PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Sound field evaluation by using closely located four-point microphone method and mixed reality technology

Wataru TERAOKA[1]; Yuta KATAOKA[2]; Yasuhiro OIKAWA[3]; Yusuke IKEDA[4]

[1-3] Waseda University, Japan

[4] Tokyo Denki University, Japan

ABSTRACT

The closely located four-point microphone method was proposed by Yamasaki in 1989. The method can grasp the spatial and temporal structures of sound reflections by estimating an image source distribution from four measured impulse responses. In recent years, mixed reality (MR) technology has rapidly developed and is now more familiar. Many sensors, display devices, and ICT technologies have been implemented in MR equipment, which enable interaction between real and virtual worlds. We proposed a MR display system for image source distributions and directivity patterns of sound reflections, which are obtained by the closely located four-point microphone method. In this paper, we propose a MR display system for some image source distributions measured at multiple measurement points. By moving in the room, the user can view the image source distribution in the room adjusting to the user's view point simultaneously. MR display enable us to observe the spatial and temporal structures of sound reflections in the room while maintaining a relationship between the positional information of the real room and the data, even when user moves around the room.

Keywords: Closely located four-point microphone method, Mixed reality, Architectural acoustics

1. INTRODUCTION

There are various ways to measure sound positions or grasp the spatial and temporal structures of sound reflections. Sound field visualization is an effective technique to understand spatial sound information. For example, in previous research, acoustical holography [1, 2], beam-forming [3, 4], and optical methods [5-14] have been proposed and studied.

One of the visualization methods, the closely located four-point microphone method, was proposed by Yamasaki in 1989 [15]. The method constructs an image source distribution from four impulse responses measured by using four microphones that are not coplanar and by using a short-time correlation technique. The image source distribution represents the spatial and temporal structures of sound reflections.

On the other hand, the HoloLens, a head mounted display developed by Microsoft, is realized by mixed reality (MR) technology in which real and virtual spaces interact with each other as if the real and virtual worlds were merged. Its applications are receiving increases worldwide attention. By taking advantage of MR technology, visualization methods of three-dimensional sound intensity maps with an optical see-through head mounted display (OSTHMD) were recently proposed [16-18].

In the previous papers, we proposed an MR display system for the image source distribution obtained by the closely located four-point microphone method [19]. We measured impulse responses only one point and display the image source distribution which can observe at any points. However, image sources depends on the measurement point. Thus it needs to change the displayed data when user translates.

In this paper, we propose a MR display system which can display image source distribution measured at multiple measurement points. At first we measured impulse responses and calculate image source

[1] tera-wata1129@fuji.waseda.jp
[2] kataoka8894@akane.waseda.jp
[3] yoikawa@waseda.jp
[4] yusuke.ikeda@mail.dendai.ac.jp

distributions at many points using closely located four-point microphone method. The system displays the image source distribution at user's position while the user is moving in the room. Thus, the user can observe the spatial and temporal structures of sound reflections in the whole room while maintaining a relationship between the positional information of the real room and the data, and also can evaluate sound field at user's position.

2. METHOD

2.1 Closely located four-point microphone method

To grasp the spatial structures of sound fields, the closely located four-point microphone method constructs an image source distribution from four impulse responses measured by using four closely-located microphones that are not coplanar and by using a short-time correlation technique [15]. The method is briefly described in the following discussion.

The distances $r_{i,n}$ between the i-th $(i = 1, 2, 3, 4)$ microphone and the n-th image source are represented by

$$r_{i,n} = Ct_{i,n}\ , \tag{1}$$

where C is the speed of sound and $t_{i,n}$ is the arrival time from the n-th image source at the i-th microphone. When four microphones are located on the origin and three points at the same distances from the origin on the rectangular coordinate axis, the coordinates of the n-th image source are estimated by

$$X_n = \frac{d^2 + r_{1,n}^2 - r_{2,n}^2}{2d},$$
$$Y_n = \frac{d^2 + r_{1,n}^2 - r_{3,n}^2}{2d}, \tag{2}$$
$$Z_n = \frac{d^2 + r_{1,n}^2 - r_{4,n}^2}{2d},$$

where d is the distance between microphones located on the origin and a rectangular coordinate axis. Therefore, if microphone distance d, arrival time $t_{i,n}$, and the speed of sound C are known, positional information of the image sources can be estimated by using Eq. (2). The arrival time $t_{i,n}$ is estimated by a short-time correlation technique.

As the result of closely located four-point microphone method, image source distribution can be displayed. Figure.1 shows a scaled sound image source distribution and its color reference. In displaying the image source distribution, the origin of the coordinates shows the sound-receiving point, the center of the sphere shows the coordinates of the image source, and the diameter of the spheres shows the energy of the image source. The arrival time from each image source to the receiving point is expressed by colors.

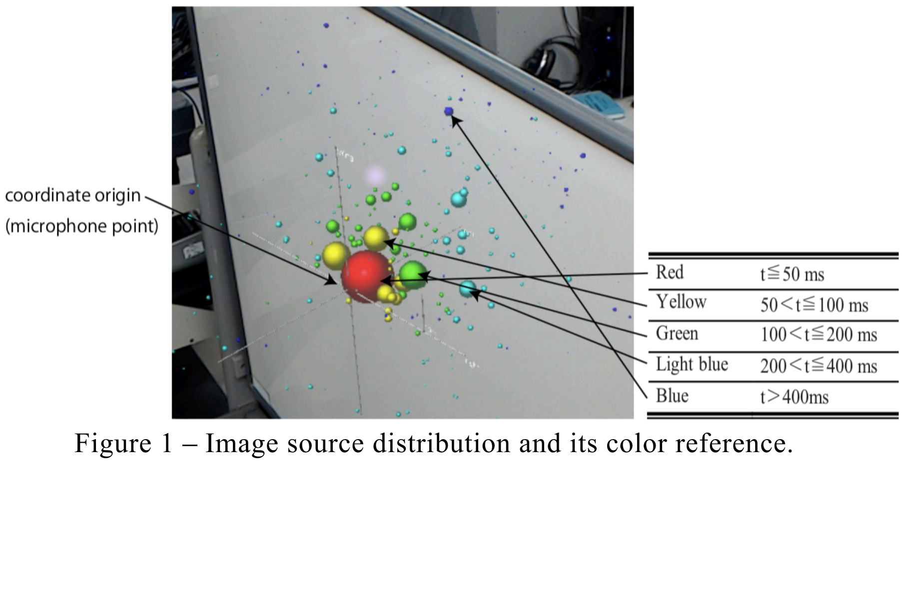

Figure 1 – Image source distribution and its color reference.

2.2 Mixed reality technology and augmented reality marker

In MR technology, the real and virtual worlds in visual information interact with each other as if these two worlds were merged. In this paper, we introduced MR technology by using Microsoft HoloLens as shown in Fig.2. HoloLens is a self-contained holographic computer with an OSTHMD and simultaneous localization and mapping (SLAM) techniques [20]. A HoloLens is a stereo transparent display that can overlay 3DCG on the real view. SLAM technology can obtain the viewing position, the direction of the user, and the shape of the room with a depth camera and image processing [21]. By using the spatial mesh of a real object's shape, 3DCG objects can be occluded by the real objects as if the virtual object were in real space. The spatial mesh continues to be update and follows the observer's translation and rotation. By continuing to change the 3DCG view on the basis of the spatial mesh and the user's translation and rotation, 3DCG is appropriately overlaid on the view of the freely moving user.

We use augmented reality (AR) marker as spatial marker. SLAM sensors attached to HoloLens continuously acquire the space shape of the room. When HoloLens recognizes the marker with SLAM sensors, marker's position represents the reference point of measured room.

In this paper, the proposed system was developed with Unity2018.3, which contains the Unity3D engine of Unity Technologies, and with Microsoft Visual Studio 2017 [22]. The signal processing of the proposed system was implemented by C# scripts. AR markers recognition were developed with Vuforia Augmented Reality SDK [23].

Figure 2 – Worn Microsoft HoloLens.

2.3 The display system

Figure 3 shows the display procedure of the proposed system. The system displays a three-dimensional image source distribution by using the HoloLens in the following steps.

Step 1 Play time stretched pulse (TSP) signal and the impulse responses in the room are measured using closely located four-point microphones. On the other hand, microphone position is gotten with two AR markers and saved on HoloLens.

Step 2 Four sound signals are convoluted with inversed TSP signal which is played in the room.

Step 3 Four convoluted signals are calculated by closely located four-point microphone method. Coordinates of image source distribution are derived.

Step 4 Coordinates are sent from PC to HoloLens by User Datagram Protocol (UDP).

Step 5 In displaying the image source distribution, 3DCG spheres are placed at the positions of the image sources.

Step 6 The arrival times of the image sources are divided into five categories and represented by colors for each category.

Step 7 Change four-point microphones' position and repeat Steps 1)-6).

Step 8 Display each image source distribution at user's position with fiducial marker.

We used two AR markers to display image source distributions. One is as fiducial marker and the other is as microphone marker. The former one is set on the floor and is spatial reference. The latter one is set on four-point microphone and get relative coordinate between fiducial marker and four-point microphone. When the user worn HoloLens approaches to the measurement position, the proposed display system displays image source distribution at user's point by using fiducial marker.

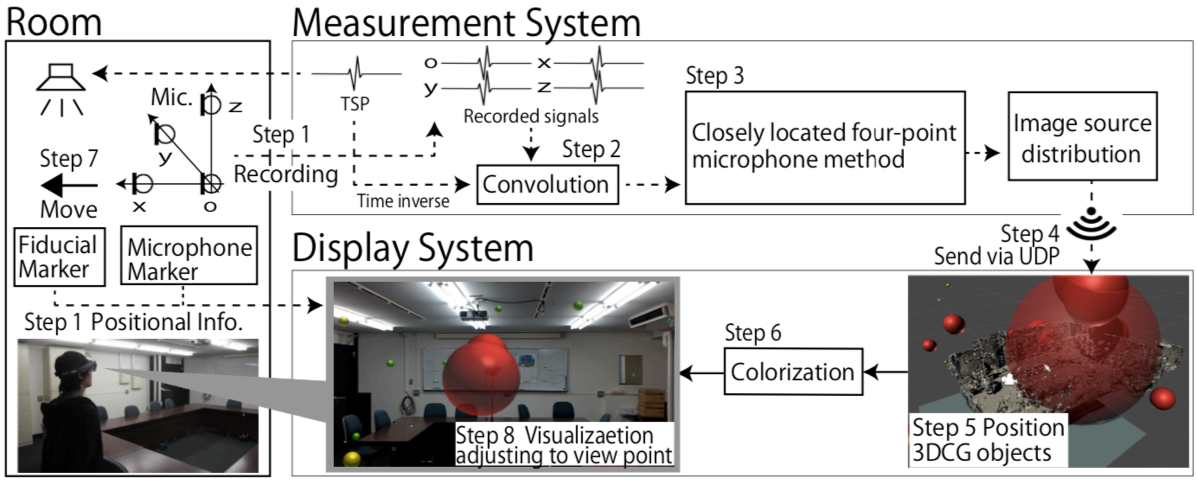

Figure 3 – Display procedure. First, the impulse responses in the room are measured using four microphones. On the other hand, the positions of fiducial marker and microphone marker are gotten. Second, image sources are estimated using the four-point microphone method to obtained the analysis results. Next, each image source is set as a sphere. Finally, the spheres are set on the fiducial marker and displayed on the HoloLens at measured point.

3. EXPERIMENTS

3.1 Measurement conditions and procedure

We measured the impulse responses and estimated the image source distributions at multiple points in an ordinary meeting room at Waseda University. Table.1 shows the measurement conditions. A loudspeaker was located at a distance of 1.5 m from the floor and 2.0 m from the nearest front wall. The four-point microphone was located at a distance of 1.2 m from the floor. We measured 24 points as shown in Fig. 4. Fiducial marker was set under the loudspeaker, in other words, at a distance of 0.0 m from the floor and 2.0 m from the nearest front wall. And microphone marker was set on four-point microphone.

We measured as per the following steps:

1) A TSP signal was played by the loudspeaker via an audio interface and a power amplifier.
2) The sound signals were recorded by four microphones with a PC via an audio interface.
3) After calculation of the impulse responses, the image source distribution was obtained by using the closely located four-point microphone method. On the other hand, the microphone position was gotten with two AR markers and HoloLens.
4) Change four-point microphone's position and repeat steps 1)-3).
5) After impulse responses were measured, each image source distribution at user's position was displayed on HoloLens.

Table 1 – Measurement conditions.

Measured room	59-04-15 Meeting room, Nishi-Waseda campus, Waseda University
Equipment	Opposed loudspeaker (Fostex FE204) Power amplifier (YAMAHA P4050) Audio interface (MOTU 8M) Microsoft HoloLens MacBook Pro (2.9GHz Intel Core i5, 8GB, 1600MHz DDR3) Four microphones (AUDIX TM1) AR markers (fiducial marker and microphone marker)
Distance between microphones [cm]	5.0
Measurement signal	Time stretched pulse
Sampling frequency [Hz]	48000
Interpolation for calculation	16

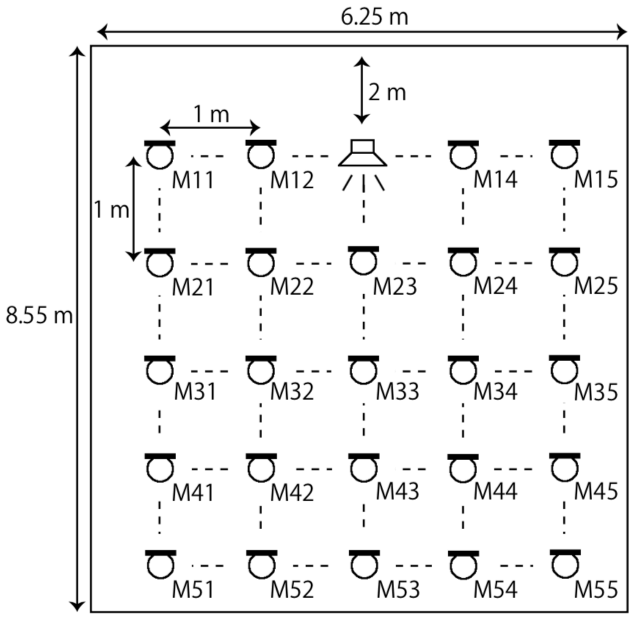

Figure 4 – Measured points in the room.

3.2 Results of experiment with the MR display

Figure 5 shows experimental results of image source distributions with the HoloLens. The left-side, center and right-side figures are the view of image source distribution at M32, M33 and M34 positions, respectively. The upper figures show the views when the user directed to the front wall. The lower figures show the views when the user directed to the left wall. To observe small sound reflections, logarithm of amplitude of each image source sphere is represented as diameter of sphere.

When a user approaches to measured position, image source distribution adjusting to the position was shown in the HoloLens display. Thus, we can understand where sound reflections occurred and how large the sound reflections were by the positional and size information of image sources. When a user approaches to the other measured position, we can understand how the feature of image source distribution changes.

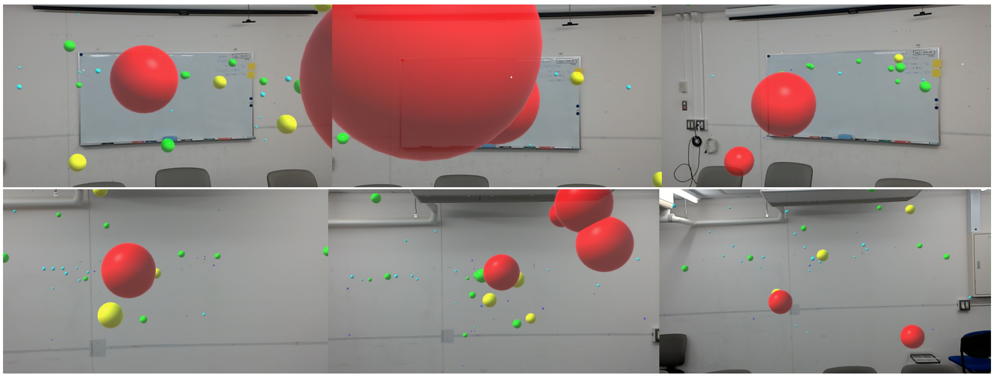

Figure 5 – View of image source distributions at different viewpoints.

4. CONCLUSIONS

In this paper, we proposed a MR display system for image source distributions measured at multiple measurement points. We measured impulse responses and estimate image source distributions at multiple points in the room by using closely located four-point microphone method and displayed them on our proposed MR display system. By using AR markers, position of measurement point can easily obtained. Then, the image source distribution can be observed at the different viewpoints. Thus, it helps us to understand how the sound reflections change.

In the future work, to evaluate whole feature of sound field completely, we will interpolate image source distributions from measured data to display image source distribution at any viewpoints. In addition, we will display image source distribution with other devices such as PC or tablets.

REFERENCES

[1]. J. D. Maynard, E. G. Williams, and Y. Lee, "Nearfield acoustic holography: I. theory of generalized holography and the development of nah," *J. Acoust. Soc. Am.*, **78**(4), 1395-1413, (1985).

[2]. S.-H. Park and Y.-H. Kim, "Visualization of pass-by noise by means of moving frame acoustic holography," *J. Acoust. Soc. Am.*, **110**(5), 2326-2330, (2001).

[3]. H. Kook, G. Moebs, P. Davis, and J. Bolton, "An efficient procedure for visualizing the sound field radiated by vehicles during standardized passby tests," *J. Sound and Vib.*, **223**(1), 137-156, (2000).

[4]. J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Proc. Mag.*, **19**(2), 30-39, (2002).

[5]. Y. Oikawa, M. Goto, Y. Ikeda, T. Takizawa, and Y. Yamasaki, "Sound field measurements based on reconstruction from laser projections," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, **4**, iv/661-iv/664, (2005).

[6]. K. Yatabe and Y. Oikawa, "PDE-based interpolation method for optically visualized sound field," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 4771-4775, (2014).

[7]. N. Chitanont, K. Yaginuma, K. Yatabe, and Y. Oikawa, "Visualization of sound field by means of Schlieren method with spatio-temporal filtering," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 509-513, (2005).

[8]. K. Yatabe and Y. Oikawa, "Optically visualized sound field reconstruction using Kirchhoff-Helmholtz equation," *Acoust. Sci & Tech.*, **36**(4), 351-354, (2015).

[9]. K. Yatabe and Y. Oikawa, "Optically visualized sound field reconstruction based on sparse selection of point sound sources," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 504-508, (2015).

[10]. Y. Ikeda, N. Okamoto, T. Konishi, Y. Oikawa, Y. Tokita, and Y. Yamasaki, "Observation of traveling wave with laser tomography," *Acoust. Sci. & Tech.*, **37**(5), 231-238, (2016).

[11]. K. Ishikawa, K. Yatabe, Y. Ikeda, Y. Oikawa, T. Onuma, H. Niwa, and M. Yoshii, "Interferometric imaging of acoustical phenomena using high-speed polarization camera and 4-step parallel phase-shifting technique," *Proc. Int. Congr. High-Speed Imaging Photonics (ICHSIP)*, 331-336, (2016).

[12]. Y. Oikawa, K. Yatabe, K. Ishikawa, and Y. Ikeda, "Optical sound field measurement and imaging using laser and high-speed camera," *Proc. Int. Congr. Noise Control Eng. (INTERNOISE 2016)*,

258-266, (2016).

[13]. K. Ishikawa, K. Yatabe, N. Chitanout, Y. Ikeda, Y. Oikawa, T. Onuma, H. Niwa, and M. Yoshii, "High-speed imaging of sound using parallel phase-shifting interferometry," *Opt. Express*, **24**(12), 12922-12932, (2016).

[14]. N. Chitanont, K. Yatabe, K. Ishikawa, and Y. Oikawa, "Spatio-temporal filter bank for visualizing audible sound field by Schlieren method," *Apple. Acoust.*, **115**, 109-120, (2017).

[15]. Y. Yamasaki and T. Itow, "Measurement of spatial information in sound fields by closely located four point microphone method," *J. Acoust. Soc. Jpn. (E)*, **10**(2), 101-110, (1989).

[16]. A. Inoue, Y. Ikeda, K. Yatabe, and Y. Oikawa, "Three-dimensional sound-field visualization system using head mounted display and stereo camera," *Proc. Mtgs. Acoust*., **29**(1), 025001, (2016).

[17]. A. Inoue, K. Yatabe, Y. Oikawa, and Y. Ikeda, "Visualization of 3D sound Field using See-Through Head Mounted Display," *SIGGRAPH '17*, Posters, (2017).

[18]. Y. Kataoka, W. Teraoka, Y. Oikawa, Y. Ikeda, "Real-time Measurement and Display System of 3D Sound Intensity Map using Optical See-Through Head Mounted Display," *SIGGRAPH Asia*, vol. Posters, pp.71, (2018).

[19]. W. Teraoka, Y. Kataoka, Y. Oikawa, Y. Ikeda, "Display System for Distribution of Virtual Image Sources by using Mixed Reality Technology," *Inter Noise*, pp. in18_1647, (2018).

[20]. https://www.microsoft.com/hololens

[21]. H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *IEEE Robot. &2Autom. Mag*., **13**, 99-110, (2016).

[22]. https//unity3d.com

[23]. https://developer.vuforia.com