed and its modes of failure, it is possible that a variety of different anomalous sounds may occur even for essentially the same anomalous operating states. In the method involving learning discrimination rules, all of these anomalous sounds have to be collected and learned, making this approach increasingly impractical.

Therefore, we used a method that only requires normal operating sounds instead of having to collect sounds from anomalous states. This method is achieved by calculating the *normalness* of acoustic features. The normalness is determined by measuring the deviation from the normal state, and the state of the machine is identified as having an anomaly when this deviation exceeds a predefined threshold. We have applied this concept to develop an acoustic feature extraction method using DNNs to increase the normalness of normal operating sounds [1, 2].

In this method, discrimination is carried out by only deciding whether or not the acoustic features of the sound to be judged are normal, which is advantageous in that it does not depend on the type of anomalous sound. This method only determines whether or not a sound is the same as a normal sound. It does not detect anomalies in a strict sense but only detects when the sound being generated is not normal. However, if it is possible to detect anomalous sounds, then it should be possible to link with systems to subsequently deal with the anomaly.

4. Noise reduction technique

It is assumed that anomaly detection in sound will be introduced in factories and other such environments. In these environments, there are many other machines running in addition to the equipment to be checked for normal/anomalous operation, so the operating sounds of the target equipment are mixed with noise consisting of the operating sounds of other equipment coming from all directions. An algorithm that detects anomalies by learning only the sounds made during normal operation can be used when the operating sounds of the equipment to be checked have been collected.

In theory, if the waveforms of the sounds generated by all noise sources are subtracted from the sound entering the collecting microphone, then only the operating sound of the target equipment will remain. This can be achieved by obtaining information including the transfer characteristics from the noise sources to the collecting microphone. However, since there are many noise sources, each with transfer characteristics that vary according to diverse factors including the shape of the room and the distance from each noise source to the microphone, it is difficult to accurately calculate the transfer characteristics of each noise source.

Rather than handling many noise sources one by one, we addressed this issue by first creating a model whereby multiple noise sources are treated as a single noise source group. That is, instead of obtaining the transfer characteristics from individual noise sources, we approximated them with a single transfer characteristic. This approximation holds because we can regard each noise source as being much further away from the sound collection microphone than the target equipment. Next, instead of strictly calculating these transfer characteristics, we approximate them in terms of the time delay and transfer gains. In this way, transfer characteristics that would previously have been calculated strictly can be modeled and estimated using only the information needed for noise reduction, thereby facilitating the noise reduction and the detection of anomalous sounds in noisy environments.

5. Experimental detection of anomalous sounds in real environments

We used the method introduced above to conduct an experiment on three machines—an air blower pump, three-dimensional (3D) printer, and water pump—to determine whether or not anomalous sounds could be detected in a real environment.
(1) Air blower pump
An overall view of the air blower pump used in the experiment is shown in **Fig. 1**, together with the microphone mounting position and a close-up view of the pump. The microphone is installed by attaching it to a pole adjacent to the pump. First, we performed preliminary learning of normal sounds by using the

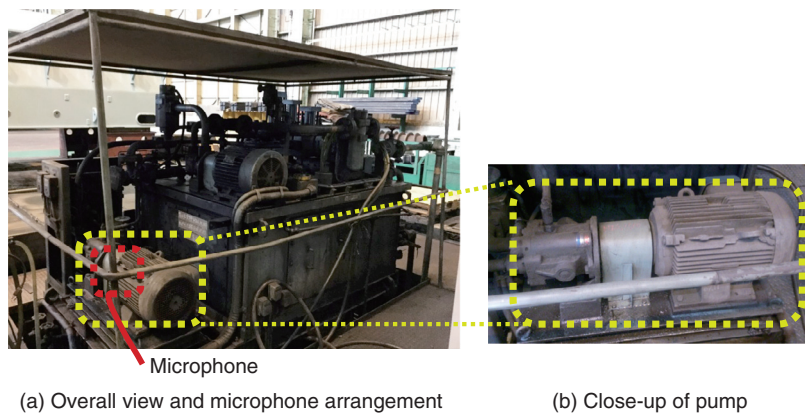(a) Overall view and microphone arrangement

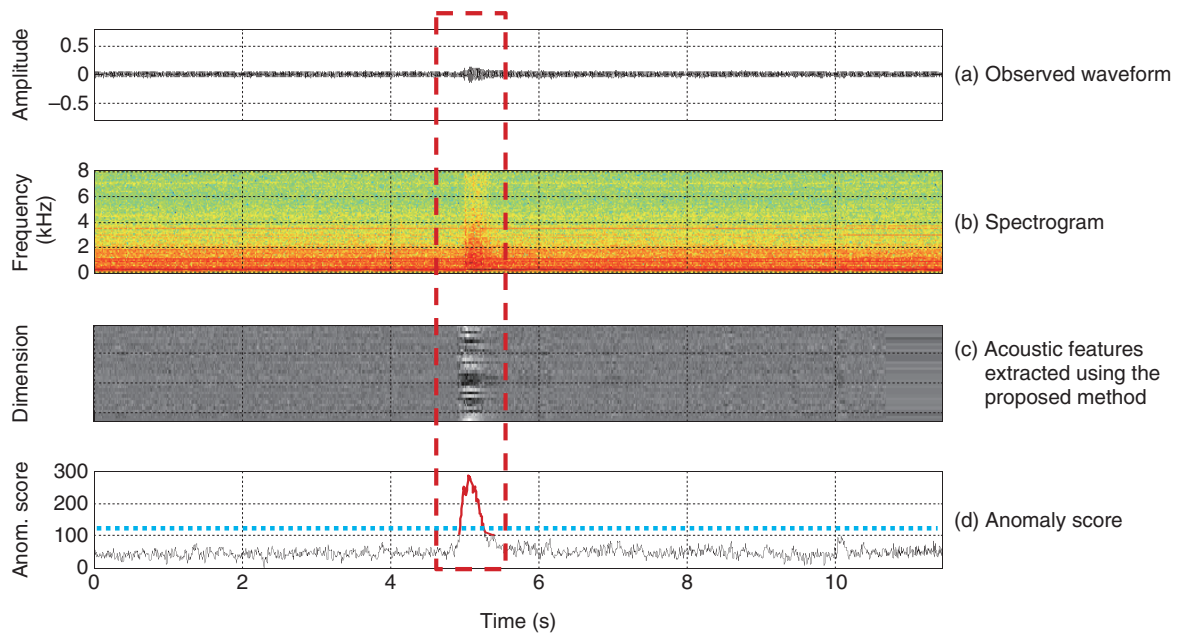(b) Close-up of pump

Fig. 1.   Air blower pump.

Fig. 2.   Air blower pump anomalous sound detection results.

normal operating sound of the air blower pump for 20 minutes.

The results of this experiment are shown in **Fig. 2**. The observed waveform (Fig. 2(a)) and spectrogram (Fig. 2(b)) clearly show that there are different waveforms at around the 5-second point. This is known to be the noise produced when a blockage occurred due to a foreign object stuck in the air blower duct. A characteristic change can also be seen in the acoustic features (Fig. 2(c)) corresponding to the period when this blockage occurred. The anomaly score (Fig. 2(d)) indicates that it was possible to detect the presence of an anomalous sound due to the foreign object blockage.

Although in this experiment we were able to ascertain the cause of the anomalous sound through constant manual observation of the air blower pump's operation, it would be impractical to constantly monitor all pumps in a real environment. Therefore, as in this method, we can expect that if it is possible to detect an anomalous state automatically, then this would yield various benefits in terms of operability
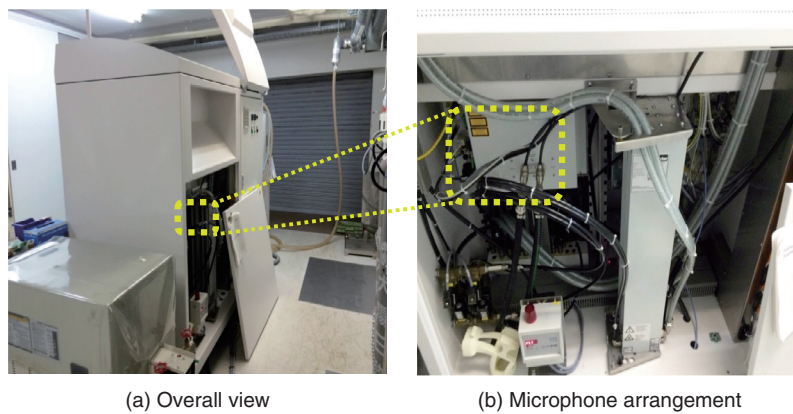
(a) Overall view (b) Microphone arrangement

Fig. 3. 3D printer.

(a) Observed waveform

(b) Spectrogram

(c) Acoustic features extracted using the proposed method
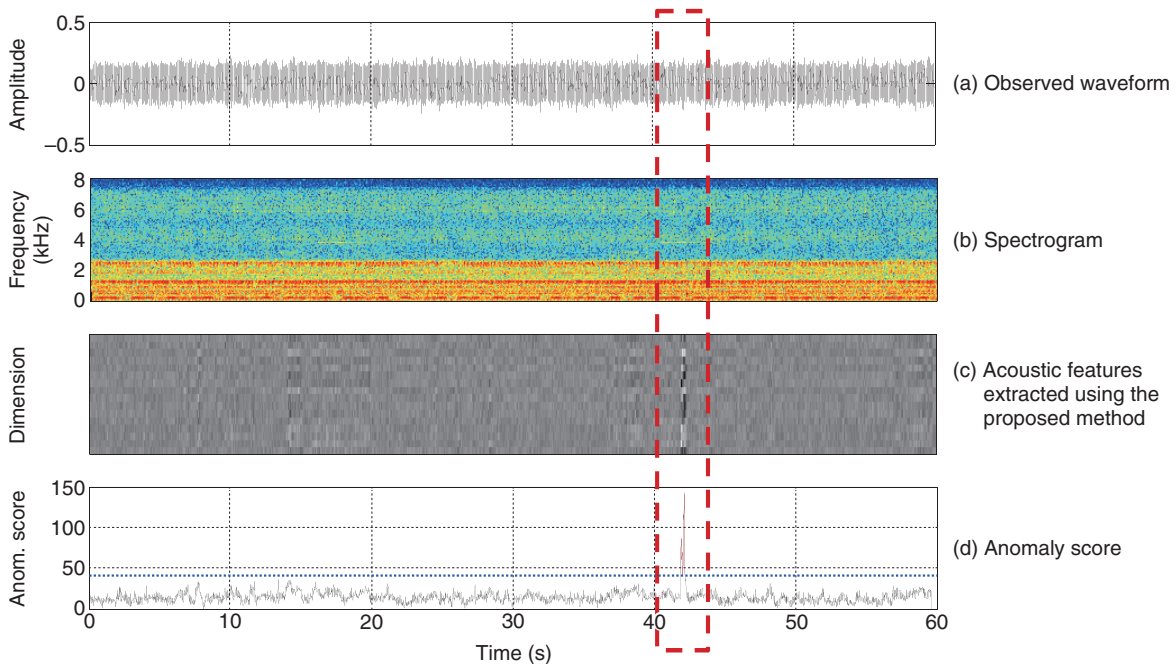
(d) Anomaly score

Fig. 4. 3D printer anomalous sound detection results.

and cost.

(2) 3D printer

An optical fabrication 3D printer was used in the experiment (**Fig. 3**). The microphone was placed inside the body of the 3D printer (the part outlined in yellow in Fig. 3) to record the operating sounds. We used 30 minutes of normal operating sounds to train the system with normal sounds.

The experimental results are shown in **Fig. 4**. The results in this experiment differed from those of the air blower pump in that the sound's observed waveform (Fig. 4(a)) and spectrogram (Fig. 4(b)) did not exhibit any changes over the entire displayed period, so it was not possible to distinguish any anomalous states from this information. However, in the acoustic features extracted using the method proposed here (Fig. 4(c)) and the anomaly score calculated based on these acoustic features (Fig. 4(d)), we observed anomaly changes at around 43 seconds (the part outlined in red). This means that a sound that does not
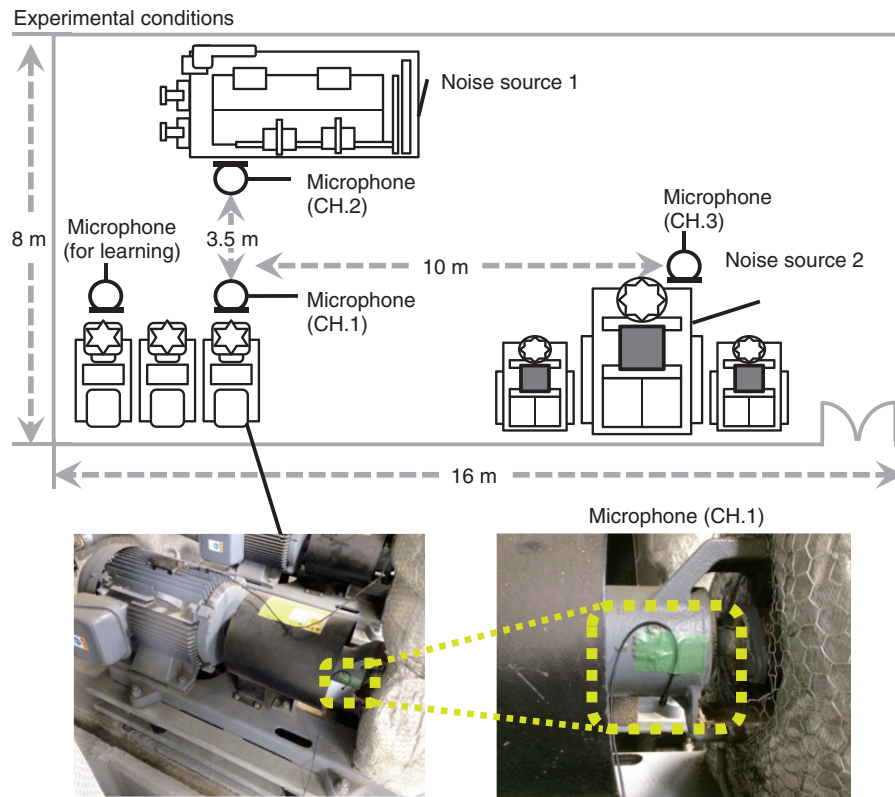
Fig. 5.   Water pump and microphone arrangement.

occur during normal operation was observed at around 43 seconds, suggesting that some sort of anomaly had occurred. In fact, the 3D printer stopped unexpectedly about 5 minutes later. Furthermore, from the results of other observations, we found that an unusual action was performed at around 43 seconds whereby the sweeper collided with the model under construction. As in the example of this result, even when the existence of anomalous sound has not been clearly found in the analysis of data such as sound waveforms and spectrograms, this proposed method can be used to clarify the existence of anomalous sounds.
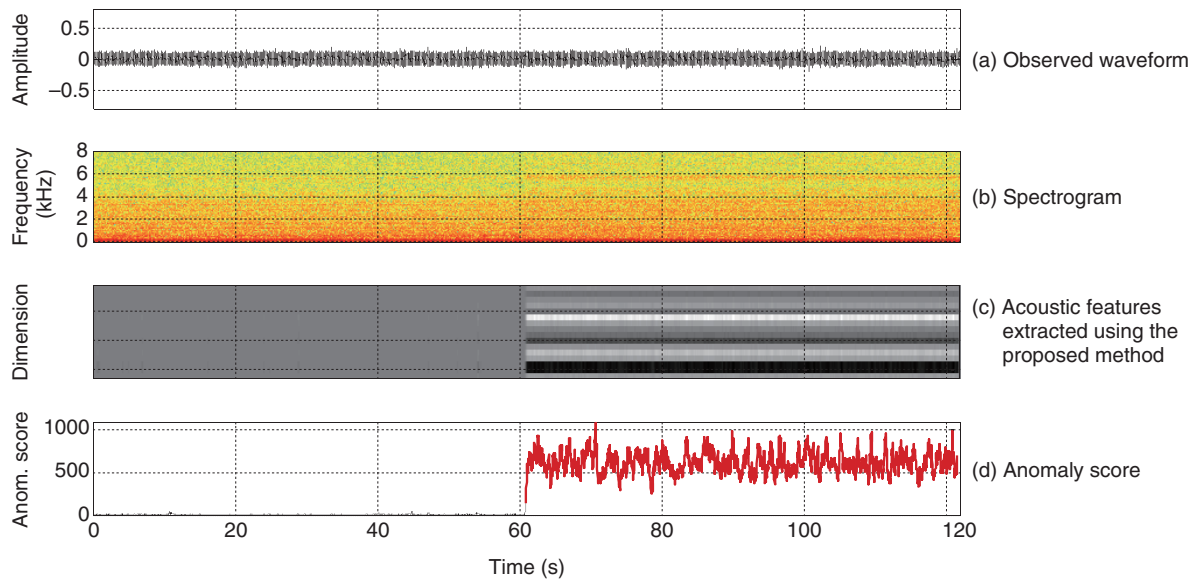
(3)   Water pump

Finally, we present an example where this method was applied to a pump used to supply water to a building. In the machine room where this water pump is installed, noise is produced by a variety of other machines in the vicinity, and the operating sound of the target water pump is masked by the noise in this environment. Therefore, in this experiment we arranged the microphone as shown in **Fig. 5**.

In the examples of the air blower pump and 3D

printer discussed above, a sudden anomalous sound was produced in the middle of normal operating sounds. With this water pump, however, anomalous sounds were continuously produced due to wearing of the pump's bearings. Thus, for the sound data for learning normalness, we used the sounds of other water pumps of the same type that were installed next to the target pump and that were operating normally.

The experimental results are shown in **Fig. 6**. In this figure, the sounds of the water pump in normal operating states for 60 seconds are shown in the first half, and in the second half these are compared with 60 seconds of sounds of the water pump in an anomalous state. In the observed waveform (Fig. 6(a)) and spectrogram (Fig. 6(b)), no major difference can be seen between the normal operating state (first half) and the anomalous operating state (second half), so from this information alone, it is not possible to distinguish between normal and anomalous states.

In contrast, there is a clear difference between the tendencies of the first half and second half with regard to the acoustic features (Fig. 6(c)) and the anomaly score (Fig. 6(d)), which clearly indicates

(a) Observed waveform

(b) Spectrogram

(c) Acoustic features extracted using the proposed method

(d) Anomaly score

The first 60 seconds are the operating sound of the water pump during normal operation, and the last 60 seconds are the operating sound of the water pump during anomalous operation.

Fig. 6.   Water pump anomalous sound detection results.

that the operating states are different. As these results show, our technique can be used to detect not only anomalous sounds that occur suddenly, as in the air blower pump and 3D printer, but also anomalous sounds that are produced continuously.

6.   Future work

In principle, the anomaly detection technique in sound introduced above can detect anomalies automatically if an equipment produces sounds that appear to be anomalous when heard by naïve listeners in an environment with hindrances such as noise. In the future, we hope to develop algorithms that can

even detect anomalous sounds that would be difficult for naïve listeners to detect but could be detected by a well-trained inspection engineer with decades of experience in this field.

References

[1]   Y. Koizumi, S. Saito, H. Uematsu, and N. Harada, "Optimizing Acoustic Feature Extractor for Anomalous Sound Detection Based on Neyman-Pearson Lemma," to appear in Proc. of the 25th European Signal Processing Conference (EUSIPCO 2017), Kos Island, Greece, Aug./Sept. 2017.
[2]   Y. Koizumi, S. Saito, and H. Uematsu, "Anomalous Sound Detection for Machine Operating Sounds using Deep Neural Networks," The 2017 Spring Meeting of the Acoustical Society of Japan, pp. 473–476, Kawasaki, Japan, Mar. 2017.

Hisashi Uematsu
Senior Research Engineer, Cross Modal Computing Project, NTT Media Intelligence Laboratories.
He received a B.E., M.E., and Ph.D. in information science from Tohoku University, Miyagi, in 1991, 1993, and 1996. He joined NTT in 1996 and has been engaged in research on psychoacoustics (human auditory mechanisms) and digital signal processing. He is a member of the Acoustical Society of Japan (ASJ).

Yuma Koizumi
Researcher, Audio, Speech, and Language Media Laboratory, NTT Media Intelligence Laboratories.
He received a B.S. and M.S. from Hosei University, Tokyo, in 2012 and 2014. Since joining NTT in 2014, he has been researching acoustic signal processing and machine learning. He was awarded the IPSJ Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ) in 2014 and the Awaya Prize from ASJ in 2017. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), ASJ, and the Institute of Electronics, Information and Communication Engineers (IEICE).

Shoichiro Saito
Research Engineer, Audio, Speech, and Language Media Laboratory, NTT Media Intelligence Laboratories.
He received a B.E. and M.E. from the University of Tokyo in 2005 and 2007. Since joining NTT in 2007, he has been researching acoustic echo cancellers and hands-free telephone terminals. He is a member of IEEE, IEICE, and ASJ.

Akira Nakagawa
Senior Research Engineer, Audio, Speech, and Language Media Laboratory, NTT Media Intelligence Laboratories.
He received a B.E. and M.E. from Kyushu Institute of Technology, Fukuoka, in 1992 and 1994. He joined NTT in 1994 and since then has been conducting research on acoustic signal processing for echo cancellation. He received a paper award from ASJ in 2002. He is a member of ASJ.

Noboru Harada
Senior Research Scientist, Supervisor, Audio, Speech, and Language Media Laboratory, NTT Media Intelligence Laboratories.
He received a B.S. and M.S. in computer science and systems engineering from Kyushu Institute of Technology, Fukuoka, in 1995 and 1997, and a Ph.D. in computer science from the University of Tsukuba in 2017. He joined NTT in 1997. His main research areas are lossless audio coding, high-efficiency coding of speech and audio, and their applications. He is an editor of ISO/ IEC 23000-6:2009 Professional Archival Application Format, ISO/IEC 14496-5:2001/Amd.10:2007 reference software for MPEG-4 ALS, and ITU-T G.711.0, and has contributed to 3GPP EVS standardization. He is a member of IEICE, ASJ, IPSJ, the Audio Engineering Society (AES), and IEEE.