Discovering a physical parameter associated with a near-field sound control: comparing HRTFs of nine loudspeakers in a non-anechoic room

Sungyoung KIM[1]; Philip GOSSELIN[2]; Hiraku OKUMURA[3]

[1] Rochester Institute of Technology, Rochester, NY, USA,

[2] McGill University, Montreal, Quebec, Canada,

[3] Yamaha Corporation, Hamamatsu, Japan

ABSTRACT

Previously, the authors proposed a novel method of near-field sound control using an electrostatic loudspeaker radiates planar wave from above the listening position. The method was based on a pseudo inside-head-localization (IHL) phenomenon similar to headphone listening experience, which generates a near-field sound image. To search for a physical parameter associated with this near-field image rendered through the electrostatic loudspeaker, the authors measured the head-related transfer function (HRTF) with eight conventional loudspeakers and compared them with the HRTF of the electrostatic loudspeaker. The HRTFs were measured at three distances (0.75, 1, and 2 m) in a non-anechoic listening room. The results were similar to the previous HRTF measurements at an anechoic room and numerical simulations, showing that the variances of group delays were associated with the distances. Moreover, the variance of electrostatic loudspeaker measured at 2 m was similar to the variance in conventional loudspeaker at 0.75 m, suggesting a reason for near-field sound images from the electrostatic loudspeaker in a reverberant room.

Keywords: Near-field sound control, Electrostatic loudspeaker, Variance in group delays
I-INCE Classification of Subjects Numbers: 25.5 & 72.9

1. INTRODUCTION

In various multichannel sound reproduction techniques, rendering a convincing near-field sound image has been a challenging task. To the best knowledge of the authors, only Higher Order Ambisonics (HoA) and Wave Field Synthesis (WFS) could render the acoustical characteristics of a near-field sound image that diverges from a target position. For instance, a time-reversal acoustic focusing could render a focused sound source near to the listening position (1).

However, focused sources rendered with previous methods appeared to “exhibit a number of artifacts that are not desirable for high-quality reproduction of sound” (2). Furthermore, effective near-field sound control via HoA or WFS requires a large number of loudspeakers.

Previously, the authors proposed a new method that controls and locates auditory images near to a listener (3). The method utilized an overhead electrostatic loudspeaker (located at a ceiling, for example). The loudspeaker is thin (1 mm), light (400 g/m2), and flexible (possible to fold, roll, and have printing placed on its cover), and therefore it can be conveniently located on a wall or ceiling.

Compared with a conventional loudspeaker, this loudspeaker rendered the perceived sound image nearer than the physical distance of a sound source. In particular, when placed overhead, this loudspeaker generated a sound image as if it being inside the listener’ head. This phenomenon, known as inside-the-head localization (IHL), often happens when we listen to sound or music with a pair of headphones. We also applied additional signal processing that eliminated spectral cues of elevation of the loudspeaker, which provided listeners with a near-field sound image. We used this near auditory image to control perceived distance between a screen and a listening position and to render sound image projecting out of a TV as a 3-D visual object does.

To better understand why the electrostatic loudspeaker generated an auditory image near a listener,

[1] sxkiee@rit.edu
[2] philip.gosselin@mail.mcgill.ca
[3] hiraku.okumura@music.yamaha.com

the authors investigated physical parameters that highlight the difference between a conventional and the electrostatic loudspeaker. Specifically, we wanted to find related physical parameters that changed according to the distances for a conventional loudspeaker yet remained constant for the electrostatic loudspeaker. Physical parameters include the interaural level difference (ILD), the interaural time difference (ITD), the interaural phase difference (IPD), and the variance in group delays of head-related transfer functions (HRTFs). Since an electrostatic loudspeaker generates planar waves, our investigation is equivalent to comparing characteristics of spherical waves radiated from a conventional loudspeaker with planar waves from an electrostatic loudspeaker.

To externally verify previous measurement results of an anechoic chamber and numerical simulation, the authors collected impulse responses (IRs) of eight conventional loudspeakers in two normal listening rooms and compared them the electrostatic loudspeaker. In following sections, the summary of previous measurement and simulation results will be introduced, followed by multi-point measurement results of the electrostatic loudspeakers, and the comparison results with conventional loudspeakers will be presented.

2. ANECHOIC MEASUREMENT AND COMPUTATIONAL SIMULATION

Previously we calculated physical characteristics from the HRTFs of two loudspeakers in two circumstances: (1) in an anechoic chamber, and (2) via a numerical simulation.

We first measured HRTFs of two loudspeakers using a Head And Torso Simulator (HATS: 4128D, Brüel & Kjær) from at three distances–0.5, 1, and 2 m–in an anechoic chamber. The electrostatic loudspeaker used in the measurement was 60 cm wide and 60 cm tall. The conventional loudspeaker was FOSTEX FF85K with a custom-made cabinet.

Subsequently, we calculated numerical simulations of HRTFs using the boundary element method (BEM) (4), which could account for reflecting objects in the acoustic field simulation. We simulated the impulse responses at both the ears of the HATS. Three-dimensional geometry of the HATS was scanned using a non-contact optical 3D digitizer (ATOS) to yield two models; one has only a head and another has a head and shoulders as shown in Fig. 1. The ear canals were blocked at their entrances. As results, two computational head models with and without shoulders were produced. In addition, we simulated two sound-source positions, in front of and above, for the head model with shoulders.

A point and planer source were employed as sound sources. The point source is an ideal volumeless point and the planar source is an infinite plane. The two wave types used in the measurement were chosen to approximate the radiation pattern of the conventional and electrostatic loudspeakers, respectively. Distances of spherical sources were identical to measured distance, and the distance of the planar wave was set at 1.8 m. The receivers were located at points adjacent to the blocked ear-canal entrances with a few millimeters spacing to the surfaces. All the boundary surfaces were assumed to be acoustically rigid. First, frequency responses were calculated at frequencies from 43 Hz to10 kHz with 43-Hz intervals, which resulted in impulse responses of 1,024 points at 44.1 kHz sampling frequency.

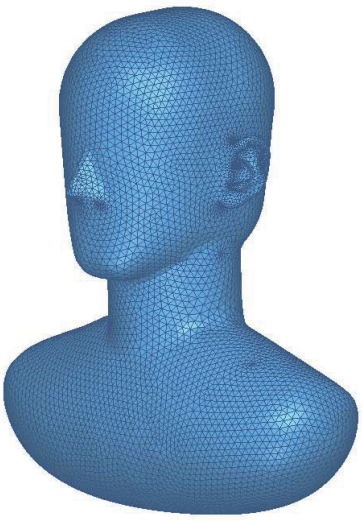

Figure 1 – Figures of a head-and-torso model.

From both the measured and calculated HRTFs, we generated various physical characteristics including the interaural level difference (ILD), the interaural time difference (ITD), the interaural phase difference (IPD), and the variance in group delays. The variance in group delays has been included because Toyama (5, pp. 216 – 219) reported that the degree of variation in differentials of a phase response can be linked to the perceived auditory distance (smaller variation for closer distance perception).

Figure 2 show the variance in group delays of measured and calculated HRTFs respectively. We calculated the variances in the selected frequency bands ranging between 400 Hz and 8000 Hz because the loudspeaker's phase responses modulated unstably at below and above this range. The result shows that the planar wave produced a smaller difference between group delays across frequency at the two ear positions than did the spherical wave radiated with 0.5 m distance. The implication is that a planar wave produced less variance in group delay (in the selected frequency range), which may be related to a near-field sound image.

Other binaural measurement results, however, did not significantly change their values according to the given distances.

The analysis further revealed that the increase in the variance in group delays was related to a reflection from a torso. As the right panel of Fig. 2 shows, when the simulation accounted for only the head model, the variance in group delays remained constant regardless of the loudspeaker distance. Yet with the torso model in it, the variances associated with spherical wave simulation increased with source distances. Furthermore, the increase became larger when the sound source was above, possibly because of direct reflections from the shoulder. For planar wave simulation, the variances remained small as the measured results.

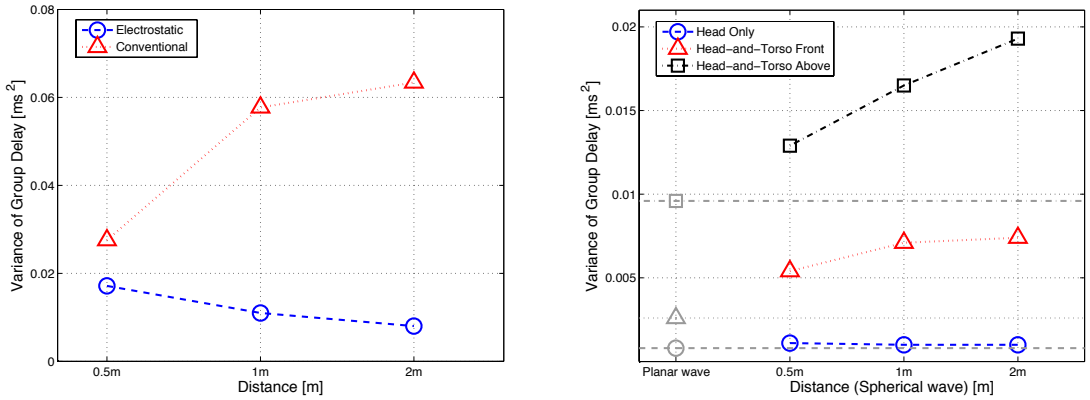

Figure 2 – [Left panel] Variance in group delays of measured HRTFs. The variance was directly related to the sound source distance for the conventional loudspeaker, while it maintained a relatively small value for the electrostatic loudspeaker. [Right panel] Variances of group delays of three numerical simulations: (1) a head-only model, (2) a head-and-torso model with the loudspeaker located in front of the model, and (3) a head-and-torso model with the loudspeaker located above. The variances of spherical waves are displayed according to their distances. Planar waves (one distance–1.8m) are displayed as single lines for the comparison to spherical waves.

3. ANALYSIS OF HRTFS OF NINE LOUDSPEAKERS MEASURED IN TWO NON-ANECHOIC ROOMS

The previous measurement and simulation results showed that the variance in group delays was a parameter associated with near field sound image generated by an overhead electrostatic loudspeaker. To externally verify this finding, we conducted additional measurement in two non-anechoic rooms. The underlying hypothesis was that a specific room acoustics could alter the variance in group delays and weaken difference between a conventional and electrostatic loudspeaker. Moreover, we wanted to check whether measurement result was influenced by a specific loudspeaker or not.

3.1 Peripherals

We mainly measured the HRTFs of loudspeakers at a listening room called MARLAB (Multichannel Audio Research LABoratory) of McGill University. The ambient noise in the room was 27 dBA measured at the center of the room and reverberation time (RT_{60}) is about 247 ms (average of an octave band at 500 Hz and 1 kHz). In addition, we used a critical listening room (A816) located at the Center for Interdisciplinary Research

Table 1 – Eight loudspeakers used for the HRTF measurement compared to an electrostatic loudspeaker

JBL LSR6326P	PMC MB2S-A	Yamaha MSP5	Yamaha MSP7
Dynaudio BM5A	Dynaudio BM15A	Genelec 8020B	Genelec 8050A

in Music Media and Technology (CIRMMT) of McGill University. The RT_{60} of the CIRMMT room is about 168 ms.

We measured the electrostatic loudspeaker with eight conventional loudspeakers located at three distance–0.75, 1, and 2 m. In this measurement, 0.75 m was the closest distance where we could locate all loudspeakers. We used two microphones–DPA 4011 for reference and B&K Head-And-Torso-Simulator (HATS) and the AFMG EASERA software for measurement. Table 1 lists the used eight loudspeakers compared with an electrostatic loudspeaker.

It should be noted that, due to the limitation of accessibility of the loudspeakers, we could measure all loudspeakers only at MARLAB. And at the critical listening room of CIRMMT, we measured only two loudspeakers (Dynaudio BM15A and the electrostatic loudspeaker) for comparison of two rooms. Also, the electrostatic loudspeaker was bigger (1 m wide and 1.4 m tall) than previously used one for anechoic measurement. We used this bigger loudspeaker since this recent model has improved frequency-magnitude response and produces higher sound pressure level. Another technical limitation was that we could not locate the loudspeakers overhead position; we instead located all loudspeakers in front of two microphones.

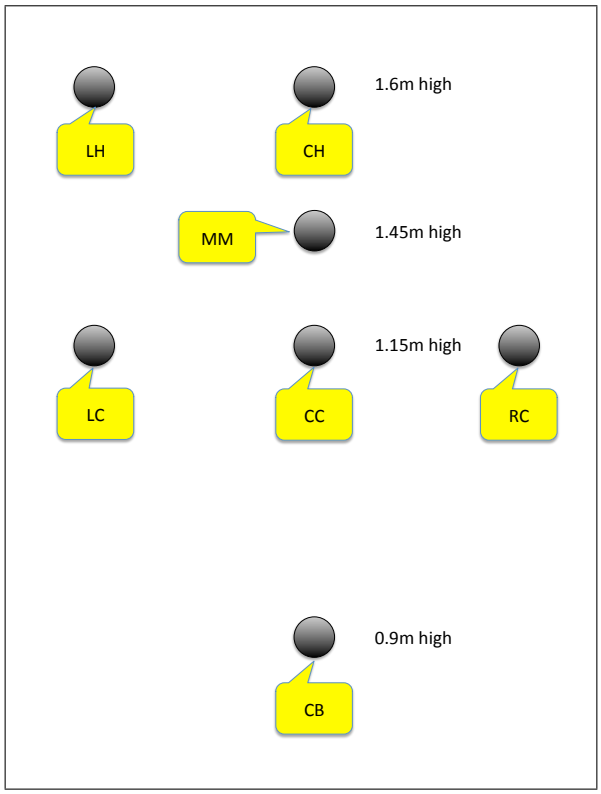

Figure 3 – Seven measurement points of the electrostatic loudspeaker and corresponding annotations.

3.2 Result I: Variance in group delays of an electrostatic loudspeaker

First, we wanted to know the variance in group delays of the electrostatic loudspeaker over its large surface area and thus measured the loudspeaker-to-microphone transfer functions at multiple positions. Figure 3 illustrates the measurement points and associated annotations. For example, “CC” indicates the horizontal center and vertical center, which is equivalent to the location of the center of the B&K Head-And-Torso-Simulator (HATS).

Figure 4 shows the variance in group delays measured at the MARLAB and the CIRMMT A816 in three different distance and seven microphone positions. The results first show that the variances measured at seven

Table 2 – The average variances in group delays of severn measured points.

	0.75 m	1 m	2 m
CIRMMT	0.0018	0.0017	0.0036
MARLAB	0.0030	0.0028	0.0072

points remained smaller than the values of conventional loudspeakers (please refer the measurement results of conventional loudspeaker in the following section). At the same time, however, the variances (in group delays) were not constant over all measurement points and also different depending on measured rooms. For instance, in the MARLAB, the position "RC" showed the biggest change but it was "LH" in the CIRMMT room. This indicates that the electrostatic loudspeaker did not generate a perfect planar wave over its surface area.

Furthermore, the results imply that the room acoustics influenced on the variances in group delays. Table 2 shows the mean variance in group delays (of seven measured points) increased according to its measuring distance while the previous measurement at an anechoic chamber (left panel of Fig. 2) showed that the variance of group delay remained constant regardless of distance.

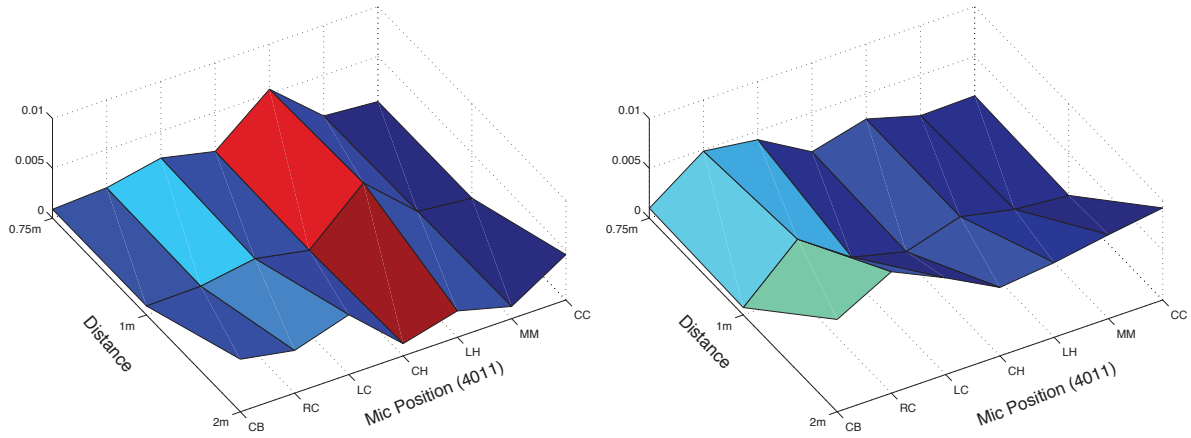

Figure 4 – [Left panel] Variance in group delays of the electrostatic loudspeaker measured at seven positions of the CIRMMT critical listening room. The variances were measured at three distances (0.75 m, 1 m and 2 m) and seven positions (as illustrated in Fig. 3) using a cardioid microphone pointing to the loudspeaker. [Right panel] Variances of group delays of the electrostatic loudspeaker measured at the MARLAB.

3.3 Result II: The variance in group delays of nine loudspeakers

Table 3 shows the measured variances in group delays of the electrostatic loudspeaker (denoted "EL") and eight conventional loudspeakers (denoted "S1" through "S8"). The variances were calculated from the left ear signal of HATS.

Two leftmost columns–S5′ and EL′–are the results measured at the CIRMMT A816 and following nine columns show measurement results of the MARLAB. The S5 and EL were measured in both rooms for comparison. The results show that the electrostatic loudspeaker generated relatively smaller variances in group delays at the two rooms; average variances of nine conventional loudspeakers (S1–S8 plus S5′) were 0.0045 (0.75 m), 0.0072 (1 m), and 0.046 (2 m), which are bigger than average variances of EL and EL′ (0.0008, 0.0008, and 0.0059) . The electrostatic loudspeaker generated small variance in group delays and rendered near-field sound images even in a non-anechoic room. Compared with the previous anechoic measurement and numerical simulation, this result indicates that room acoustics increased the variances of the electrostatic loudspeaker (at 2 m) in both rooms. However, the variances in group delays of the electrostatic loudspeaker located at 2 m were greater than all of conventional loudspeakers (except S1) located at 0.75 m.

Table 3 – The variances in group delays of nine loudspeakers measured in non-anechoic rooms–MARLAB and CIRMMT A816. S1 through S8 indicate conventional loudspeakers (as listed in Table 1) and EL indicates the electrostatic loudspeaker.

	S5′	*EL′**	*EL**	S1	S2	S3	S4	S5	S6	S7	S8
0.75 m	0.0039	*0.0006*	*0.0010*	0.0082	0.0033	0.0037	0.0074	0.0018	0.0046	0.0026	0.0053
1 m	0.0096	*0.0007*	*0.0009*	0.0134	0.0037	0.0035	0.0096	0.0084	0.0096	0.0018	0.0048
2 m	0.0332	*0.0054*	*0.0065*	0.0698	0.0464	0.0469	0.0411	0.0392	0.0423	0.0412	0.0540
	CIRMMT A816		MARLAB								

*Electrostatic Loudspeaker

4. CONCLUSION

The authors measured the HRTFs of nine loudspeakers in two non-anechoic rooms, and compared their variances in group delays, which is known to covary with the receiver-loudspeaker distance. Previously we showed that a planar wave maintain smaller variance regardless of the distance and produces a near sound image. Current study results show that the variances of an electrostatic loudspeaker remained smaller than other eight loudspeakers in both rooms, although the room acoustics increased the variances of the electrostatic loudspeaker measured at 2 m distance. The results externally support that the electrostatic loudspeaker generated a near-field sound image due to the smaller value of variance in group delays. Due to this physical characteristic, the electrostatic loudspeaker could render convincing near-field sound images. The near-field sound images will integrate auditory images with new high-definition three-dimensional visual images, and deliver increased reality and immersive presence for future media reproduction.

ACKNOWLEDGEMENTS

This study has been supported by Spatial Acoustic Solution Group in the Research & Development Division of Yamaha Corporation.

REFERENCES

1. Yon S, Tanter M, Fink M. Sound focusing in rooms: The time-reversal approach. J Acoust Soc Amer. 2003;113(3):1533 – 1543.

2. Spors S, Wierstorf H, Geier M, Ahrens J. Physical and Perceptual Properties of Focused Virtual Sources in Wave Field Synthesis. In: Proc. Audio Engineering Society 127th Int. Conv. New York, USA: AES; 2009. Preprint 7914.

3. Kim S, Okumura H, Sakanashi H, Sone T. Auditory depth control: a new approach utilizing a plane wave loudspeaker from above a listener. In: Proc. Audio Engineering Society 131st Int. Conv. New York, USA: AES; 2011. Preprint 8484.

4. Otani M, Hirahara T, Ise S. Numerical study on source distance dependency of head-related transfer functions. J Acoust Soc Amer. 2009;120(5):3253–3261.

5. Toyama M. Signal Analysis and Acoustics. Springer Japan KK; 2007. (Japanese).