1. INTRODUCTION

Human senses are nonlinear in nature to physical quantities, and follow Weber Fechner’s law. Therefore, while studying sensitivity, it is important to investigate consciousness and cognition of a weak stimulus. It is the same for tones. Humans perceive small differences in tones and receive information regarding the sound’s source. The evaluation of noise and tone, which includes the evaluation of musical instruments and audio equipment, such as speakers and amplifiers, can be performed. Our goal is to investigate what types of relations exist between the sensory qualities and the physical quantities of a tone.

The sensory expression of a tone can be described with adjectives, such as “soft sound” and “warm sound.” These properties of tones have been studied for many years, which included approaches using tone expression words [1]. Studies about expression words have been performed in Japanese as well [2-4]. Moreover, the relation between expression words using a synthetic sound has been investigated [5].

Although various studies on tones consider loud sounds, the relationship between tones and the physical quantities that humans perceive are unclear.

Furthermore, few differences are observed even if there are human-perceptible differences in a tone for an amplifier (in audio equipment) [13]. That is, humans can perceive the difference in a tone by the difference in very minute sounds that are difficult to measure. Moreover, since specialist evaluations of audio equipment may agree, it is possible that certain common physical quantities are perceived.

The human hearing characteristic is nonlinear, and a minute sound is greatly amplified when perceived. Therefore, a minute sound is assumed to have the same influence which it has on a tone. So, we believed that the method of adjusting a minute sound would be more effective than that with the adjustment by alignment amplification (like the graphic equalizer as a tone method of preparation).

Until now, it has been considered effective to display physical quantities using a sensory scale, to investigate the relation between the physical characteristics and sensory evaluation of sounds by humans. Therefore, the method of transposing the amplitude value of a sound into “stev” and displaying the sound using a waveform was proposed; following which, studies confirmed that the vision and hearing information were in agreement [14]. Furthermore, the amplitude compression/expansion (ACE) method for...
changing the tone quality of minute sounds in certain frequency bands was developed, and tone evaluation was performed [15]. However, because the method changed relatively loud sounds, the influence of minute sounds on tone evaluation could not be investigated.

Therefore, in this study, the threshold amplitude compression/expansion (ThACE) method was suggested. The ThACE method uses a threshold value at which only minute sounds (sound of 50 dB or less) are changed. The difference in the evaluation between sound sources under tone adjustment was shown using the Scheffe method of paired comparison. As a result, when a minute sound influenced sensuous evaluations and elongated 5300–6400 Hz, it turned out “soft”, which would be thought as “metallic” if compressed. Thus, it became clear that a change of minute sound influences a tone.

2. TONE CONTROL METHOD

2.1 ACE method

Processing techniques similar to the ACE method were developed early by Goldstein and others. Goldstein proposed an amplification circuit for hearing-aids [16]. In this study, ACE is used for tone control.

Figure 1 shows the signal processing flow for tone control. The frequency band that adjusts the tone uses a band pass filter (BPF). The amplitude is adjusted for the selection using compression or expansion. The odd-harmonic distortion that arises because of the compression or expansion is removed by a low pass filter (LPF) in the next stage. The signal, excluding the signal in the frequency band that undergoes tone control by the band stop filter (BSF), is added and obtained as output.

When the instantaneous value that is extracted from the signal of the band, which undergoes tone control by the BPF and was normalized at maximum, is set to \( v \), the signal after processing is given by the following equation:

\[
\text{Eq. (1)} \quad w = \text{sgn}(v)c|v|^n
\]

where, \( \text{sgn}() \) is the signum function, \( c \) is a linear amplification coefficient, and \( n \) is a compression/expansion coefficient. In a hearing aid, \( 0 < n < 1.0 \) holds, and the amplitude band is compressed. On the contrary, the amplitude band is expanded if \( n > 1.0 \). The input-and-output characteristics of the ACE method up to 100 dB, using the sound pressure level (SPL), are shown in Figure 2, when using Eq. (1). The solid line indicated by ① is the condition of \( n=1.0 \), i.e., the condition at which the ACE method is not performed. The dashed lines indicate the positive and negative cases of linear amplification; they represent adjustment by +20 dB and -20 dB, respectively. The solid lines of ② and ③ are the cases in which the signal is compressed at \( n=0.8 \) and \( n=0.9 \), respectively; these lines indicate that a minute sound is more strongly amplified (relatively speaking). The solid lines of ④ and ⑤ are the cases in which \( n \) increases to \( n=1.1 \) and \( n=1.2 \), respectively, which further decreases the minute sound. Therefore, it becomes impossible, before processing, to hear parts of the sound in the auditory region.

Thus, by changing the compression/expansion coefficient \( n \), a control of minute sounds is possible. Goldstein had proposed this technique for the development of hearing-aids, and it was observed to be effective.

However, changes in a sound at a comparatively high amplitude cannot be disregarded in this method. Therefore, the processing needed to be changed when...
it aimed at investigating a change in a minute stimulus. So, the ThACE method shown in the following paragraph is proposed in this study.

2.2 ThACE method

For human hearing, sensitivity falls, and the perceived volume of the sound rapidly decreases below 40 dB SPL. Therefore, in recent years, it was common to reduce the gain of automatic gain control circuitry in the amplitude band below 40 dB, in digital hearing-aids. Goldstein established a threshold value, carried out the compression process of the level above the threshold value, and proposed a method of alignment for the lower level [17]. Although using the Goldstein method was effective in hearing-aids, it was not suitable for tone control. Here, to make changes in only minute sounds and perform tone adjustment, the amplitude compression elongating method in Eq. (1) was changed as follows:

\[
w = \text{sgn}(v) c \left| \frac{v}{v_{th}} \right|^n
\]  

(2)

Since only minute components were to be changed, this method was referred to as \( n = 1.0 \) when \( v > v_{th} \). The values of \( n \) were adjusted in the amplitude-band to be not more than one. The input-and-output characteristics are shown in Figure 3. The solid lines of ② and ③ are the cases in which \( v \) has a value equivalent to 50 dB SPL; below this point, the values of \( n \) were set to \( n = 0.6 \) (compression) and \( n = 1.4 \) (expansion), respectively. It is possible to change only a minute component, without changing the loud component of a sound by employing this method.

When the ThACE method was applied, as described previously, odd-harmonics distortion occurred. To remove odd-harmonics distortion, one must apply the ThACE method, after restricting the frequency band using a BPF and removing the distortion using an LPF. In this study, the filter bank shown in Table 1 was used.

### Table 1: Characteristics of the filter bank

<table>
<thead>
<tr>
<th>Central frequency [Hz]</th>
<th>Cut-off frequency [Hz] (low)</th>
<th>Cut-off frequency [Hz] (high)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>150</td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>250</td>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td>350</td>
<td>300</td>
<td>400</td>
</tr>
<tr>
<td>450</td>
<td>400</td>
<td>510</td>
</tr>
<tr>
<td>570</td>
<td>510</td>
<td>630</td>
</tr>
<tr>
<td>700</td>
<td>630</td>
<td>770</td>
</tr>
<tr>
<td>840</td>
<td>770</td>
<td>920</td>
</tr>
<tr>
<td>1000</td>
<td>920</td>
<td>1080</td>
</tr>
<tr>
<td>1170</td>
<td>1080</td>
<td>1270</td>
</tr>
<tr>
<td>1370</td>
<td>1270</td>
<td>1480</td>
</tr>
<tr>
<td>1600</td>
<td>1480</td>
<td>1720</td>
</tr>
<tr>
<td>1850</td>
<td>1720</td>
<td>2000</td>
</tr>
<tr>
<td>2150</td>
<td>2000</td>
<td>2320</td>
</tr>
<tr>
<td>2500</td>
<td>2320</td>
<td>2700</td>
</tr>
<tr>
<td>2900</td>
<td>2700</td>
<td>3150</td>
</tr>
<tr>
<td>3400</td>
<td>3150</td>
<td>3700</td>
</tr>
<tr>
<td>4000</td>
<td>3700</td>
<td>4400</td>
</tr>
<tr>
<td>4800</td>
<td>4400</td>
<td>5300</td>
</tr>
<tr>
<td>5800</td>
<td>5300</td>
<td>6400</td>
</tr>
<tr>
<td>7000</td>
<td>6400</td>
<td>7700</td>
</tr>
<tr>
<td>8500</td>
<td>7700</td>
<td>9500</td>
</tr>
<tr>
<td>10500</td>
<td>9500</td>
<td>12000</td>
</tr>
<tr>
<td>13500</td>
<td>12000</td>
<td>155000</td>
</tr>
</tbody>
</table>

3. CONDITIONS OF EVALUATION

3.1 Evaluation of the sound source

In this study, a comparison of an audition was performed using a sound source that applied ACE to only one frequency band among 25 frequency bands. The compression/extension coefficient \( n \) had three values (i.e., \( n = 0.6, 1.0, \) and 1.4). First, the sound source of SACD (Super Audio CD) was downloaded to a PC. The SACD data was sampled from the analog output of an SACD player (Marantz DV-12S2) at 24 bits and 96 kHz using an A/D converter (Roland UA-25). Subsequently, the data was uploaded to a PC. The frequency band that performs tone control for the sound source was extracted using a BPF (\( f_L - f_H \)). ACE processing was performed, and the signal was removed by an LPF (cut off frequency: \( f_H \)). Then, the signal was added to the signal excluding the frequency band (\( f_L - f_H \)) blocked by the BSF, which was considered as the evaluation sound source.

The original sound source was used for 3 s after the start of Hélène Grimaud’s “Cred” (SACD Deutsche Grammophon 028947486923: only piano sound). Beforehand, we tried listening to the ThACE processing sound source. The ThACE processing of the frequency band of 5300-6400 Hz, for sounds described as “soft” or “metallic,” was
performed to create the evaluation sound source. The evaluation word pair was confirmed and defined by a screening test. Creation of the evaluation sound source performed by exponential function processing below the threshold value $v_{th} = 50$ dB. Figure 4 shows an original sound source waveform. As compared with the amplitude of the original sound source, the amplitude of the sound in this frequency band (5300-6400 Hz) is quite small. Figure 5 shows the waveform of the limited frequency band after ThACE processing. Using the ThACE method, minute sounds are greatly amplified as compared to loud sounds.

3.2 Evaluation method

To statistically process an evaluation result, an examination that uses the Scheffe method of paired comparison was used. However, for paired comparison evaluating a pair of sound sources, the standard is not clear; this method of evaluation of a tone does not allow easy judgment.

The preliminary examination, in which the frequency component of 5300-6400 Hz was changed with the graphic equalizer, was conducted. The results confirmed feeling it as “metallic”, when carrying out the 4 dB alignment amplification with the volume of about 60 dB, and feeling it as “soft”, when below 4 dB. In the examination using a graphic equalizer, since the large parameter of sound pressure is changed, its influence is also considered. Therefore, the evaluation method investigated the influence of changes in a minute sound by the proposed ThACE method in detail.

The evaluation words were “soft” and “metallic”. The differences in the tones of the prepared evaluation sound sources were very small, and difficult to differentiate after a single listening in many cases. Therefore, the evaluation sound sources were presented to examinees to listen repeatedly. That is, the effect of order is not taken into consideration. A total of three comparative evaluations were performed for one pair of evaluation words. The rating scale was a five-step evaluation from -2 to 2. If evaluation sound source B was close to an evaluation word for sound source A, evaluation sound source A would have been considered close to an evaluation word for sound source B, which was considered a drawback.

Because the tone evaluation was a Kansei measurement, it needed to consider the change of decision-making based on experience [18]. Therefore, the method was unified by explaining the meaning of each evaluation word to the subject before the examination. However, examinees that could not perform the exact evaluation might have been counted. Therefore, we decided to investigate the significance of individual differences, and the existence of a circular triad in the analysis of Scheffe method. When a disqualified examinee became clear, in this case, we decided to exclude the subject from the data.

4. RESULT AND DISCUSSION

The listening test was administered in a soundproof chamber, with environmental noise less than 20 dB (A). The evaluation sound source was written in DVD-Audio (DVD-A). The player used was a Marantz DV-12S2. The amplifier was an Accuphase E530. The headphones were Sennheiser HD650. There were six subjects, aged between 22–24.

The evaluation results for evaluation words (“soft” and/or “metallic”) were analyzed using analysis of variance (ANOVA). For the evaluation words, because the main effect was significant at 5%, results of by an individual difference were not significant. Figure 6 indicates that the standard of 5% significance level was measured for each result. Results showed that “soft” was applied if the amplitude expansion was carried out for the frequency band of 5300-6400 Hz. Moreover, “metallic” was applied when the amplitude compression was carried out for the
Nonlinear Tone Control Method by ThACE for Evaluation of Minute Sounds

frequency band of 5300-6400 Hz. As mentioned above, it was possible to distinguish the tone of the evaluation sound source using the ThACE method. Therefore, results showed that the differences in minute sounds influence the perceived tone.

5. CONCLUSION

Human senses are nonlinear in nature to physical quantities, and a change in a minute stimulus is considered to bring a change in any senses. This study focused on hearing, and developed and investigated the ThACE method, which presented the influence of minute sounds on consciousness.

In this study, the ThACE method was used and a tone control method for minute sounds was developed to analyze the relation between the physical differences in minute levels of electric sound reproduction and the sensory evaluation. Evaluation and analysis were conducted using the Scheffe method of paired comparison and a processing method for an evaluation sound source that adjusted the piano sound at a minute level in several frequency bands.

As a result, by controlling the sound at the minute level for a specific frequency band, it was possible to change the tone. Results confirmed that this method was effectively able to analyze minute sounds and the properties of a tone. As future research, we will evaluate the reproduction of sounds, other than those of a piano, and the case in which the processed frequency band is changed.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Dr. Itsuo Ishikawa, a former professor at the University of Yamanashi for his advice and cooperation.

REFERENCES


