Acoustics'08
Paris
June 29-July 4, 2008
www.acoustics08-paris.org

An Investigation of Width and Depth Perception toward a Sound Image Constructed of Multiple Variant Sound Waves Emitted from a Loudspeaker Array

Yoko Yamakata, Toshiyuki Kimura, Munenori Naoe and Michiaki Katsumoto

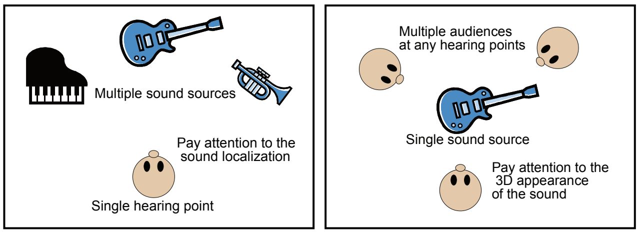
National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi,
Koganei, 184-8795 Tokyo, Japan
yamakata@nict.go.jp

Many musical instruments, including violins and guitars, vibrate their resonant bodies differently over their surfaces while making a sound. This paper aims to reveal the influences of such vibration variations of a soundboard surface on the width and depth perceptions of a sound image when listeners are present in the near-field—at a distance of 50 cm or 100 cm from the soundboard. In this paper, a loudspeaker array mimics surface vibrations as each loudspeaker makes a corresponding sound independently and cooperatively. Three types of sounds—synthesized single tone, double tone and instrumental—were used as the sources. To determine what factors affect the perception of the sound image, various test sound sets were prepared by varying the amplitude or delay of an original sound set for each frequency for each loudspeaker. On the basis of Scheffe's pair comparison method, eleven subjects were asked to identify which sound image in a pair of test sounds was wider or farther than the other. The results show that a test sound set with delay variation, which mimics the sounds emitted by bending vibrations propagating on a soundboard, obviously influences the perception of the sound image width and that the amplitude variation does not have much influence.

1 Introduction

Many researches about sound-field reproduction have aimed to reproduce a sound field of wide spaces such as an entire room or musical hall where multiple sounding objects are sparsely positioned (as shown in Figure 1(a)). In such cases, it is inherently assumed that only one listener listens to the sound at a specific spot (or at least in a small area). On the other hand, our research aims to reproduce the entire sound radiated from a single sounding object towards all directions—front, back, up and down (as shown in Figure 1(b)). Many listeners can simultaneously listen to the sound at any point and from any angle.

Therefore, we are constructing a new audio reproduction device that can radiate the same sound with the original sound in all directions [1].



(a) **Common sound-field reproduction**: reproduce sounds of wide sound fields at one healing spot.

(b) **Our purpose**: reproduce entire radiated sound of a single sounding object.

Figure 1: Different purposes of sound-field reproduction.

Most studies on audio reproduction devices have assumed omnidirectional sounds. However, in reality, most instruments, including violins and guitars, never emit such omnidirectional sounds [2, 3]. When a sound is produced, its resonant body vibrates differently over its surface. The emitted sound from each place of the surface interferes with the emitted sounds from other places, and then the total sound radiated from the soundboard has various spectral patterns that vary in direction and frequency. This sound directivity strongly depends on the sounding object characteristics such as 'size', 'shape', 'material' and so on [2, 4]. In the near-field—a distance of 50 cm or 100 cm from a sounding object, the sounds appearing at the right and left ear positions have obviously different frequency spectral patterns. It is possible that human beings can learn the relationship between radiated sound directivity and the characteristic of the sounding object. Correspondingly, can human beings perceive such characteristics of a sounding object just by listening to the radiated sound? If so, how much they can?

To answer these questions, in this paper, we investigate human perception towards vibration variations on a soundboard surface, particularly in the near-field from the surface.

To control the surface vibrations of a sounding object, a loudspeaker array is used to mimic the sounding object as each loudspeaker emits a sound at the corresponding position independently and cooperatively. The loudspeaker array consists of 21 speaker units that are set on 3 × 9 matrix positions with 4-cm intervals on a front panel and 7 units of middle low, except the units of both ends, are used because these speaker units are entirely surrounded by speaker units and then their frequency responses are almost same.

Since we consider that the object size exerts the strongest influence on the directivity of the radiated sound, we focus on width perception towards the sound image when listeners are in the near-field of the front panel of the loudspeaker array.

To determine the factor that affects the width perception of the sound image, various types of test sound sets are prepared by varying an original sound set in amplitude or phase for each frequency and for each loudspeaker. Then, the most effective factor is discovered by conducting listening tests by human subjects on the basis of the Scheffe's pair comparison method.

2 Surface vibration and radiated sound directivity of a violin

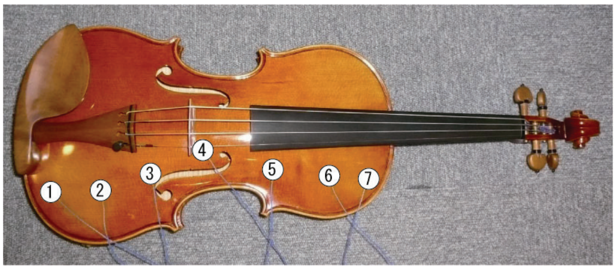
In this section, a violin is cited as an example of a sounding object. By measuring the surface vibration of the resonant body of a violin using accelerated pickups, it is illustrated that each place of the surface vibrates differently and the total of all the emitted sounds forms a directional sound.

First, the surface vibrations on the face board of the resonant body were measured at seven points using seven accelerated vibration pickups while bowing the open string of #A (tuned to 442 Hz). The accelerated vibration pickups are ONOSOKKI NP-2100, which are very small and weigh only 0.5 g. They were connected to one recording device ONOSOKKI Graduo (DS-2000 series). The measuring data were captured by PC via PCMCI card bus in the ORF format and converted into the MATLAB format afterward with a sampling rate of 51.2 kHz and 16-bit amplitude resolution. The ID number of each accelerated pickup is shown in Figure 2(a). Hereafter, the measuring data captured by a pickup whose ID is *n* is denoted as V_n.

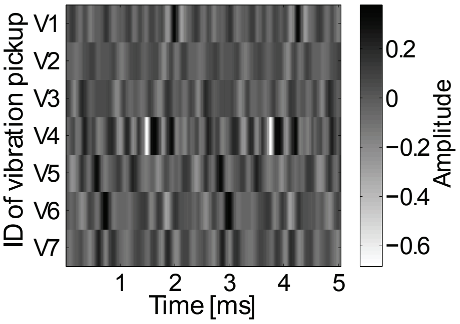
The waveforms of the measuring vibrations are shown in Figure 2(b). In this figure, the y-axis denotes the ID number of the pickup, x-axis denotes the duration in seconds and grey intensity indicates the acceleration of vibration. As shown in the figure, the timings of the highest acceleration differ between the pickups. This implies that the phases of the vibration differ on different parts of the front board.

When a board makes a sound, a bending vibration propagates on it while emitting a sound in the air. Therefore, the vibration of the edge gets delayed by a short time as compared to the vibration of the central part, for instance. This propagation delay induces a phase difference at different locations on the soundboard.

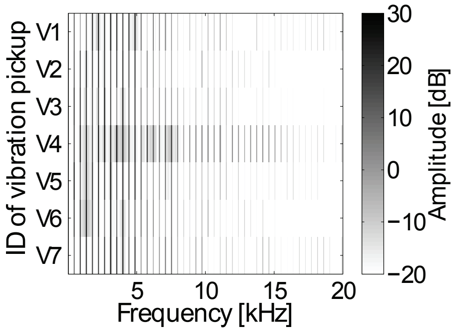
It is noteworthy that the amplitudes also differ for each location. V_4 has a wide amplitude range, while V_2 or V_3 have narrow ranges. The bending wave reflects on the edge of the soundboard and interferes with the other waves. They are sometimes reinforced or destructed depending on their wavelengths, which depend on the



(a) Setting position of accelerated pickups.



(b) Measuring waveforms with bowing opening string-A.



(c) Frequency spectra of measuring data

Figure 2: Measure bending vibration on the resonant body of a violin.

frequency. Therefore, the amplitude also differs at different locations because of the frequencies.

The frequency spectral patterns of the measuring vibrations are shown in Figure 2(c). In this figure, the y-axis denotes the channel number of the acceleration pickups, x-axis denotes the frequency in kHz and grey intensity indicates the amplitude of each frequency component: the darker region has higher amplitude than the brighter region. As shown in the figure, all the measuring vibrations are composed of waves of 442 kHz—the fundamental frequency—and its overtones, but the amplitude patterns of the overtones differ. For instance, V_2 and V_6 show high amplitudes at low-frequency overtones, while V_1 and V_7 show high amplitudes at high-frequency overtones.

It is well known that the sound radiated from a soundboard that vibrates differently at different sections has directivity that differs with frequency [1–3]. This occurs because of interference between the sounds emitted from each part of the soundboard. In particular, in the near-field from a soundboard, sounds observed at two points at short distances, such as the left and right ears of a human being, have obviously different waveforms. If humans have the capability to perceive these differences, it is inappropriate to reproduce such sound radiated from a non-uniformly vibrating board as a point sound source.

3 Human perception test by emulating surface vibrations using loudspeaker arrays

In this section, a method for testing human perception towards the radiated sound from non-uniformly vibrating surfaces is described.

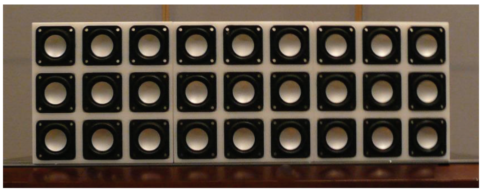
To reveal which factor of the non-uniformity yields which type of human perception, we need to independently control the phase and amplitude at each place on the soundboard surface. Therefore, we used the loudspeaker array to discretely resemble the surface vibration.

3.1 Loudspeaker array system for experiments

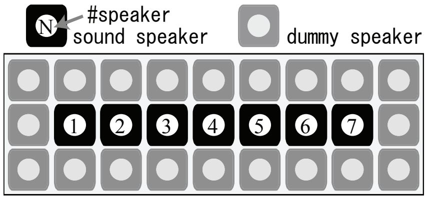
The picture and configuration illustration of the loudspeaker array system are shown in Figure 3. Here, 27 loudspeaker units (AURASOUND NSW1-205-8A) are set at 3 × 9 matrix points at 4-cm intervals on a frame with a width of 36 cm and height of 12 cm. The system has an enclosure with a depth of 4.5 cm that is made of ABS plastic.
Since the outer loudspeaker units next to the frame edge have different frequency responses from the central speakers due to the luck of reflections from the outer side, only seven central speakers in the middle low, which are numbered in Figure 3(b), were used for the following experiments. Hereafter, these seven loudspeaker units are denoted as SP_1,..., SP_7. These loudspeaker units can be controlled independently and synchronised using the Digidesign ProTools HD system. The sound source emitted from SP_n is denoted as S_n, and $S = \{S_1,..., S_7\}$ is termed the 'test sound set'.

3.2 Test sound sets

As shown in Section 2, vibrations on the same soundboard are composed of the same fundamental frequency and its multiple overtones with different phases or amplitudes. Therefore, we generated test sound sets by independently changing the phases or amplitudes for each loudspeaker unit for comparison purposes. Since it is easy to assume that sound perception depends on the type of sound, we prepared three types of test sound sets. These are synthesized single tone with multi-overtones, synthesized double tones with multi-overtones and recorded violin sound by bowing opening string-A.



(a) Photograph of loudspeaker array system.



(b) Configuration of loudspeaker array system.

Figure 3: Loudspeaker array system used in the experiments.

3.3 Experimental condition

In the experiment, a human subject listened to the test sounds by sitting on a stool and facing towards the front panel of the loudspeaker array. The centre position between both his/her ears was coincident with the centre of SP_4 and at a distance of 50 cm or 100 cm from the loudspeaker array surface. The background noise level was 21 dB and the reverberation time was 180 ms. The sound pressure level of the test sounds was adjusted to 60 dB at the hearing point.

At a time, every subject compared every pair of the test sounds in the normal and reverse orders. The test sound pairs were ordered randomly and differently for each subject. In each comparison, the first test sound was played for 1.5 s, followed by a short pause for 1 s, and then playing the second sound for 1.5 s; a middle-length pause of 2 s was taken. The system repeated this for five times and gave a long pause for 20 s. During this long pause, the subject answered the following two questions:

> Q1: The width of the sound image of the second sound was wider than the first sound.
>
> Q2: The location of the sound image of the second sound was felt to be farther than the first sound.

These questions were answered on an ascending scale of –2 to +2 (1. Strongly Agree [+2]. 2. Agree [+1]. 3. Neither Agree nor Disagree [0]. 4. Disagree [–1]. 5. Strongly Disagree [–2]). The total duration of the experiment for one subject was approximately 1.5 h including three short breaks lasting 2–3 min. In this, 11 subjects participated and 7 of them were professionals in music for more than 10 years. The following three experiments were conducted in the order of 'Experiment 1', 'Experiment 2' and 'Experiment 3'. Further, 5 subjects began these experiments first at a distance of 50 cm and then at 100 cm, while the remaining conducted this in the opposite order. The experimental results are analyzed according to Ura's modified procedure of Scheffe's pair comparison method [5].

4 Experiment 1: Single tone with multi-overtones

4.1 Test sound sets of Experiment 1

The test sound sets of Experiment 1 were generated by synthesizing sine waves of 442 Hz (with fundamental frequency F_0) and its overtones as follows:

$$S_n = \sum_m Amp_n^m \sin(2\pi F_m t + Dly_n^m), \tag{1}$$

while m indicates the order of the overtone. The frequency of the mth overtone, which is denoted as F_m, is as follows:

$$F_m = mF_0, \quad m = 0,1,\ldots,M.$$

Because the frequency responses of the loudspeaker units were almost flat from 400 Hz to 16 kHz, M was the maximum number as such that MF_0 did not exceed 16×10^3. Amp_n^m and Dly_n^m denote the amplitude and delay (phase difference) of the mth overtone, respectively.

Five test sound sets (T1–T5) were prepared. First, Amp_n^m and Dly_n^m were initialized to 0 and converted according to the functions in Table 1.

Test set T5 refers to the bending vibrations propagating on a thin plate as the resonant body of a violin. Because the propagating velocity of the bending vibrations is proportional to the square root of the frequency,

$$D(n,m) = \frac{a \cdot l_n}{\sqrt{F_m}}, \tag{2}$$

while $a = 9.1$ and l_n denotes the distance between the corresponding loudspeaker unit and the centre of the loudspeaker array.

ID	Meaning Test set generation function
T1	Only SP_4 sounds #A4.
	$Amp_4 = 1$
T2	SP_1, SP_4 and SP_7 sound #A4.
	$\{Amp_n^m = \frac{1}{\sqrt{3}} \mid n = 1,4,7 \wedge \forall m\}$
T3	All loudspeaker units sound #A4.
	$\{Amp_n^m = \frac{1}{\sqrt{7}} \mid \forall n \wedge \forall m\}$
T4	SP_2, SP_4 and SP_6 sound even-ordered overtones, SP_1, SP_3, Sp_5 and SP_7 sound odd-ordered overtones
	$\{Amp_n^m = \frac{1}{\sqrt{3}} \mid n = 2,4,6 \wedge m = 2k\}$, $\{Amp_n^m = \frac{1}{\sqrt{4}} \mid n = 1,3,5,7 \wedge m = 2k+1\}, k \in N$
T5	The loudspeaker array emulates bending vibrations propagating on a plate.
	$\{Amp_n^m = \frac{1}{\sqrt{7}}, Dly_n^m = D(n,m) \mid \forall n \wedge \forall m\}$

Table 1: Test sound sets of Experiment 1

4.2 Results of Experiment 1 concerning width perception

ID	50 cm	100 cm
T1	0.536	–0.254
T2	0.181	0.064
T3	0.072	–0.073
T4	–0.227	–0.073
T5	0.509	0.336

(a) Average ratings

50 cm	100 cm
T1 < T2**	
T1 < T3**	
T1 < T5**	T1 < T5**
T2 > T4*	
T3 < T5**	T3 < T5*
T4 < T5**	T4 < T5*

(b) Reliable order with 99% (**) or 95% (*) confidence interval

Table 2: Experiment 1 results concerning width perceptions of sound images

In the result of the question about the sound image width, the F value (hereafter v1) of the main effect was 18.8 with four degrees of freedom (hereafter v2), and the difference in the main effect among individuals yielded v1 = 3.47 and v2 = 40 for a distance of 50 cm from the loudspeaker array. Both these effects were significant ($p < 0.01$). Both the combination effect and sample order effect were not significant. The average ratings of the test sound sets are listed in Table 2(a). The orders of the pair listed in Table 2(b) were reliable (95% or 99% confidence intervals).

As shown in Table 2(b), the subjects answered that T2 and T3 was significantly wider than T1 at a distance of 50 cm, but these perceptions were not significant at a distance of 100 cm. This means that humans perceive the width of a sound image significantly at 50 cm; however, it becomes imperceptible at a distance as far as 100 cm from the sound sources. The order relation between T2 and T3 was not significant even at a distance of 50 cm. This means that the perception towards the sound image width is controlled by the outer edge of the sound source and the sound density within the sound source is not very important. Because T5 was felt to be significantly wider than T3 and T2, it has been suggested that changing the amplitude or phase at different sections of the sound surface yields a certain influence towards the perception of sound source width.

Although all the loudspeaker units were sounded in T4, T4 was felt to be significantly narrower than the others. It is know that when the two formants have the same fundamental frequency, listeners generally report hearing only a single sound, even though the formants were obtained from different ears [6]. In this experiment, multiple sound sources are considered to be integrated into a single sound in the subjects' perceptions, and they felt the sound image to be as narrow as T2. However, there is no explanation for why T4 was significantly narrower than T2. At any rate, it is said that changing the phase for each frequency and for each loudspeaker unit can yield stronger effects with regard to the width perception of the sound image comparing to changing the amplitude.

4.3 Results of Experiment 1 concerning depth perception

ID	50 cm	100 cm
T1	0.209	–0.082
T2	–0.045	0.100
T3	–0.082	0.100
T4	0.082	0.127
T5	–0.164	–0.245

(a) Average ratings

50 cm	100 cm
T1 > T5*	
	T2 > T5*
	T3 > T5*
	T4 > T5*

(b) Reliable order with 99% (**) or 95% (*) confidence interval

Table 3: Experiment 1 results concerning depth perceptions of sound images

In the result regarding the question about sound image depth, at a distance of 50 cm, the main effect (v1 = 3.36 and v2 = 4) was significant ($p < 0.005$). The difference in the main effect among individuals (v1 = 6.21 and v2 = 40) was also significant ($p < 0.001$). At a distance of 100 cm, the main effect (v1 = 4.05 and v2 = 4), the difference in the main effect among individuals (v1 = 4.86 and v2 = 40) and the effect of order (v1 = 11.99 and v2 = 1) were significant. The average ratings are listed in Table 3(a). The orders of the pairs listed in Table 3(b) are reliable (95% or 99% confidence intervals).

The subjects felt that T5 was the nearest as compared to the others. In the foregoing experiment, T5 was felt as the widest with regard to the sound image width. Since the sound pressures were the same for all the sound test sets at the hearing point, if the subjects felt a constant sound image size, then wider the sound image was felt, the nearer it was felt. In fact, there was a strong correlation between the answers regarding width and depth perceptions.

5 Experiment 2: Double tone with multi-overtones

ID	Meaning Generation function
T6	Only SP_4 sounds #A4 + #C4. $Amp_4 = 1, Amp_4' = 1$
T7	SP_2, SP_4 and SP_6 sound #A4 and SP_1, SP_3, SP_5 and SP_7 sound #C4. $\{Amp_n = \frac{1}{\sqrt{3}} \mid n = 2,4,6\} \wedge \{Amp_n' = \frac{1}{\sqrt{4}} \mid n = 1,3,5,7\}$.
T8	All loudspeakers sound #4 + #C4. $\{Amp_n^m = \frac{1}{\sqrt{7}}, Amp_n^m{}' = \frac{1}{\sqrt{7}} \mid \forall n\}$

Table 4: Test sound sets of Experiment 2

5.1 Test sound sets of Experiment 2

The test sound sets of Experiment 2 were generated by synthesizing multi-overtones of #A4 and #C4. Hereafter, the amplitudes of #A4 and #C4 are denoted as Amp_n and Amp_n', respectively. Each multi-overtone was synthesized according to function (1). Four test sound sets were generated by initializing both Amp_n and Amp_n' to 0 for all n and according to the functions described in Table 4.

5.2 Results of Experiment 2 concerning width perception

The result of the question concerning sound image width did not show any significant effect. The average ratings are listed in Table 5.

ID	50 cm	100 cm
T6	–0.227	0.106
T7	0.045	–0.045
T8	0.182	–0.061

Table 5: Average ratings for Experiment 2 results concerning the width perceptions of sound images

There was no significant difference between T6 and T8. Because the relation between T6 and T8 seems to be very similar to the relation between T1 and T3 in Experiment 1 (only SP_4 sounded on T1 and T6 and all the loudspeaker units sounded with the same waveform on T3 and T8), a comparison of their results suggests that a multi-tone tends to be felt as being wider than a single tone. This is attributed to the fact that usually double tones are emitted from two discrete objects and therefore the sounds of a double tone are recognized as being emitted from a wider space.

6 Experiment 3: Captured sound of a violin

6.1 Test sound sets of Experiment 3

The test sound sets of Experiment 3 were generated by synthesizing real vibration data that was measured at seven points on a resonant body of a violin, as illustrated in Section 2. We captured the violin sound by bowing open string #A in one bow stroke for more than 2 s. Then, we clipped out a wave with a duration of 1.5 s and tapered the beginning and ending of the clipped sound. This procedure was performed for each channel of the accelerated pickup. The clipped sound wave that was captured at the ith turn using the accelerated pickup of channel number k is denoted as V_k^i. The generation function for the sound test sets is described in Table 6.

ID	Meaning Generation function
T9	All loudspeaker units play the same wave measured at the centre position. $\{S_n = V_4^1 \mid \forall n\}$
T10	Each loudspeaker unit plays the wave measured at the corresponding position at the same turn. $\{S_n = V_n^1 \mid \forall n\}$.
T11	Each loudspeaker unit plays the waves measured at the centre position and captured at different turns. $\{S_n = V_4^n \mid \forall n\}$

Table 6: Test sound sets of Experiment 3

6.2 Results of Experiment 3 concerning width perception

ID	50 cm	100 cm
T9	–0.773	–0.606
T10	0.348	0.394
T11	0.424	0.212

(a) Average ratings

50 cm	100 cm
T9 < T10**	T9 < T10**
T9 < T11**	T9 < T11**

(b) Reliable order with 99% (**) or 95% (*) confidence interval

Table 7: Experiment 3 results concerning width perceptions of sound images

In the result of the question concerning the sound image width, at a distance of 50 cm, the main effect ($v1 = 29.54$ and $v2 = 2$) and its difference between individuals ($v1 = 3.77$ and $v2 = 20$) were significant ($p < 0.01$). At a distance of 100 cm, the main effect ($v1 = 26.26$ and $v2 = 2$) and its difference between individuals ($v1 = 6.61$ and $v2 = 20$) were also significant ($p < 0.01$). The average ratings are shown in Table 7(a). The orders of the pair listed in Table 7(b) were reliable (95% or 99% confidence intervals).

T10 and T11 were felt wider than T9. This means that just playing the completely same wave from a wide area is insufficient to emit a wide sound image with respect to human perception. On both T10 and T11, all the loudspeaker units sounded at the same time as well as T9, but they played different waves for each loudspeaker unit. This result is consistent with the result of Experiment 1. T11 was synthesized from the waves captured at the same position but at different timings. Despite the fact that the phase modulation of T11 was not based on physics unlike T10, no significant difference was observed between T10 and T11. These results reveal that any phase difference between the sounding surfaces makes a listener feel that the sound image is wide. However, on T11, the centre of the sound image moved slightly and busily because there was no correlation between the emitted sounds from the loudspeaker units and also the phase delay was not always bilaterally symmetric.

7 Discussion

The result of Experiment 1 (single tone with multi-overtones) showed that changing the vibration phase at different sections of the sounding surface induced a feeling of a wider width perception of the sound image. This can be attributed to the following: when the size of a sounding object is large, the time for the wave propagating to the edge of the object becomes longer, and the phase differences in the emitted sound from the surface also becomes large. Conversely, the phase differences in a sound along the horizontal direction could possibly yield such perception that the sound has been emitted from a wider area and the sounding object might be large.

The result of Experiment 2 (double tone with multi-overtones) showed that the subjects tended to feel that a sound image composed of two tones—which had different fundamental frequencies—was wider than the sound image composed of a single tone. This can be attributed to the fact that the subjects felt that two objects were simultaneously sounded. It is common knowledge that one object tends to emit a sound of a single tone, making people think that two disconnected objects are sounding when a sound composed of two tones is produced; moreover, two objects occupy more space than one object.

The result of Experiment 3 (captured sound with bowing single open string of a violin in one stroke) almost supported the result of Experiment 1. Phase differences along the horizontal direction made the subjects feel that the sound image was wide. However, even though the phase difference was not based on physics, the subjects also felt the sound image was wide.

By combining these results, we propose the following hypothesis:

> *When a sound is composed of a single tone and multi-overtones, changing the phase of the emitted sound for every section of a sounding surface and for each frequency makes a listener feel that the sound image is wider.*

This hypothesis will be validated in the following work.

8 Conclusion

Many musical instruments, including violins and guitars, vibrate their resonant bodies differently over their resonant surfaces while making a sound. This paper investigated the influences of such vibration variations of a soundboard surface towards the width perception of the sound image when listeners were in the near-field—at a distance of 50 cm or 100 cm from the soundboard.

A loudspeaker array composed of seven loudspeaker units was used to emulate the surface vibration as each loudspeaker made a corresponding sound independently and cooperatively. Three types of sounds—synthesized single tone with multi-overtones, double tone with multi-overtones and captured sound of a violin—were used as the sound sources. To determine what factors influence the width perception of the sound image, various test sound sets were prepared by varying an original sound set in amplitude or delay for each frequency and for each loudspeaker. Eleven subjects were asked to identify which sound image in a pair of test sounds was wider or farther than the other according to Ura's modified procedure of the Scheffe's pair comparison method.

The results showed that a test sound set with phase variation along the horizontal plane, which mimics sounds emitted by bending vibrations propagating on a soundboard, was felt wider than the synchronizing phase when the sound was composed of a single tone. Therefore, we formulated the following hypothesis: *when a sound is composed of a single tone and multi-overtones, changing the phase of the emitted sound for every section of a sounding surface and for each frequency makes a listener feel that the sound image is wider.* This hypothesis will be validated in the following work.

References

[1] Y. Yamakata, T. Kimura, M. Katsumoto, "Directional sound radiation system using a large planar diaphragm incorporating multiple vibrators", *Proc. ICASSP2008, AE-P6.D4*, 337-340 (2008)

[2] N.H. Fletcher, T.D. Rossing, "The physics of musical instruments, Springer", 2nd edition, Springer-Verlag, (2005)

[3] G. Weinreich, "Radiativity revisited: theory and experiment ten years later", *Proc. SMAC 93, Ed. A. Friberg, et al. Royal Swedish Acad. Music, Stockholm* (1994)

[4] R.V. Dinther, R.D. Patterson, "Perception of acoustic scale and size in musical instrument sounds", *J. Acoust. Soc. Am.* 120, No. 4, 2158-2176 (2005)

[5] E. Masuyama, S. Kobayasi, "Sensory evaluation (in Japanese)", 173-180, Kakiuchishuppan (1989)

[6] C.J. Darwin, R.P. Carlyon, "Auditory grouping" in B. C. J. Moore (Ed.), "Hearing", Academic Press, 387-424 (1995)