Application of fractal dimension to the evaluation of environmental sound

Yoshiaki MAKABE[1]; Kenji MUTO[2]

[1, 2] Shibaura Institute of Technology, Japan

ABSTRACT

We propose an evaluation method that uses fractal dimension for the analysis of environmental sound. In previous studies, it was shown that the sense of hearing is able to identify fractal dimensions. Fractal dimension is therefore considered to be a useful parameter for the evaluation of acoustic environments. However, there are still many issues left to study, partly because fractal dimensions are not widely known about, and partly because they are affected by non-stationary sounds. In this report, we use the concept of entropy (which is widely used to evaluate quantities of information in the field of information theory) to demonstrate the effects of fractal dimensions on non-stationary sounds, and we investigate the validity of applying this knowledge to the analysis of acoustic environments.

Keywords: Environmental sound, Fractal, Entropy I-INCE Classification of Subjects Number(s): 50

1. INTRODUCTION

The analysis and evaluation of environmental sounds is essential for maintaining a favorable acoustic environment. Frequency analysis is considered to be the most common method for achieving this, and has been used in many cases including the analysis of airplane noise. One of the parameters that should be considered when aiming for a more comfortable acoustic environment is the relationship between environmental sounds and the auditory perception of these sounds. Specifically, it is known that even small sound pressure levels can cause feelings of discomfort.

We have been studying the complexity of environmental sound waveforms and their relation to acoustic environments. The complexity of waveforms can be evaluated using fractal dimension (1, 2, 3), and it has been experimentally confirmed that differences in fractal dimension can be perceived by hearing (4). Therefore, the use of fractal dimension can be regarded as an effective way of evaluating environmental sounds in that it allows their relationship to the perception of sounds to be taken into consideration.

Fractal dimensions can be evaluated based on the complexity of sound waveforms, in which case the dimension of a sound lies between the values of 1.0 and 2.0. A value of 1.0 corresponds to a sound waveform with the strongest fractal quality (self-similarity), a value of 1.5 corresponds to Brownian motion, and a value of 2.0 corresponds to white noise (5).

Although this fractal dimension can easily be used to evaluate environmental sound waveforms, it cannot be used for the evaluation of acoustic environments, and further research is still needed in this regard. Therefore in this study we evaluate waveforms based on the concept of entropy, which is widely used for the evaluation of information content in the field of information theory, and we investigate the characteristics of acoustic environments by comparing their entropy and fractal dimension. We used four types of audio source in this study: white noise, white noise with an added sine-wave signal, and sounds recorded in a boiler room and at a bus terminal. As a result of using fractal dimension and entropy to evaluate these four types of sound, we found that the evaluation scores sometimes did not change with the fractal dimension or even with non-stationary sounds. Therefore, when evaluating acoustic environments in terms of fractal dimension, their fluctuations are often observed in combination with entropy, and the importance of evaluating acoustic environments while viewing each of these fluctuations is discussed while using real analysis data.

[1] yoshiaki.makabe@gmail.com
[2] k-muto@shibaura-it.ac.jp

2. Fractal dimension and entropy of waveforms

2.1 Measuring the fractal dimension of waveforms

There are several ways of measuring fractal dimension, including the box-counting dimension, Hausdorff dimension and similarity dimension. Of these, the box-counting dimension is the easiest to apply to measurements for the analysis of environmental sound waveforms, whereas the Hausdorff dimension and similarity dimension can be computationally difficult and have a limited range of applications. For this study, we therefore decided to measure fractal dimension using the box-counting dimension discovered by Benoit Mandelbrot.

The technique is described with reference to Fig. 2.1, which shows how the fractal dimension of a waveform can be determined. First, a square grid (mesh) with a cell size of d is overlaid on the entire waveform. The shape is actually covered with squares (called "cells" or "boxes") with a diameter of $\sqrt{2} \times$d, but the factor of $\sqrt{2}$ can be ignored because we will be performing calculations with logarithms. Next, we count up the number of cells N(δ) that overlap with part of the shape. Since the cells are called boxes, this is known as the box counting method, and the resulting dimension is designated as the box-count dimension.

The same operations are repeated with a series of different values of d to prepare data on the variation of N(δ). As shown in Fig. 2.2, the results are plotted on a log-log graph of log(δ) on the horizontal axis versus log(N(δ)) on the vertical axis. Since we are only concerned with the gradient of this graph, the absolute values are not important. If a relationship of the form N(δ)=cδ −D holds, then

$$\log N(\delta) = \log c - D \log \delta \qquad (2.1)$$

resulting in a straight line on the log-log graph. The gradient of this straight line is the box-count dimension.

Normally, it is essential to ensure that the data N(δ) obtained as described above follows a straight line over a sufficiently wide range. Where possible, it is preferable to check for a linear relationship over at least two orders of magnitude on a common logarithmic scale.

Unlike mathematical objects, shapes that appear in nature do not exhibit perfect self-similarity. Since there is a limit to the range where self-similarity is satisfied, even when plotted on a log-log graph, the results are often seen to deviate from a straight line for large and small values of δ, as shown in Fig. 2.2.

For curved shaped such as Fig. 2.3, the divider method can be useful. The divider method involves using a divider with a fixed gap of d to place marks on a curved line, and then approximating the curved line by joining these marks together. The number of line segments N(δ) is counted up for different values of δ. Since environmental sound waveforms also have this sort of curved shape, in this paper we will use this method to obtain fractal dimensions.

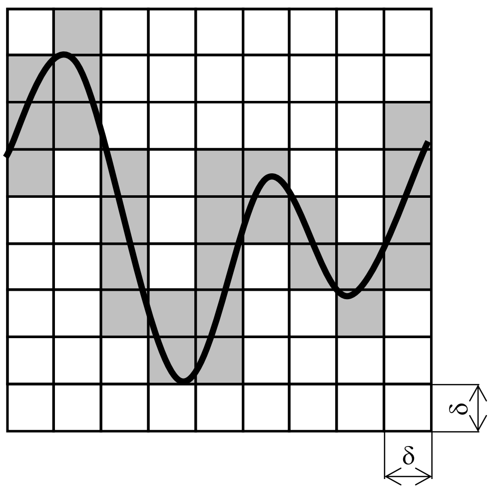


Figure 2.1 – Illustration of the box-counting method (2)

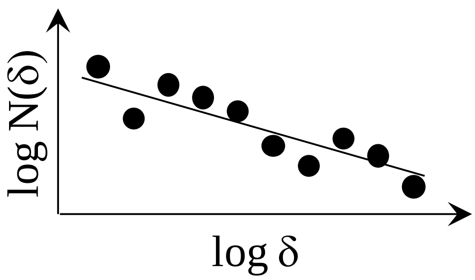


Figure 2.2 – Relationship between log(δ) and logN(δ) (2)

2.2 Entropy of a waveform

The entropy of a waveform is expressed in terms of the probability p_i of its instantaneous amplitude at time t lying within the *i*th interval. In this case, the entropy H is given by the following formula:

$$H = -\sum_{i=0}^{\infty} p_i \log_2 p_i \tag{2.2}$$

In this paper, the amplitude is normalized to the range from −1 to +1, using a distribution of 25 equal intervals. The normalized amplitude and the number of equal intervals are both arbitrary. The maximum value of the entropy occurs when the amplitude pi has equal probability of landing in any of the intervals, i.e., when it is classified with a probability of 1/25=0.04. In this case, the entropy is equal to:

$$H_{\max} = -\sum_{i=1}^{25} 0.04 \log_2 (0.04) = 4.64 \quad \text{[bit]} \tag{2.3}$$

This is the result obtained when evaluating white noise, and is equivalent to a fractal dimension of FD=2.0.

3. Fractal dimension and entropy of environmental sound

Here, we evaluate waveforms based on their fractal dimension and entropy. We focused our attention on the four types of waveform listed in Table 3.1, each of which was analyzed by evaluating 10-second audio clips with a sampling frequency of 10 kHz. Table 3.2 shows the results of evaluating the fractal dimension and entropy of these data sources.

White noise has a fractal dimension of 2.0 and entropy of 4.6 bits, which is in agreement with theory. As shown in Fig. 3.1, the amplitudes are uniformly distributed.

White noise with an added sine-wave signal simulates a non-stationary waveform. Since the intervals where the sine wave is applied occur regularly, the entropy is reduced to 3.85 from the value of 4.64 for white noise. Although the fractal dimension does not vary, the appearance of the graph shown in Fig. 2.2 is very different, and the amplitude distribution is also very unlike the graph for white noise shown in Fig. 3.2. Since the addition of the sine wave did not change the fractal dimension, it can be said that entropy is an effective way of judging the stationary nature of sounds.

Next, we looked at the boiler room noise. This is a real environmental sound that was selected as a representative example of a steady waveform. The fractal dimension and entropy both have constant values in all sections. Figure 3.3 shows a histogram of the distribution of amplitudes. In the boiler room noise, there are no large changes in the histograms of each time period, and there is a continuous distributed state close to a sinusoidal distribution.

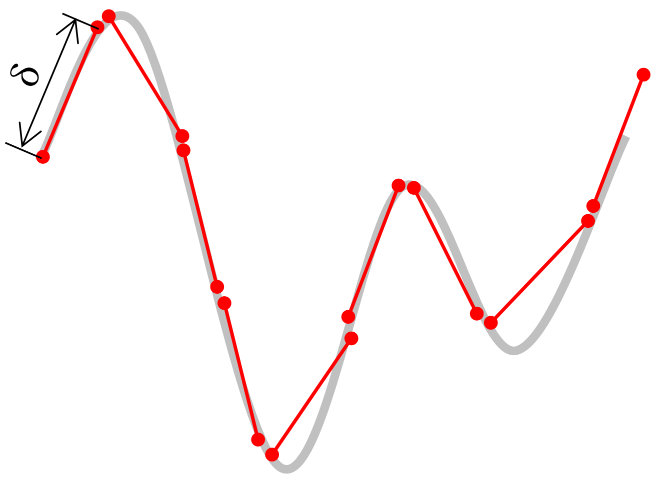


Figure 2.3 – Illustration of the divider method (2)

Table 3.1 – Signals used in the analysis (1/2)

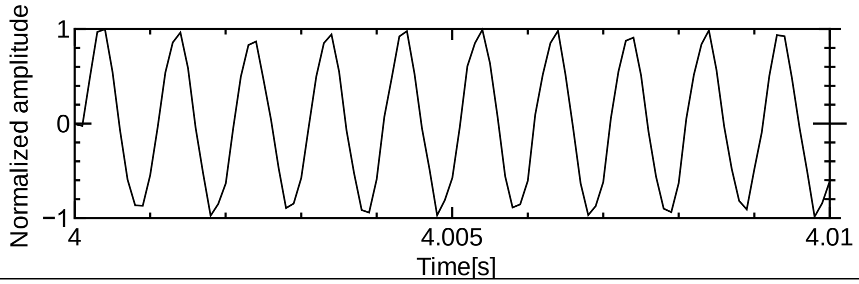
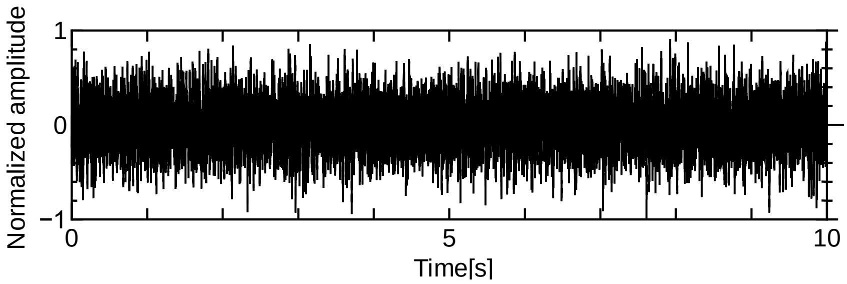
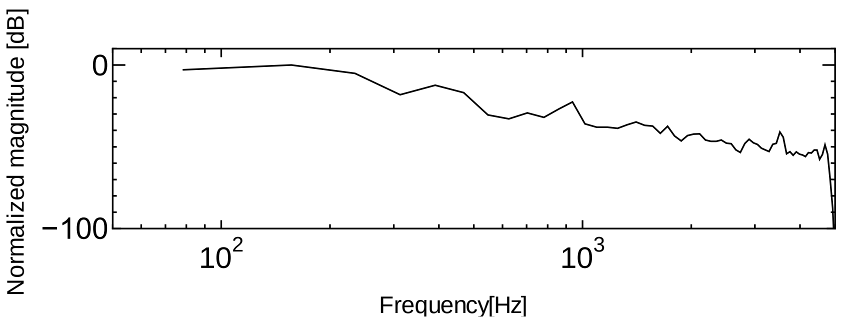
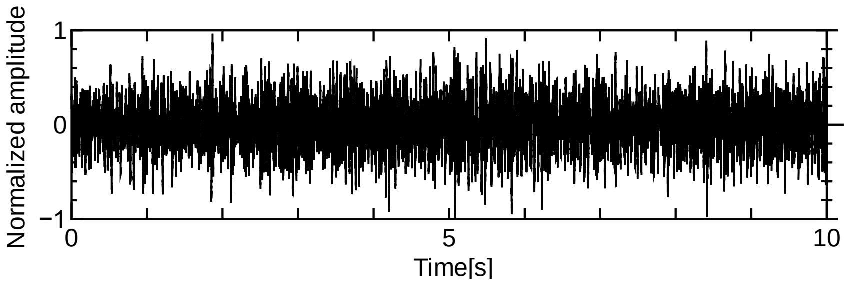
No.	Signal	Features, etc.
1	White noise	Uniform white noise generated by a LabVIEW7 (National Instruments).
2	White noise (with added sine wave)	As No. 1 above, with a pure sine wave added in bursts of 3–8 s. The sine waves had an amplitude of10 times of white noise and a frequency of 1 kHz. However, these settings were made arbitrarily. 
3	Boiler room noise	Indoor sounds recorded close to a boiler. From the Advanced Telecommunications Research Institute International environmental sound database, vol. 2 (indoor environmental sounds). The 48 kHz source audio was downsampled to a sampling rate of 10 kHz. For the evaluation, we used the first 10 seconds of this recording.    

Table 3.1 – Signals used in the analysis (2/2)

<table>
<tr><td>4</td><td>Bus terminal sounds</td><td>Outdoor sounds recorded near a bus terminal. From the Advanced Telecommunications Research Institute International environmental sound database, vol. 3 (outdoor environmental sounds). The 48 kHz source audio was downsampled to a sampling rate of 10 kHz. For the evaluation, we used the first 10 seconds of this recording.



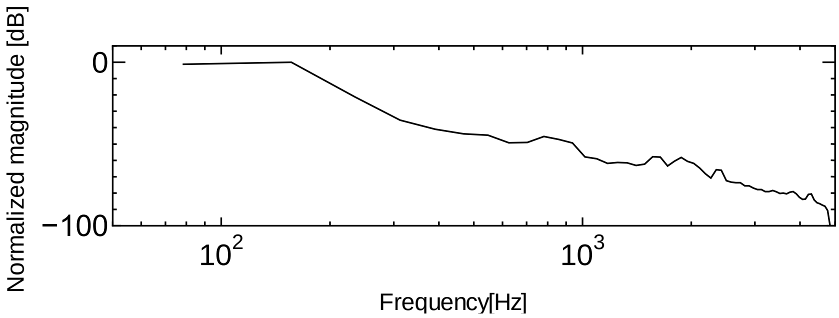

</td></tr>
</table>

Table 3.2 – Calculation results

No.	Signal name	Evaluation method	0–1 s	1–2 s	2–3 s	3–4 s	4–5 s	5–6 s	6–7 s	7–8 s	8–9 s	9–10 s
1	White noise	Fractal dimension	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
		Entropy [bit]	4.64	4.64	4.64	4.64	4.64	4.64	4.64	4.64	4.64	4.64
2	White noise (with added sine wave)	Fractal dimension	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
		Entropy [bit]	4.64	4.64	4.64	3.85	3.85	3.85	3.85	4.64	4.64	4.64
3	Boiler room	Fractal dimension	1.65	1.66	1.65	1.65	1.66	1.66	1.66	1.62	1.62	1.65
		Entropy [bit]	3.94	3.92	3.69	3.90	3.90	3.86	3.88	3.67	3.82	3.74
4	Bus terminal	Fractal dimension	1.43	1.48	1.49	1.57	1.50	1.44	1.47	1.51	1.70	1.68
		Entropy [bit]	3.95	3.82	3.64	3.92	3.45	4.16	3.68	4.01	3.50	3.92

The bus terminal noise consisted of non-stationary signals including the generation of sudden sounds over its entire duration, such as the footsteps of passengers coming and going, and the hiss of air from the buses themselves. This was reflected both in the fractal dimension and in the entropy, which were not constant. However, the behavior mechanisms of the fractal dimension and entropy are unrelated, and further investigation is needed.

As in the case of the boiler room, the amplitude distribution is shown in Fig. 3.4. It can be seen that the distribution states differ widely from one second to the next. In fact, when we listened to one-second units of this sound source, we found that there were a number of features in each section. These features are listed in Fig. 3.4. Outdoor sounds such as this contain added non-stationary noise, so if we only determine their fractal dimension then it can be difficult to extract local features including the generation of sudden sounds such as these.

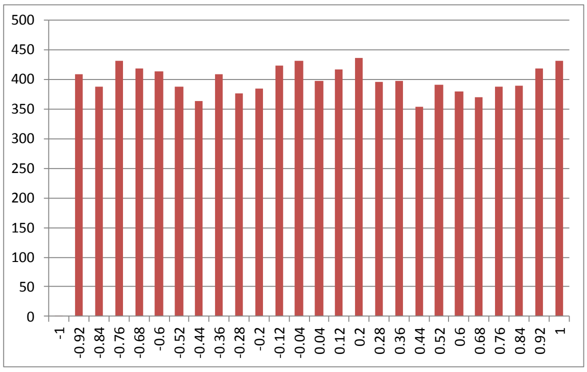


Figure 3.1 – Amplitude distribution of white noise (Horizontal axis : amplitude / Vertical axis : number of occurances)

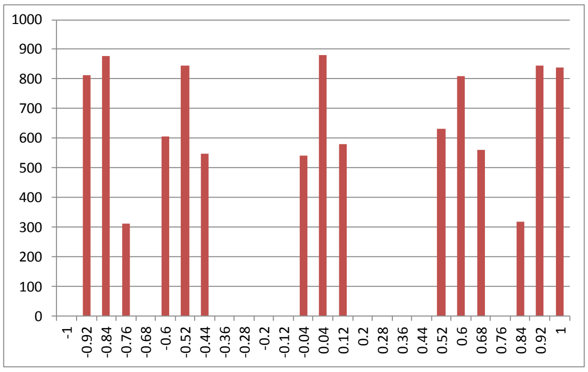


Figure 3.2 – Amplitude distribution at places with added sine-wave signal (Horizontal axis : amplitude / Vertical axis : number of occurances)

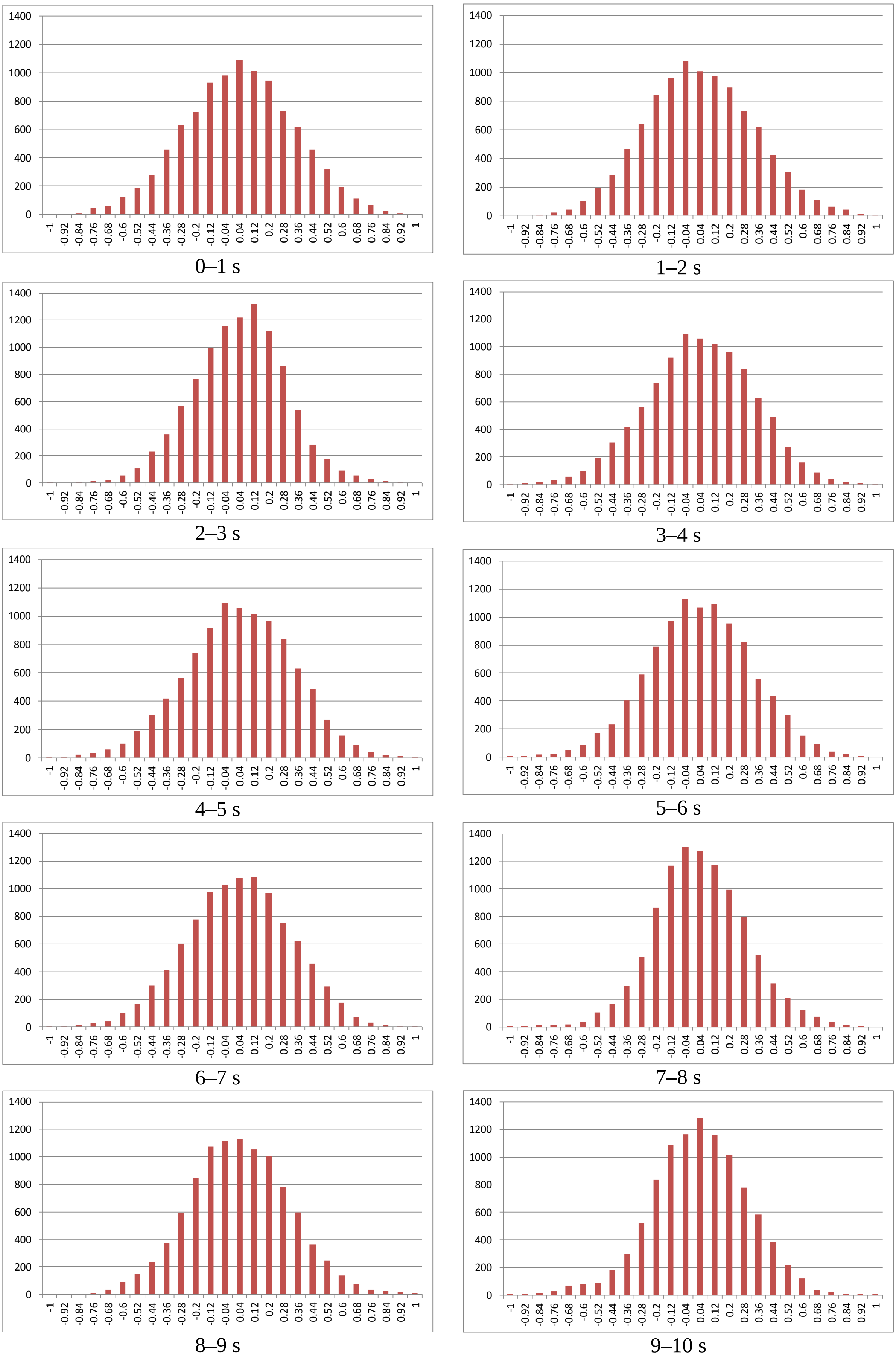
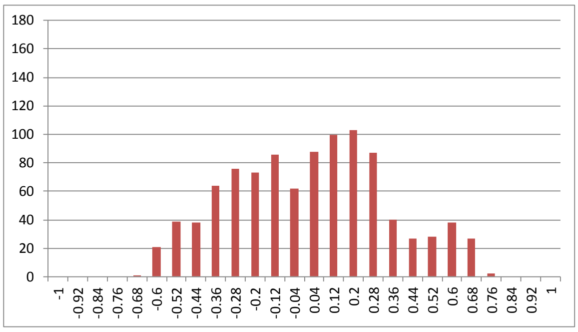
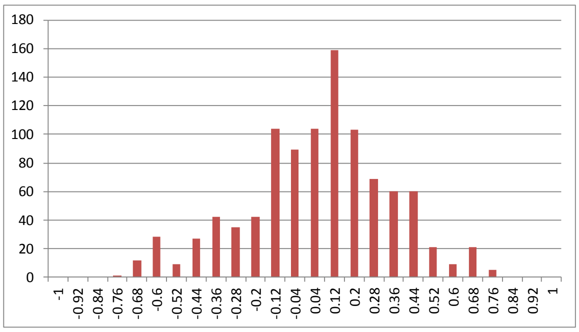


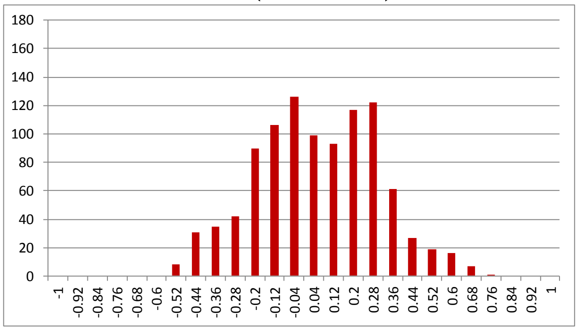
Figure 3.3 – Amplitude distribution of boiler room noise
(Horizontal axis : amplitude / Vertical axis : number of occurances)



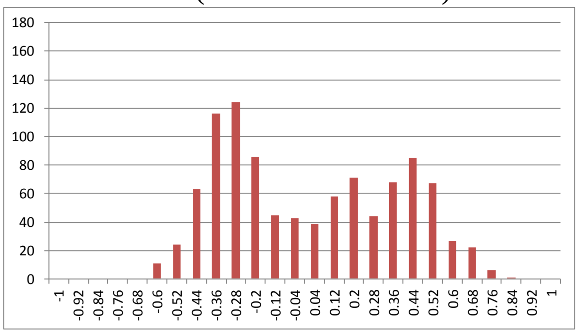
0–1 s (no features)



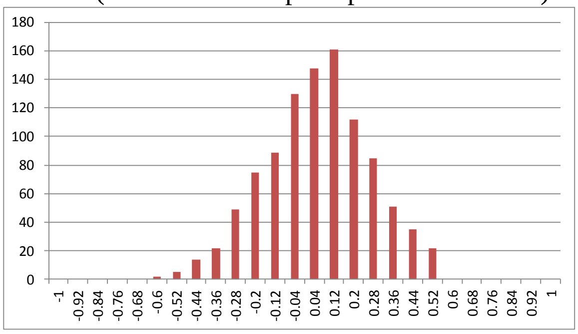
1–2 s (hiss of air from a bus)



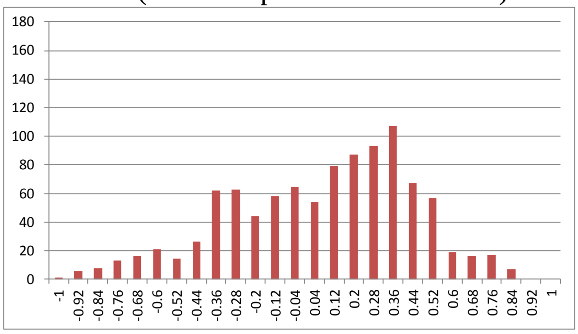
2–3 s (Feature: Footsteps of person in sandals)



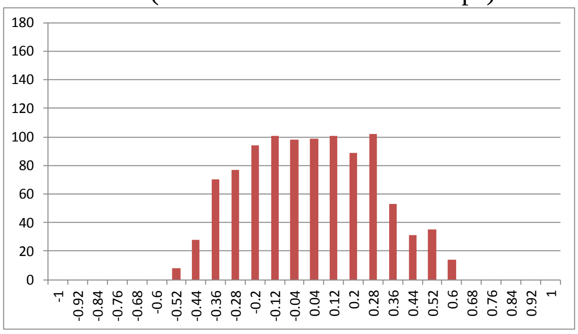
3–4 s (Feature: Spoken announcement)



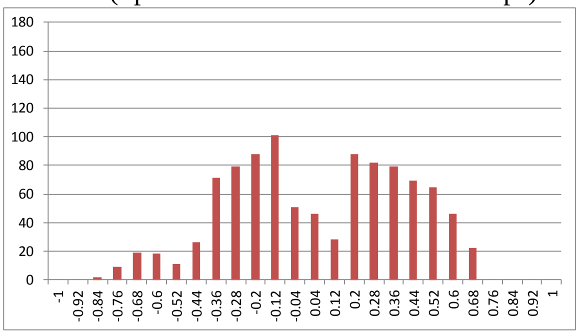
4–5 s (Feature: Voices and footsteps)



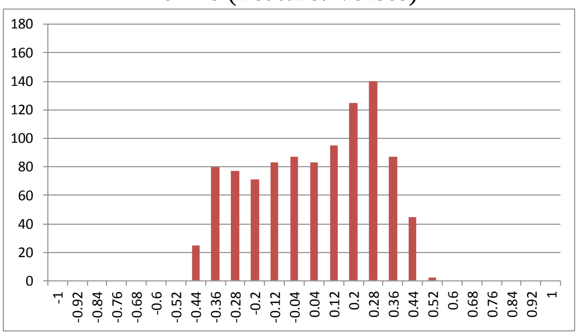
5–6 s (Spoken announcement and footsteps)



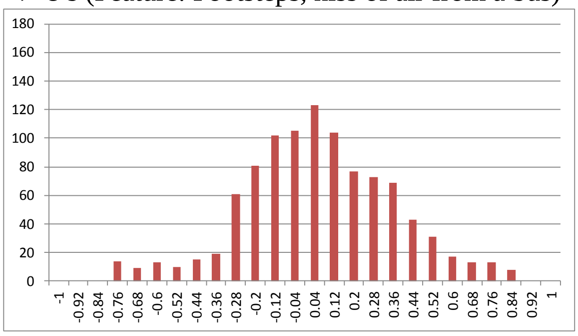
6–7 s (Feature: Voices)



7–8 s (Feature: Footsteps, hiss of air from a bus)



8–9 s (Feature: Air hiss)



9–10 s (Feature: Air hiss, voices)

Figure 3.4 – Amplitude analysis of bus terminal
(Horizontal axis : amplitude / Vertical axis : number of occurances)

4. Discussion

When analyzing environmental sounds, it is not possible to tell if the sounds are stationary or non-stationary. In most cases, there is likely to be a mixture of stationary and non-stationary sounds, like the bus terminal sounds discussed in this report. To grasp the characteristics of this sort of acoustic environment, it seems that the analysis of short time periods as shown in this report may help in grasping the characteristics of acoustic environments if repeated over extended periods. It is also necessary to accumulate more data and investigate the relationships between fractal dimension, entropy and locality. Furthermore, in the analysis of bus terminal sounds, we found that the fractal dimension and entropy undergo different fluctuations. This shows that fractal dimension reflects fractal qualities, i.e., the extent to which similar shapes are repeated, while entropy reflects the uncertainty of noise, i.e., the degree of impossibility of predicting future amplitudes. In previous studies, fractal dimension could be identified by hearing, but no studies have sought to ascertain changes of entropy. If it is also possible to recognize changes of entropy by hearing, then this will become a useful measure for the evaluation of environmental sounds.

5. Conclusion

Evaluation by fractal dimension is one way of capturing the characteristics of environmental sound. We have found that it is affected by the continuity of the acoustic environment. Here, by evaluating two types of real environmental sounds, we have found that the fractal dimension and entropy were both constant for the stationary boiler room noise, but fluctuated for the bus terminal sounds. In particular, when an amplitude distribution diagram was presented as an explanation for entropy fluctuations, there were found to be large differences in the distributions evaluated for each one-second segment, and these segments were also perceived to have different characteristics when checked by listening. Changes such as these tend to destabilize the entropy. Fractal dimension is thought to be a very useful measure for numerically expressing the complexity of a waveform. In the future, rather than proposing entropy as a new measure of auditory quality, we need to study how fractal dimension can be used in the construction of effective methods, and then see how entropy can be used as a parameter in these methods.

REFERENCES

1. Katsuya Honda, Fractals (Asakura Shoten, Tokyo, 2002), pp. 21–29 (in Japanese).
2. Jens Feder, Fractals, Plenum Press, New York, 1988.
3. Ikuo Matsuba, Nonlinear Time Series Analysis (Asakura Shoten, Tokyo, 2002), pp. 73–75 (in Japanese).
4. Yoshiaki Makabe, Hideo Shibayama and Tomohiro Okubo, Creating membership functions in the auditory domain by using Weierstrass-Mandelbrot functions with modified fractal dimension, Institute of Electronics, Information and Communication Engineers, Vol. J88-A, No. 1, pp. 91–95 (2005) (in Japanese).
5. Ikuo Matsuba, Nonlinear Time Series Analysis (Asakura Shoten, Tokyo, 2002), pp. 81–92 (in Japanese).