# Midair Haptic Display to Human Upper Body

Shun Suzuki[1†], Ryoko Takahashi[2], Mitsuru Nakajima[1], Keisuke Hasegawa[2], Yasutoshi Makino[1] and Hiroyuki Shinoda[1]

[1]Graduate School of Frontier Sciences, University of Tokyo, Chiba, Japan
(Tel: +81-4-7136-3777; E-mail: {suzuki, nakajima}@hapis.k.u-tokyo.ac.jp,{Yasutoshi_Makino, Hiroyuki_Shinoda}@k.u-tokyo.ac.jp)
[2]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan
(Tel: +81-3-5841-6368; E-mail: takahashi@hapis.k.u-tokyo.ac.jp, Keisuke_Hasegawa@ipc.i.u-tokyo.ac.jp)

**Abstract:** In this paper, we present a calibration method for multi-unit midair haptic display that can cover a workspace larger than one-meter cube. In typical previous midair haptic displays, the workspace was smaller than a 30-cm cube, where the display target was limited to a hand and fingers. To achieve a large workspace covering a human upper body, we develop a system where arbitrarily located multiple display units with sensors can determine the relative positions and postures by a simple calibration process. In a prototype system using two units of 45 cm by 57 cm phased arrays, we examine the ultrasound convergence from the multiple units. We also show that the midair haptic stimulation is perceivable through clothes on the upper body.

**Keywords:** Human Interfaces, midair haptics

## 1. INTRODUCTION

Midair haptics using Airborne Ultrasound Tactile Display (AUTD) originating from [1] enables noncontact tactile feedback to a human body surface. Such a midair haptics alongside wearable haptics is a promising option to provide haptic feedback in 3D interfaces including VR/AR. The strength of midair haptics including an airflow device [2] is that no devices are required on the user. Especially in AUTD, the stimulation position is freely and precisely determined at arbitrary timing, and the amplitude modulation gives arbitrary temporal force patterns in the workspace.

AUTD is a phased array in which ultrasonic transducers are arranged typically in a lattice pattern [3, 4]. By controlling the phase of the transducers, AUTD creates ultrasound focus at an arbitrary position in the air. In HaptoMime developed by Momani et al. [5], an aerial touch panel with tactile feedback was prototyped. In HaptoClone developed by Makino et al. [6], 3D images of real objects were transferred to a nearby workspace and haptic interaction with the 3D images were produced. These studies showed the effect and necessity of haptic feedback that improved the usability and enhanced reality in interaction with aerial images. In these previous systems, hands and fingers are the main haptic display targets, and the workspace was limited to the range of a 30 cm cube.

The next intriguing step is to expand the workspace and examine the haptic effect in it. In this paper, we consider a one-meter cube workspace where a human upper body can freely move and 40 kHz ultrasound absorption
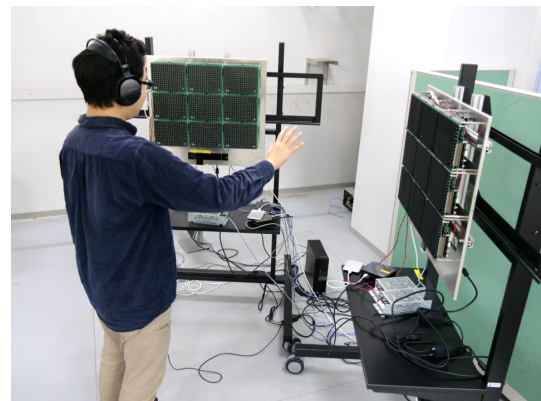
---

† Shun Suzuki is the presenter of this paper.



Fig. 1 Multi phased array unit system.

in the air is acceptable.

A practical problem in expanding the workspace is to digitize the positions and postures of the multiple ultrasound transducers. To obtain a desirable convergence in a large workspace, the workspace must be surrounded by many transducers. A practical way is to surround the workspace with multiple phased array units (AUTD units). Figure 1 shows a case of two AUTD units. A workspace is a flexibility designed by surrounding the workspace with multiple AUTD units of various directions. The problem posted here is that to create the focus of ultrasound in such a system, the control computer must know the positions and angles of all AUTD units precisely.

In this paper, we present a method to semi-automatically perform position/posture calibration of freely arranged AUTD units. Using the system, we can freely change and expand workspace. Such a

semi-automatic calibration system makes the workspace construction much easier and its application practical. Though there was a research to examine a large workspace [7], only a single large-aperture phased array was examined and the maximum distance between the phased array and focal point was 60 cm in the research.

Using this system, we set up a workspace and evaluate the ultrasound convergence at 1 m from AUTD units. In addition, we examine the haptic perception in the workspace and confirm the stimulus is perceivable at the chest through clothing.

## 2. SYSTEM CONFIGURATION

The system configuration used in this study is shown in Fig. 2. Each AUTD unit has a slave PC and depth sensor (Kinect v2, Microsoft Corp.) to detect the three-dimensional coordinates. The slave PCs are connected to the Master PC via LAN. All the units are synchronized and measure a point target simultaneously in each coordinate system and send the coordinate data of the point target to the master PC. The master PC calculates the positional relationship among the units and controls the AUTD units.

The appearance and design drawing of the unit is shown in Fig. 3. Nine AUTD subunits are fixed to an aluminum frame. An AUTD subunit has 249 transducers and an AUTD unit has 2241 ones in total. A depth sensor is also fixed to the frame. The positional relationship between the AUTD unit and depth sensor has been determined in the unit design. Therefore, by performing position calibration between the depth sensors, AUTD calibration is virtually completed. The weight of an AUTD unit is 8 kg. The unit is mounted on a TV stand with casters for convenience.
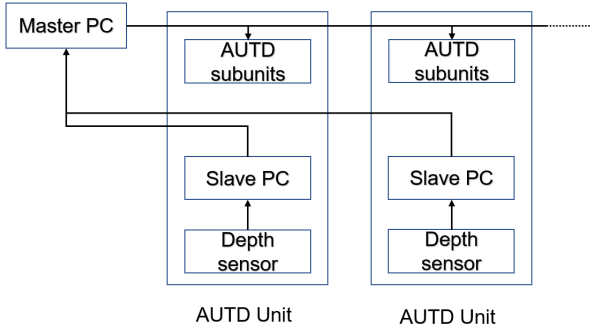


Fig. 2 System configuration.

## 3. CALIBRATION AMONG MULTIPLE ARRAYS

In this section, we describe the method of calculating the rotation matrix and the translation vector between two AUTD units. Extension to the general case of multiple AUTD units is straightforward, but we focus on a two-unit case since we have prototyped only the two-unit system. Next, we will describe the method of obtaining the point set using the depth sensors.

### 3.1. General

Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N\}, Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_N\}$ be sets of points measured in two coordinate systems fixed to two AUTD units. $N$ is the number of points. Suppose that correspondence between the two point-sets is given. We want to obtain the rotation matrix and the translation vector between these point sets. That is, we want to find a rotation matrix $R$ and a translation vector $\boldsymbol{t}$ that satisfies

$$\boldsymbol{y}_i = R\boldsymbol{x}_i + \boldsymbol{t} \tag{1}$$

for any $i$. There are varieties of methods to solve this problem. Here we use a method using singular value decomposition (SVD) [8]. First, we shift the center of $X$ and $Y$ to the origin as

$$\bar{X} = \{\boldsymbol{x}_1 - \boldsymbol{x}_C, \boldsymbol{x}_2 - \boldsymbol{x}_C, ..., \boldsymbol{x}_N - \boldsymbol{x}_C\}, \tag{2}$$
$$\bar{Y} = \{\boldsymbol{y}_1 - \boldsymbol{y}_C, \boldsymbol{y}_2 - \boldsymbol{y}_C, ..., \boldsymbol{y}_N - \boldsymbol{y}_C\}, \tag{3}$$
$$\boldsymbol{x}_C = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i, \ \boldsymbol{y}_C = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i. \tag{4}$$

Second, perform SVD of $\bar{X}\bar{Y}^{\mathrm{T}}$ as

$$\bar{X}\bar{Y}^{\mathrm{T}} = U\Sigma V^{\mathrm{T}} \tag{5}$$

which gives the rotation matrix $R$ as

$$R = V\mathrm{diag}\left[1, 1, \det(V^{\mathrm{T}}U)\right]U^{\mathrm{T}}. \tag{6}$$

Finally, the translation vector $\boldsymbol{t}$ is calculated as

$$\boldsymbol{t} = \boldsymbol{y}_C - R\boldsymbol{x}_C. \tag{7}$$

This method can easily extend to the case of $i$-th data has weight $w_i$. By using Tukey's biweight function as weight, we can remove outlier values. Tukey's biweight function $w_{\mathrm{Tukey}}(i)$ is defined as

$$w_{\mathrm{Tukey}}(i) = \begin{cases} \left[1 - \left(\frac{d_i}{t}\right)^2\right]^2 & (d_i < t) \\ 0 & (otherwise), \end{cases} \tag{8}$$
$$d_i = |\boldsymbol{x}_i - R^{\mathrm{T}}(\boldsymbol{y}_i - \boldsymbol{t})|, \tag{9}$$
$$t = 2 \times \frac{\mathrm{median}\{d_i\}}{0.6475}. \tag{10}$$

We remove outliers by the following method.
1. Initialize with $w_i = 1$.
2. Perform SVD of $\bar{X}W\bar{Y}^{\mathrm{T}} = U\Sigma V^{\mathrm{T}}$ with $W = \mathrm{diag}(w_1, ..., w_N)$ and calculate $R, \boldsymbol{t}$ by Eq. (6),(7).
3. Update weight with $w_i = w_{\mathrm{Tukey}}(i)$.
4. Repeat 2 and 3 until the weights convergence.

### 3.2. Optical measurement of point target

A simple and practical way to obtain above $X$ and $Y$ is to move a small marker (target object) in the workspace by hand and to measure the trajectory by the two AUTD units simultaneously. The three-dimensional position vector of the point target is obtained by the depth sensor (Kinect v2, Microsoft Corp.) using infrared rays.

A rhombicuboctahedron with the side of 10 mm as shown in Fig. 4 is used as the target object. To obtain a high contrast from the environment, the surface has a retroreflective film. We use the rhombicuboctahedron target instead of a sphere to stick a hard retroreflective film.
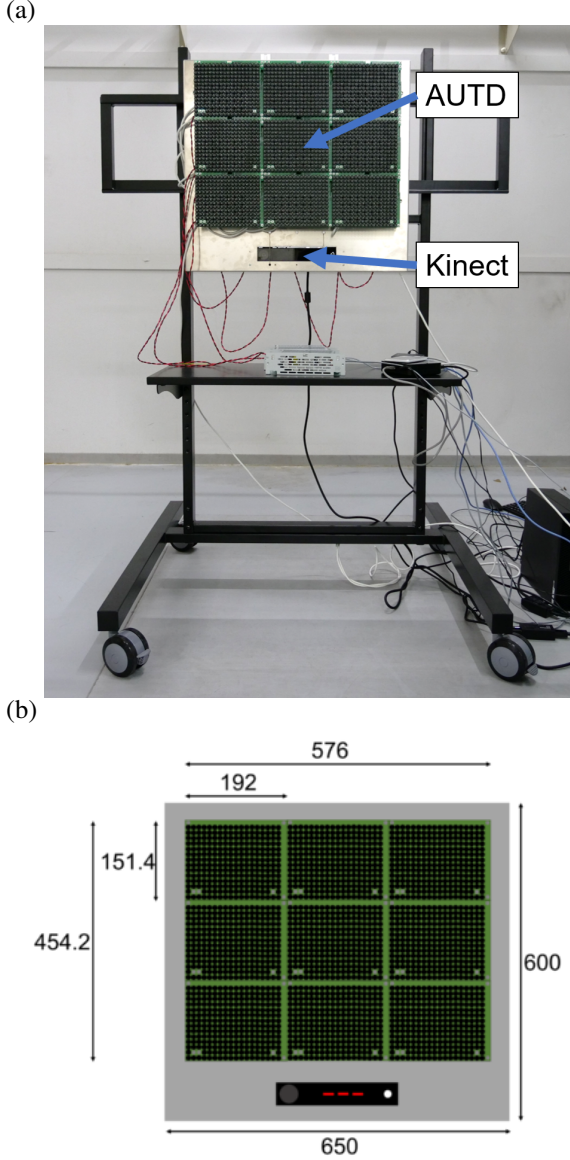
(a)

(b)



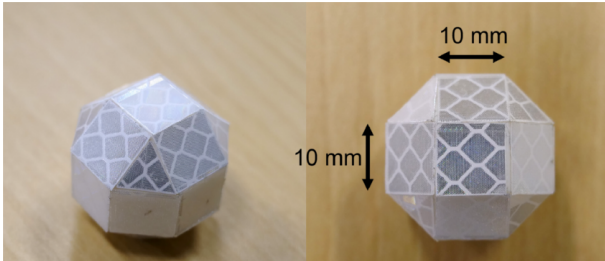Fig. 3 (a) Appearance and (b) drawing of Unit. Dimensions are in mm.



Fig. 4 Point target used in calibration.

Since the reflection from a retroreflective film can be too intense which causes saturation, we covered the infrared camera of the depth sensor with an infrared-cut film. Figure 5 shows an actually obtained depth image with the depth sensor.

Since the depth sensor acquires the distance to the target object surface, we estimate the center of the target
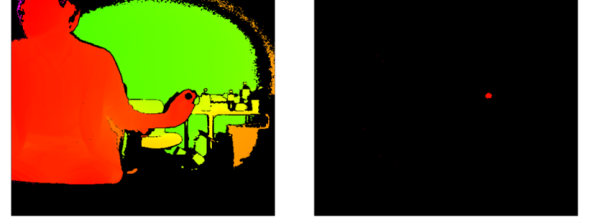


Fig. 5 (Left) Raw depth image. (Right) Image after covering the depth sensor with an infrared-cut film.

object from the obtained data. Suppose that a square part on the target object surface is perpendicular to the target direction and the depth sensor obtain the distance to the square part surface. In this case, we should estimate the target center $\boldsymbol{p}_C$ from the measured position vector $\boldsymbol{p}$ as

$$\boldsymbol{p}_C = \frac{|\boldsymbol{p}| + r}{|\boldsymbol{p}|}\boldsymbol{p}, \tag{11}$$

where $r = (5 + 10/\sqrt{2})\,\text{mm} = 12.07\,\text{mm}$ is the distance from the square part surface to the target center. In this experiment, we apply the above equation to the general cases.

To ensure the multiple depth sensors measure an identical point, we use time information. We synchronize the clock of the Slave PCs to that of Master PC and store the timestamp with the target coordinates taken by the depth sensor. We identify the corresponding point data of the other depth sensor by the timestamp. The point data having the closest timestamp is considered as the corresponding point.

## 4. EXPERIMENT ON FOCUSING

In this section, we set up the system, conduct the proposed calibration and confirm the ultrasound convergence by acoustic field measurement using a standard microphone.

### 4.1. Experimental Setup

The experimental setup is shown in Fig. 6. Two units are roughly arranged, without measuring the precise positional relationship in advance. First, after arranging the units, calibration is performed by the method described in the previous section. The point target was moved in the workspace randomly so that the trajectory covers the edges of the workspace. The number of points used for the calibration was 119 points, and the required time was about 3 minutes. The positional relationships of the units obtained as a result of the calibration are shown in the Fig. 7. Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_{119}\}, Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{119}\}$ be point sets obtained by depth sensor. The root mean square error $\sigma$ was

$$\sigma = \sqrt{\frac{1}{119}\sum_{i=1}^{119}\left[\boldsymbol{y}_i - (R\boldsymbol{x}_i + t)\right]^2} = 1.8\,\text{mm} \tag{12}$$

where $R$ and $\boldsymbol{t}$ are rotation matrix and translation vector calculated from $X$ and $Y$, respectively. This means the
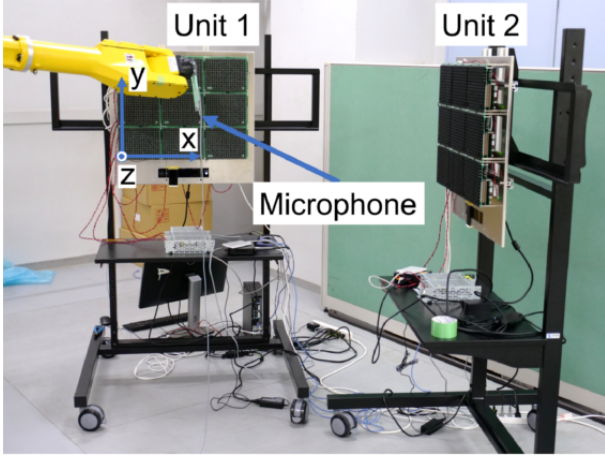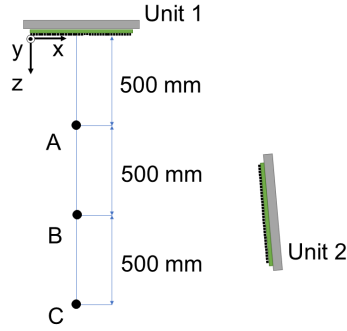
Fig. 6 Experimental setup.



Fig. 7 System coordinates.



Fig. 8 Result of focusing. The acoustic energy density around points A, B, C shown in Fig. 4.

error of the focal position will be about $\sigma$ if the error of the depth sensor is uncorrelated between the two AUTD units. Second, a focal point of the convergent ultrasound is formed at three points A, B, and C as shown in Fig. 7.

### 4.2. Result

Figure 8 shows the acoustic energy density measured around the focal point. Figures (A), (B), and (C) show the results for the focal points produced at A, B, and C, respectively. The acoustic energy density was measured by a microphone attached to a scanning robot arm as shown in Fig. 6. The $x, y$ coordinates of A, B, and C are commonly $(x, y) = (302, 225)\,\mathrm{mm}$, and $z$ of A, B, and C are $500, 1000$, and $1500\,\mathrm{mm}$, respectively. The acoustic energy density (elasticity energy density) $e$ proportional to the radiation pressure is calculated as

$$e = \frac{p^2}{\rho c^2} \tag{13}$$

where $\rho$ is measured sound pressure, $\rho$ is the air density $(1.2\,\mathrm{kg/m^3})$ and $c$ is the sound speed at room temperature $(344\,\mathrm{m/s})$. This value $e$ provides the rough estimation of radiation pressure when a sound reflector is placed at the point. For example, $e = 100\,\mathrm{J/m^3} = 100\,\mathrm{Pa} = 10\,\mathrm{mN/cm^2}$ indicates that the radiation pressure on the reflector is approximately $10\,\mathrm{mN/cm^2}$ though the direction of the reflector is not considered.
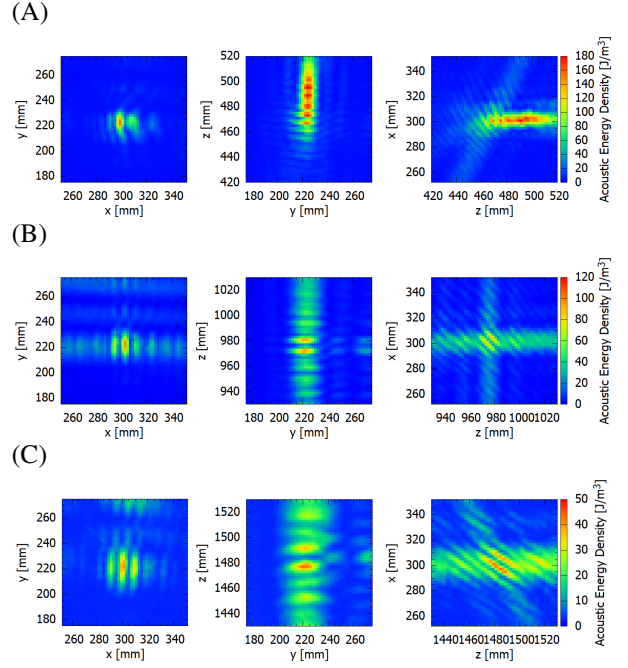
Let $e_{\max}$ be the maximum acoustic energy density

around point B. On each plane, the region of points where the acoustic energy density is higher than $e_{\max}/\sqrt{2}$ falls within a range of a circle with a diameter of $13\,\mathrm{mm}$. And, the same region around point C falls within a range of a circle with a diameter of $30\,\mathrm{mm}$.

As shown in Fig. 7, the $z$ coordinate of the pressure maximum deviates from the theoretical value by 20-30 mm. This means that the position of unit 2 to unit 1 in calibration is deviated about 20 mm from the true position. The fact that error $\sigma$ in Eq. (12) is 1.8 mm and there is no deviation in the height of the beam indicate that the depth sensor has an offset in the z-axis direction. Since the measurement error of Kinect v2 is 2-4 mm in the range used in this method [9], the offset seems to have occurred due to the retroreflective film and the IR cut filter. It is necessary to verify the error of the depth sensor and search for an appropriate calibration method for it.

## 5. UPPER BODY STIMULATION

In this section, we investigate the haptic perception of a human upper body, especially the effect of wearing clothes. We create a focus of converging ultrasonic waves at the center of the chest and examine whether we can display haptic sense through clothing.

### 5.1. Experimental Setup

The experimental setup is shown in Fig. 9. The two AUTD units are calibrated by the procedure of the previous experiment. The participant stands in front of the two units, and a focal point is placed at about 1 m away from each of the two units. The participants wear headphones to prevent ears from the ultrasound. In addition, pink noise is applied from the headphones for eliminat-
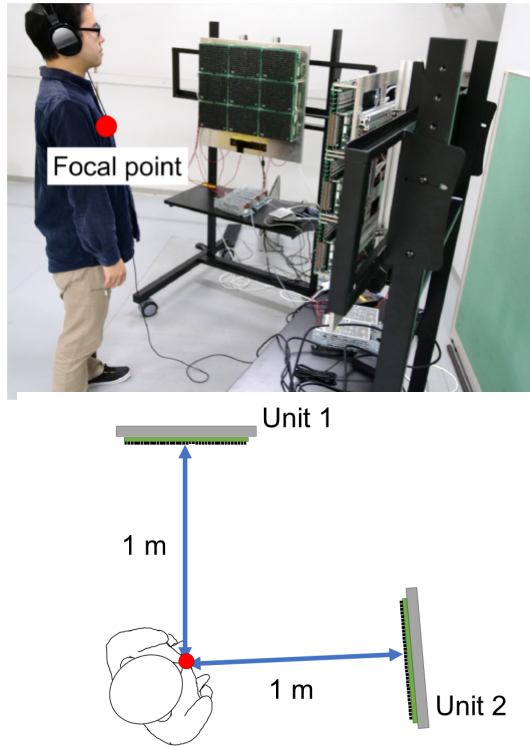
Fig. 9 Focal point of ultrasound projected at center of chest through clothing.



Fig. 10 Result of haptic display to chest through clothing.

ing the operation sound of AUTD. Participants keep their eyes closed. A focal point of ultrasound is projected to the participant's chest. Ultrasound is modulated into 25, 50, and 100 Hz sine waves and projected ten times respectively in random order. We choose these frequencies in the frequency range of Merkel cells and Rufini corpuscle. A focal point of ultrasound is displayed for 1 second, and the participants are asked to report by raising their hand when they felt haptic sense. The participants are wearing their own clothes, and they conducted experiments in the two conditions: wearing an inner wear only and wearing an inner wear and an outerwear. Note that we don't controll the clothes. Seven men in their twenties and a man in his thirty participated in this experiment. The experimental procedure was approved by the Ethical Committee of the University of Tokyo.

**5.2. Result**

The results are shown in Fig. 10. The result shows that haptic sensation is displayable to the chest through a single innerwear. On the other hand, in the condition of an outerwear on an inner wear, the correct answer rate decreased, and the perception was unstable. In all the cases, 100 Hz sine modulation gave the highest correct answer rate. When the frequency was 100 Hz, a significant difference was seen between Inner and Inner + Outer with the $p-$value of $4.2 \times 10^{-3}$.
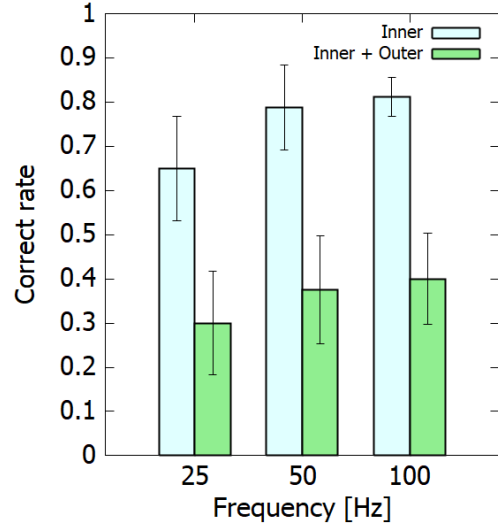
## 6. CONCLUSION

In this paper, we presented a calibration method for co-operative operation of multiple AUTD units. We demonstrated that we can expand the workspace of the midair haptic display by a prototype system, and we confirmed haptic sensation is displayable to the chest through an inner wear, at a distance of 1 m from the device. In the current system, the perception was unstable in the case of wearing an outerwear on the innerwear.

Since the relative positions and angles among multiple AUTD units are semi-automatically measured by the sensors on the units, it is straightforward to increase the number of the devices surrounding the workspace. In future works, we will increase the number of AUTDs and examine the effect of midair haptics with sufficient intensity in a large workspace. In addition, we improve the calibration method into a method measuring acoustic intensity at the target object.

## ACKNOWLEDGEMENTS

## REFERENCES

[1] T. Iwamoto, M. Tatezono, and H. Shinoda, "Noncontact Method for Producing Tactile Sensation Using Airborne Ultrasound," *Proc. Eurohaptics 2008*, pp.504–513, 2008.

[2] R. Sodhi, I. Poupyrev, M. Glisson, A. Israr, "AIREAL: interactive tactile experiences in free air," *ACM Transactions on Graphics (TOG)*, Vol. 32, No. 4, Article No.134, 2013.

[3] T. Hoshi, M. Takahashi, T. Iwamoto, and H. Shinoda, "Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound," *IEEE Trans. on Haptics*, Vol. 3, No. 3, pp.155–165, 2010.

[4] T. Carter, S. A. Seah, B. Long, B. Drinkwater, S. Subramanian, "UltraHaptics: multi-point mid-air haptic feedback for touch surfaces," *Proc. of the 26th ACM symposium on user interface software and technology (UIST '13)*, pp.505–514, 2013.

[5] Y. Monnai, K. Hasegawa, M. Fujiwara, K. Yoshino, S. Inoue, and H. Shinoda, "HaptoMime: Mid-Air Haptic Interaction with a Floating Virtual Screen," *Proc. 27th Annu. ACM Symp. User interface Softw. Technol.*, pp.663–667, 2014.

[6] Y. Makino, Y. Furuyama, S. Inoue, and H. Shinoda, "HaptoClone (Haptic-Optical Clone) for Mutual Tele-Environment by Real-time 3D Image Transfer with Midair Force Feedback," *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst.*, pp.1980–1990, 2016.

[7] K. Hasegawa, H. Shinoda, "Aerial Display of Vibrotactile Sensation with High Spatial-Temporal Resolution using Large-Aperture Airborne Ultrasound Phased Array," *Proc. IEEE World Haptics Conference 2013*, pp.31–36, 2013.

[8] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp.698–700, 1987.

[9] L. Yang, L. Zhang, H. Dong, A. Alelaiw and A. E. Saddik, "Evaluating and improving the depth accuracy of Kinect for Windows v2," *IEEE Sensors Journal*, vol. 15, no. 8, pp.4275–4285, 1015.