


Aerial Vibrotactile Display Based on Multiunit Ultrasound Phased Array

Keisuke Hasegawa , *Member, IEEE* and Hiroyuki Shinoda, *Member, IEEE*

Abstract—In this paper, we report on an airborne vibrotactile display with a multiunit ultrasound phased array synthetic aperture. The system generates an ultrasound field with a location-tunable focus in the air, which exerts time-variant acoustic radiation pressure on the users skin, resulting in perceivable localized vibrotactile stimuli. The paper contains three major new contributions from previous related works. The first is an experimental validation of large-aperture focusing with improved synchronization offering an enlarged workspace in which sufficient acoustic power concentration is guaranteed. From the experiments, it is expected that perceivable vibrotactile focus can be generated 1 m away from a four-unit array system. The second is an experimental evaluation of the presented pressure for producing a broad variety of tactile perception, which shows that the generated ultrasound focus can serve as an vibrotactile actuator that has flat frequency characteristics in the domain of perceptual stimuli. The third is a psychophysical result of the detection threshold curve for sinusoidal stimuli offered by the system. The obtained curve shows similarity with conventionally known results, which have minimum values at approximately 200 Hz.

Index Terms—Airborne ultrasound, vibrotactile display, radiation pressure

1 INTRODUCTION

1.1 Advent of Mid-Air Tactile Technologies

ONE of the unique properties of the tactile modality is that it handles physical interactions through contact with skin surfaces. On the contrary, the visual and auditory modalities receive optic and acoustic stimuli, which enable us to perceive physical events apart from us. This is owing to their physical characteristics of these stimuli as propagating waves. Users can receive audiovisual information regardless of their posture and location around displays without any physical constraints. The same is not the case with most tactile displays, because they require physical contact with body surfaces. This is inconvenient for practical use and constrains users' physical movements once these tactile displays are attached to users' bodies. Moreover, these devices can present stimuli only on the regions in contact with the users.

Mid-air touch displays are one of the solutions free from these limitations. There have been several studies focusing on the physical principle of producing mid-air tactile stimuli. The most straightforward method is the use of air flow [1], which generates strong pressure near jet nozzles. However, because air spreads once it is emitted from nozzles, it is in principle difficult to achieve localized perceivable tactile stimuli at a significant distance from the device. There is

another approach that ejects impulsive air lumps for increasing the range. These air lumps are called vortex rings, which can travel over a significant distance while remaining isolated from the surrounding air. Aerial [2] uses this technique to remotely generate fluffy texture onto the bare skin of the users. Although a point-like stimulation would be difficult to produce, a vortex ring of several centimeters in diameter was produced. It also should be noted that no visual interference occurred with these technique, which enables the superposition of such systems onto images without optical occlusion.

There are several proposals for mid-air tactile displays that are different from the "air-based" techniques introduced above. A quasimid-air tactile technology is also proposed that converts energy conveyed through the air into tactile stimuli via passive materials mounted on the skin surface. One such example makes use of a photoelastic material for converting irradiated laser light into vibrotactile stimuli [3]. Because laser light travels over a great distance while maintaining its spot size and is easy to accurately localize, this technique guarantees a vast workspace as long as there are no laser blockages.

Another completely different challenge is mid-air tactile stimuli produced by direct laser irradiation to the skin for relatively short period [4]. It also generates a mid-air image as well as stinging texture in the air. Unlike the laser-based technique described above, this method supports bare-hand visuotactile interaction.

1.2 Widespread Use of Ultrasound

There is another potent method for realizing localized mid-air tactile stimuli on bare skin, which is the use of focused ultrasound. This technology was originally proposed and demonstrated by Iwamoto et al. [5] with an array of ultrasound transducers creating a localized focus in the air.

• The authors are with the Graduate School of Frontier Sciences, University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba-ken 277-8563, Japan.
E-mail: Keisuke_hasegawa@ipci.u-tokyo.ac.jp, Hiroyuki_Shinoda@k.u-tokyo.ac.jp

Manuscript received 7 Sept. 2017; revised 21 Dec. 2017; accepted 22 Jan. 2018. Date of publication 30 Jan. 2018; date of current version 14 Sept. 2018.
(Corresponding author: Keisuke Hasegawa.)

Recommended for acceptance by C. Duriez.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TOH.2018.2799220

Although ultrasound emitted from each transducer spreads as spherical wave, the entire output acoustic energy of the array can converge into a focus by controlling the phase delay on each transducer. This technique of realizing steerable directivity by integrating multiple wave sources is called the “phased array” technique, which was originally used in radio wave communications. The highly condensed acoustic energy yields “radiation pressure,” a nonlinear acoustic phenomenon that exerts a static force on the surface of an object blocking the sound propagation [6]. Based on this physical principle, skin deformation is caused by focused ultrasound and is felt as touch sensations. Because ultrasound is free from optical occlusion, a number of visuo-tactile applications have been proposed [7], [8], [9].

The ultrasound-based framework possesses several specific advantages owing to its physical properties as propagating waves. First, the minimum possible latency of the technique is determined by the sound velocity, which is at least tens of times faster than that of vortex rings. Supposing the sound velocity to be 340 m/s at room temperature on the ground, it takes approximately 3 ms for the ultrasound emitted from a phased array to reach users standing at a distance of one meter from the array. Okamoto et al. have shown that the human noticeability threshold of vibrotactile latency is 40 ms [10], which corresponds to a distance of 13 m from ultrasound sources. Thus, a low-latency system in a considerably large workspace is guaranteed in principle.

Second, the spatial and temporal variation of the produced tactile texture has been reported. Several studies have shown that an appropriate combination of phase delays for individual transducers can generate a spatial distribution of radiation pressure, which is obtained as a solution of inverse problems [11], [12], [13], [14], [15]. Inoue et al. have experimentally demonstrated that the produced pressure distribution yields volumetric textures that can be touched with freely exploring fingers [13], [16]. These textures have a spatial resolution determined by the wavelength. For example, the resolution is 8.6 mm at 20 °C when the frequency is 40 kHz. For temporal textures, the amplitude modulation of ultrasound waves creates vibrotactile sensations according to the AM envelope of the waveform [17]. It should be noted that the frequency characteristics of the “displayed” vibration are in principle expected to be flat over all perceivable frequencies (0–1000 Hz), which has been shown experimentally, as discussed in the following part of the paper.

Third, the produced stimuli show excellent reproducibility. This is owing to the wave-like physical nature of ultrasound, which travels over distance while maintaining its temporal structure. Compared to air flows and vortex rings, the behavior of ultrasound exhibits much less unpredictability.

1.3 Wide-Band Vibration for Various Tactile Experiences

Our proposed technique focuses on vibrotactile sensation on a bare skin surface. In other words, it cannot generate a strong static force or torque that leads to a kinesthetic experience. Nevertheless, it should be noted that vibrotactile sensation often offers various types of realistic tactile experiences without perfect physical reconstruction of the corresponding phenomena, when its waveform is properly designed.

Hayward et al. demonstrated that a rod-shaped device with a vibrator can make users perceive a “virtual rolling stone” inside it [18]. TECHTILE Toolkit [19] offers a framework to add vibrations to generic objects held by the users, which are activated in very natural manners. There are many other examples offering realistic and practical vibrotactile experiences [20], [21]. Thus, the vibration-based approach can be regarded as a promising tactile display method that has attracted broad interest among researchers.

The common feature of these technique is that the actuators produce broad-band vibrations, which is to say that a high-Q vibrator that only offers a sinusoidal vibration cannot yield this kind of realistic tactile experience. As stated above, our technique could yield such vibrations, because of the flat frequency characteristics of the produced vibration.

1.4 Contribution of the Paper

As described with respect to several previous studies, the mid-air vibrotactile technique is one of the most promising methods to broaden and diversify the use of tactile technologies. To make the most of the technique, we have fabricated a multiunit phased array system that widens the workspace in which the generated focus reaches while maintaining its convergence. By integrating multiple array units, we have an enlarged emission aperture, which guarantees unblurred focusing in longer region than was allowed with a single array unit [22]. There is another similarly motivated work based on a multiunit strategy that intends to produce three-dimensional touchable midair shapes [23].

This study is based on a conference proceeding that reported the above multiunit system [17], yet with the following improvements in the implementation and the performance assessment:

- We implement periodic synchronization among units. Each unit generates ultrasound based on its internal clock, and there are slight frequency differences among the units. The previous system [17] was incomplete in the synchronization originating from the clock variation. In this study, we address the previous problem by shortening the period of synchronization with a common external signal. As a result, the focal amplitude has been intensified.
- We perform spatial measurements of the acoustic amplitude distribution around focal regions with a variety of focal depths. We conduct two sets of measurements: one with a single phased array unit activated and another with four units activated. These measurements have shown that an enlarged aperture significantly improves both the focal intensity and concentration, particularly in regions distant from the emission plane. This is theoretically obvious, but no experimental data have been reported that ensure the focusing by a large aperture.
- We conduct psychophysical studies to demonstrate that the system can offer vibrotactile stimuli with a wide variety of waveforms. The system provides participants with sinusoidal and rectangular vibrations. We measure the detection threshold for several vibrational frequencies and find that for sinusoidal stimuli, the obtained detection curve shows similarity with

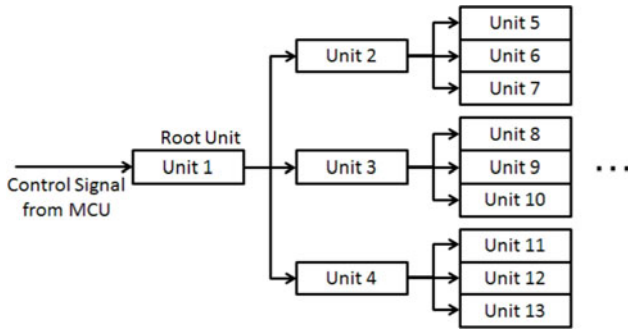


Fig. 1. Network topology of the fabricated multiple-AUTD system.

the conventionally obtained curves from the literature. We also find that the obtained detection thresholds are much lower with rectangular pulses when compared with the sinusoidal pulses of the same frequency.

2 PHYSICAL PRINCIPLE

The contents of this section have much in common with the previous relevant work [24]. We briefly summarize these for the sake of understandability of the subsequent parts of this section.

2.1 Acoustic Radiation Pressure

The proposed technique makes use of acoustic radiation pressure, which appears as static pressure apart from the time-variant sinusoidal sound pressure when the sound energy is considerably intense. The radiation pressure P [pa] is proportional to the square of the sound pressure p [pa]

$$P = \alpha \frac{p^2}{\rho c^2}, \quad (1)$$

where ρ [kg/m³] is the density of the medium, c [m/s] is the sound velocity in the medium, and α [-] is a coefficient whose value is between 1 and 2, as determined by the acoustic impedance ratio between the medium and the pressed object[6]. In the case of perfect reflection in an open space, which is applicable to most solid objects irradiated by focused airborne ultrasound, the value of α is set to 2.

2.2 Phased Array Technique

The phased array technique is a method of integrating multiple point wave sources in an array to produce a localized spatial distribution of energy. Let r_i be the three-dimensional position of the i th transducer among N transducers and r_f be the desirable three-dimensional (including the depth of the focus) focal position. The phase shift of each transducer θ_i is calculated so that it compensates the phase delay through distance

$$\theta_i = k|r_i - r_f|. \quad (2)$$

Here, k denotes the wavenumber. This technique enables the generation of a spot of radiation pressure at an arbitrary position in the air.

2.3 Aperture and Focusing Performance

A phased array of ultrasound transducers acts as an ultrasound lens to produce a focal spot. Its focal point is variable,

as mentioned above. As the focusing performance of an optical lens depends on its aperture size, the size of an ultrasound phased array directly limits the depth range of proper focusing. In general, according to the distance from the emission surface, the focus is blurred and the perceivable intensity of the vibrotactile stimuli diminishes. Since the degree of this blurring becomes more evident with a smaller aperture, creating a large aperture is a straightforward solution for this issue, which can be achieved by developing a multiunit system.

2.4 Amplitude Modulation of Sound Pressure

The quantitative relation between the sound pressure and the acoustic radiation pressure (1) holds temporally if the amplitude of the sound pressure p changes slowly enough compared to its frequency. Suppose that $p(t)$ is the product of a carrier sinusoidal wave $p_0 \sin(\omega_c t)$ and modulating wave $f(t)$

$$p(t) = p_0 \sin(\omega_c t) f(t). \quad (3)$$

Here, $f(t)$ yields the envelope of the carrier sinusoidal wave $p_0 \sin(\omega_c t)$. We assume that the highest frequency component of the envelope $f(t)$ is ω_f and $\omega_f \ll \omega_c$. In this case, the amplitude of $p(t)$ varies with $f(t)$. Then, Eq. (1) can be rewritten as

$$P(t) = \alpha \frac{p_0^2 [f(t)]^2}{\rho c^2}. \quad (4)$$

The important point here is that the resulting radiation pressure is proportional to the squared envelope of the output ultrasound waveform. As described in the following, desirable vibrotactile stimuli from the radiation pressure are obtained by appropriately designing the waveform envelope.

3 HARDWARE ARCHITECTURE

3.1 Signal Flow

We have developed a new airborne ultrasound tactile display (AUTD) system comprising multiple ultrasound transducer units. Each AUTD unit has a single input signal port and three output ports. The signal transmitted to the input port comes out from all of the output ports with a delay of 40 ns. Because this corresponds to a phase delay of less than 1 degree in the 40 kHz ultrasound employed in the fabricated system, thousands of units can be chained in a virtually synchronized manner. These ports can be mutually connected via eight-pin LAN cables. Multiple AUTD units are connected in a tree network topology, such that signals entering the input port of the root node unit are relayed to every other unit. This signal network structure is depicted in Fig. 1. The ports are connected to a field-programmable gate array (FPGA) equipped on a unit. The FPGAs employed here were from the Altera Cyclone III series. Each unit contains five FPGAs and the referenced one is known as the “master” FPGA of a unit. The remaining four are “slave” FPGAs, which receive the internal control signal from the master FPGA and apply driving signals to the transducers of the unit.

The whole multiple-AUTD system comprises several AUTD units connected as described above, as well as a microcontroller unit (MCU) and a computer. We used

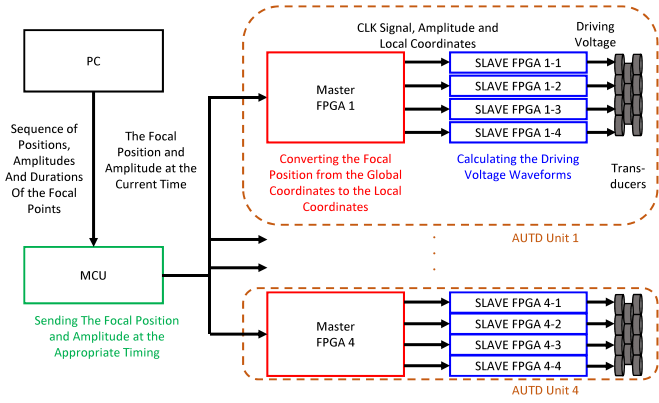


Fig. 2. Signal flow of the multiple-AUTD system.

mbed LPC1768 MCUs for the system. The MCU receives the control signals from the computer and transmits the focus signal packets into the input port of the root AUTD unit. The signal packet contains the three-dimensional focal position (48 bits in all, including 16 bits each for x , y , and z coordinates), amplitude (9 bits, 320 levels) and the header and footer bits. Each unit receives the sequence of the packet transmitted from the MCU at precise intervals, namely $500 \mu\text{s}$. The focal position and amplitude change at every received focus packet. The control signal conveys several commands to the MCU, such as generating/removing the vibrotactile waveform stored on the internal memory of MCU, initializing the positional setup of each AUTD unit, and so forth.

A diagram of the entire signal flow is shown in Fig. 2. The slave FPGAs are directly connected to the transducers and they calculate the proper phase delay of each transducer for focusing at the desired location. They receive the focal position and amplitude from the master FPGA. The focal positions in the packet are defined in a global coordinate system. The master FPGAs convert them to the local coordinates according to the position of each AUTD unit in the global coordinates.

3.2 Necessity of Intra-Unit and Inter-Unit Synchronization

It is required for the AUTD system that every transducer in every AUTD unit must be driven synchronized. In the previous system [14], the synchronization of each unit to the common clock signal was done only upon activating the system. As a result, it yielded incomplete focusing for a fixed focus, where the acoustic pressure had a beat, as shown in Fig. 3. In this study, we newly implemented inter-unit synchronization every $500 \mu\text{s}$, which ensured that the delay variation was less than $1 \mu\text{s}$ and produced stable and intensified focusing. (Fig. 3).

3.3 Generating Rectangle Pulse Trains Controlling Output Pressure

As described in Section 2.4, the temporal profile of the vibrotactile stimuli is created by the envelope of the carrier sinusoidal ultrasound wave. Their relation is given by Eq. (4). Here, we describe a practical method to generate the output ultrasound wave with its desired temporal profile.

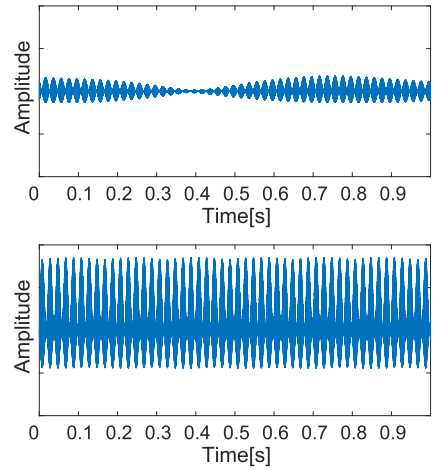


Fig. 3. Observed waveform of acoustic amplitude at the focal point without (upper) and with (lower) inter-unit synchronization. The beat effect is observed to be cancelled when the synchronization is enabled.

As detailed in [24], a pulse width modulation (PWM)-based technique is employed in the system. A rectangle pulse of period T [s] includes its $f_0 = 1/T$ [Hz] component by the amplitude

$$p = p_M |\sin(\pi d)|, \quad (5)$$

where d denotes the duty cycle and p_M denotes the amplitude when $d = 1/2$. This is obtained by a Fourier series decomposition. It indicates that we can control the ultrasound amplitude at $f_0 = 40$ kHz by the duty cycles. Once the duty cycle and time delay τ [s] are determined, the driving signal $V(t)$ [V] of the transducer is given by

$$V(t) = \begin{cases} V_M, & (nT - \tau < t < nT + dT - \tau) \\ 0, & (nT + dT - \tau < t < (n+1)T - \tau) \end{cases}, \quad (6)$$

where $T = 1/f$ [s] is the period of the ultrasound and n is an arbitrary integer. Here, the relation between V_M and p_M is experimentally determined.

3.4 Specification

The fabricated AUTD array unit is shown in Fig. 4. It is implemented with 249 ultrasound transducers, resulting in an emission aperture of $180 \text{ mm} \times 140 \text{ mm}$. Two of the four edges of a unit have 10 mm marginal spaces with no transducers. The maximum gross static force generated by stationary focused ultrasound onto a $12 \text{ mm} \times 12 \text{ mm}$ region using four units was approximately 20 mN with all transducers driven at the maximum power, consuming 0.2 W per transducer. The total radiation force applied to the reflecting surface was 40.8 mN, which was measured with an electronic scale. We experimentally confirmed that the focal location was controllable horizontally in less than 1 mm by actually generating a moving focus at a water surface, which is sufficient compared to the perceptual spatial resolution for vibrotactile ultrasound stimuli [25]. The FPGAs used in each unit were supplied with 25.6 MHz clock signals on the individual AUTD unit. The total power consumption per unit was estimated to be approximately 50 W at the maximum power.

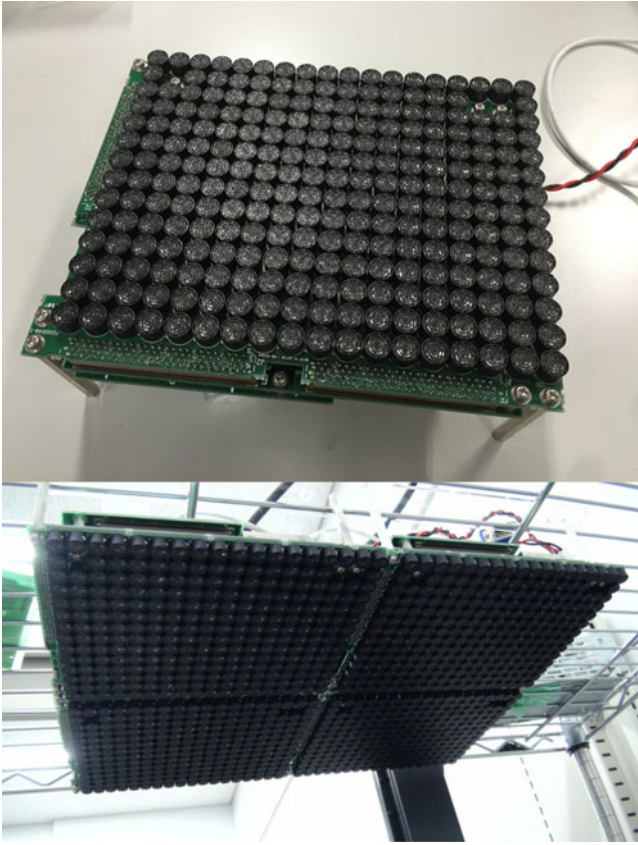


Fig. 4. Ultrasound transducer array unit (upper) and a four-unit emission aperture (lower).

4 VIBROTACTILE SIGNAL DESIGN

4.1 Creating Vibrotactile Sensation with Pulse Width Modulation

In the previous section, we treated the case in which the duty cycle of the input pulse train is temporally constant. According to Eq. (4), the resulting acoustic waveform is controlled by the duty cycle.

From Eqs. (4) and (5), when the duty cycle is expressed as a time-variant function $d(t)$, the resulting radiation pressure can be expressed as

$$\begin{aligned} P_R(t) &= \frac{\alpha}{\rho c^2} \{p(t)\}^2 \\ &= \frac{\alpha}{\rho c^2} p_M^2 \sin^2(\pi d(t)). \end{aligned}$$

Using this relation, the driving pattern $d(t)$ for a given radiation pressure waveform $P_R(t)$ is given by

$$d(t) = \frac{1}{\pi} \sin^{-1} \left(\sqrt{\frac{\rho}{\alpha} \frac{c}{p_M}} P_R(t)^{\frac{1}{2}} \right). \quad (7)$$

The term in the inverse sine function is a normalized output waveform of the radiation pressure, whose value is expected to be between -1 and 1 . The relation between the three functions $V(t)$, $p(t)$, and $P_R(t)$ is depicted in Fig. 5. Thus, the time-variant radiation pressure with a desired temporal profile is generated by varying the duty cycle of the driving voltages in accordance with the above equation. Note that negative pressure cannot be generated by

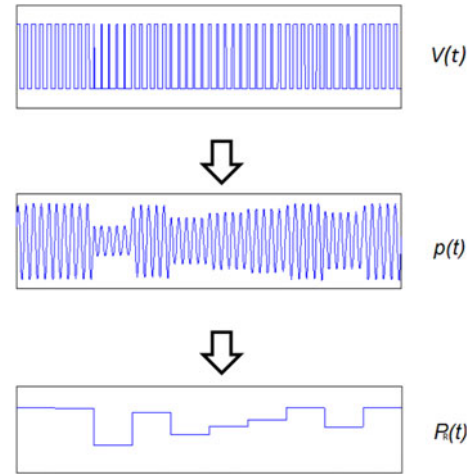


Fig. 5. Relation between the driving signal $V(t)$, sound amplitude $p(t)$, and time-variant acoustic radiation pressure $P_R(t)$.

radiation pressure. Designing the optimal vibration to produce tactile perception with the highest fidelity remains as an essential future task.

4.2 FIR Filtering of the Series of Duty Cycles

The MCU updates the focal position and amplitude at 2 kHz, which results in 20 sequential pulses of constant duty cycle in the output voltage. This means that the envelope waveform of the output ultrasound is stair-shaped, containing harmonic components of its fundamental frequency (2 kHz). These harmonic components in the envelope waveform yield audible noises.

To attenuate the noises, it is effective to tune the fundamental frequency of the envelope to its maximum possible value, namely 40 kHz. This is equivalent to making the envelope waveform as smooth as possible by switching the duty cycle of each PWM pulse. Consequently, the harmonic components in the envelope are suppressed.

We implemented a finite impulse response (FIR) filtering scheme in the FPGAs on the units that yielded a series of duty cycles varying at every pulse out of the input signals from the MCU. In other words, the FIR resamples the discrete duty cycle trains originally sampled at 2 kHz into 40 kHz with low-pass filtering.

5 MEASUREMENTS OF ACOUSTIC RADIATION PRESSURE

In this section, we show several measurements of generated ultrasound focus to assess the physical properties of the fabricated system. In all these experiments, the waveform of the acoustic pressure was recorded by a standard microphone (Brüel & Kjær Type 4138). The recorded voltage was amplified with a pre-amplifier and power amplifier pair (Brüel & Kjær 2670A and Nexus 2690). The microphone possessed a flat frequency response from 6 to 100 kHz. In the measurements, the output intensity of the ultrasound was adjusted to be lower than its maximum in order to prevent the focus from exceeding the maximum input range of the microphone and causing waveform saturation. We drove a four-unit AUTD system in all of the experiments.

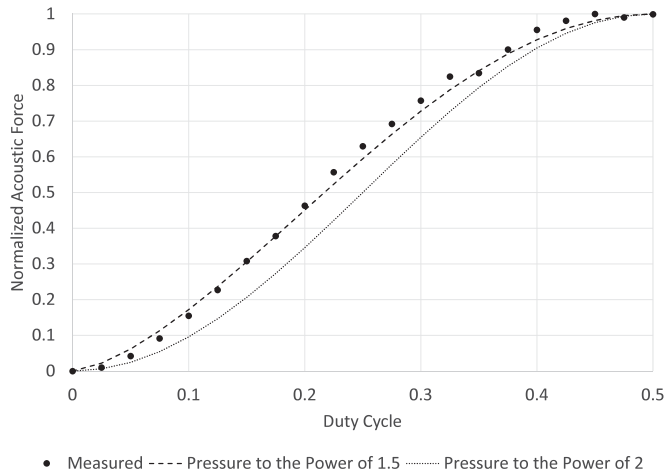


Fig. 6. Normalized radiation pressure against the duty cycle of the driving signal of the transducers. The dotted lines indicate the results of the theoretical calculations.

5.1 Waveform Characteristics

We first confirmed the actual relation between the duty cycle $d(t)$ and the resulting radiation pressure $P_R(t)$. Namely, we experimentally investigated whether (7) holds. We placed a four-unit ultrasound transducer array system on the ceiling of a metal framework, such that its emission plane faced downward. We placed an electronic scale under the arrays. We attached a force receiving jig on the table of the scale so that it could transmit radiation pressure on its receiving surface of 12 mm × 12 mm to the scale. The size of the receiving surface was almost equal to that of the ultrasound focus.

Fig. 6 shows the normalized radiation pressure plotted against the duty cycle. It is seen that the measured radiation pressure is better understood by the model $P_R \propto p^{1.5}$ compared to $P_R \propto p^2$. We consider that this incongruence with the prevalent physical model (1) is due to the characteristics of the employed transducers and their driving circuits. Fig. 7 shows the measured waveforms designed with the criteria $P_R \propto p^1, p^{1.5}$, and p^2 . It shows that the resulting saw wave had the most straight increasing slope when the criteria was $P_R \propto p^{1.5}$. To derive the radiation pressure waveforms, we applied a 5th-order infinite impulse response filter to the

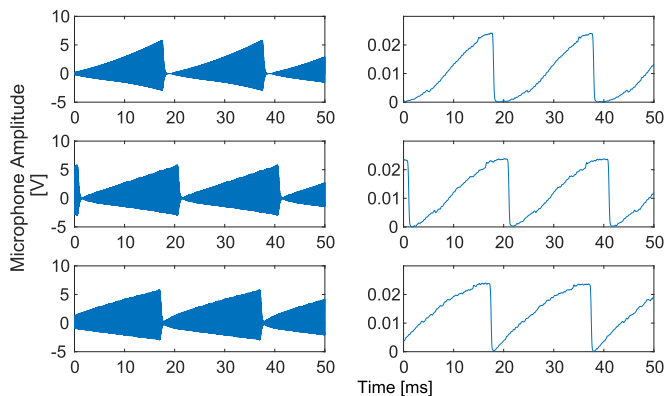


Fig. 7. Observed waveforms of instantaneous acoustic pressure generated with the waveform designed based on the assumption of $P_R \propto p^1$ (upper left), $P_R \propto p^{1.5}$ (middle left), and $P_R \propto p^2$ (lower left), and that of the radiation pressure generated with the waveform designed based on the assumption of $P_R \propto p^1$ (upper right), $P_R \propto p^{1.5}$ (middle right), and $P_R \propto p^2$ (lower right).

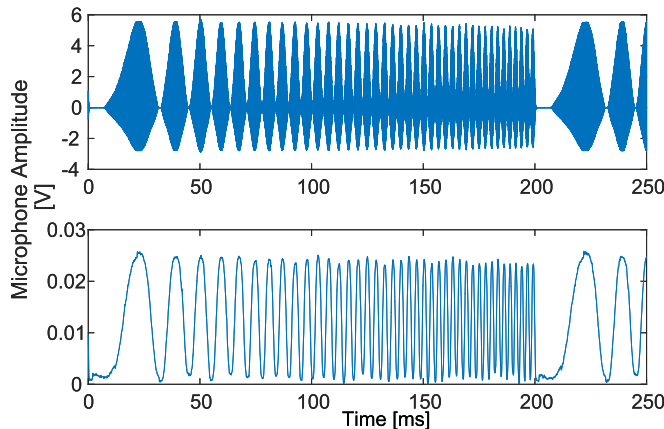


Fig. 8. Observed acoustic pressure (upper) and corresponding radiation pressure (lower) from a designed chirp signal.

observed amplitude waveforms. The filter coefficients were: $b = [0.0749 \ 0.3747 \ 0.7494 \ 0.7494 \ 0.3747 \ 0.0749] \times 10^{-14}$, $a = [1.0000 \ -4.9939 \ 9.9756 \ -9.9635 \ 4.9757 \ -0.9939]$ when the sampling frequency was 10 MHz (Fig. 7) and $b = [0.0229 \ 0.1144 \ 0.2289 \ 0.2289 \ 0.1144 \ 0.0229] \times 10^{-10}$, $a = [1.0000 \ -4.9695 \ 9.8785 \ -9.8184 \ 4.8794 \ -0.9700]$ for 2 MHz (Fig. 8). a and b correspond to the polynomial coefficient series of the filtering transfer function in its denominator and numerator, respectively. Note that the DC component cannot be measured with microphones. Therefore, we added an offset value to the derived radiation pressure waveforms depicted in Fig. 7, such that the minimum value corresponds to zero.

Based on these results, the following measurements and experiments were conducted with radiation pressure outputs whose waveforms were designed with the criteria of assuming $P_R \propto p^{1.5}$. We used the experimental relationship here.

Next, we investigated the frequency characteristics of the acoustic pressure at the focus. Fig. 8 shows the observation of a chirp sinusoidal envelope. The peak-to-peak amplitude of the radiation pressure is seen to be constant in the chirp signal waveform. The frequency characteristics of the root mean square (RMS) acoustic pressure obtained from the sinusoidal modulation are depicted in Fig. 9. In this study,

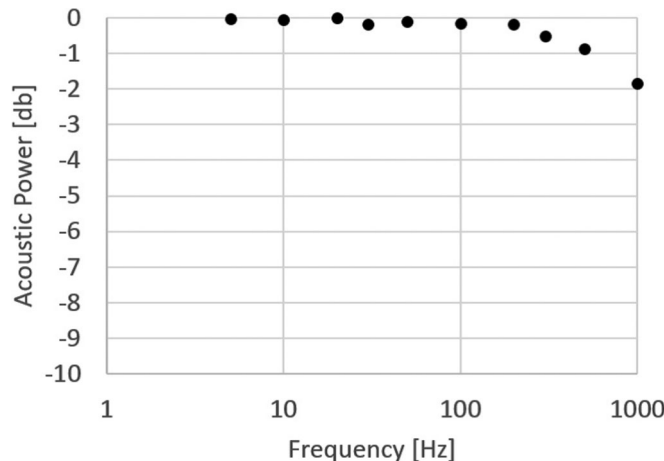


Fig. 9. Frequency characteristics of the acoustic pressure plotted against the sinusoidal modulation frequencies.

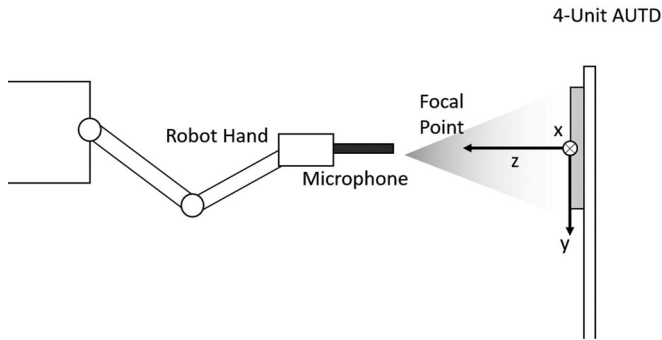


Fig. 10. Experimental setting of the pressure distribution measurements.

the definition of the acoustic pressure and radiation pressure in dB is given as $20 \log_{10} p + C$, where p [Pa] is the RMS value and C is a constant offset. It was experimentally shown that the differences in output amplitude were less than 2 dB in the frequency range of 5 to 1,000 Hz. With respect to the result of generating saw waves, it can be safely said that the focused ultrasound serves as a remote vibrotactile actuator that covers the primary frequency range required for tactile stimuli.

5.2 Spatial Distribution

We measured the spatial distribution of acoustic pressure around the focal point. Fig. 10 shows the experimental setup. We mounted a four-unit AUTD system on a metal framework. We also mounted a standard microphone on the tip of a robot hand (FANUC M710-iC 20L) that performed spatial scanning. The xy -coordinates corresponded with the lattice of transducers and the z -axis was

perpendicular to the emission surface. The array size was $364 \text{ mm} \times 302.8 \text{ mm}$. We performed the measurements under two different conditions: (1) activating all four units and (2) activating only unit. The origin of the xy -plane in the case of the single-unit measurements was shifted to the focal center. Under both conditions, the focal depth was set to $z = 100, 200, 400, 600, 800,$ and $1,000 \text{ mm}$, and was centered with respect to the array.

Fig. 11 depicts the RMS value of the observed acoustic pressure distribution in the xy -plane at $z = 400 \text{ mm}$. As expected, the widened array produced an acoustic pressure field that was more intense and spatially concentrated. Fig. 11 also shows cross-sectional amplitude distributions around the focal region with varied focal depths. It is seen that by enlarging the aperture, the focal blurring effect can be suppressed within a longer depth range. These results strengthen the argument that widening aperture with multi-unit array units can guarantee a widened workspace.

6 EXPERIMENTS ON DETECTION THRESHOLD

As we emphasized previously, another important contribution of the system is that it can present physical vibration of arbitrarily designed waveforms on human skin.

6.1 Method

We conducted a subjective experiment to psychophysically evaluate the fabricated system. We generated sinusoidally modulated focus onto the palms of 12 experimental participants (aged from 22 to 32, including both males and females), and determined the detection threshold curve among several frequencies (5, 10, 20, 30, 50, 100, 200, 300,

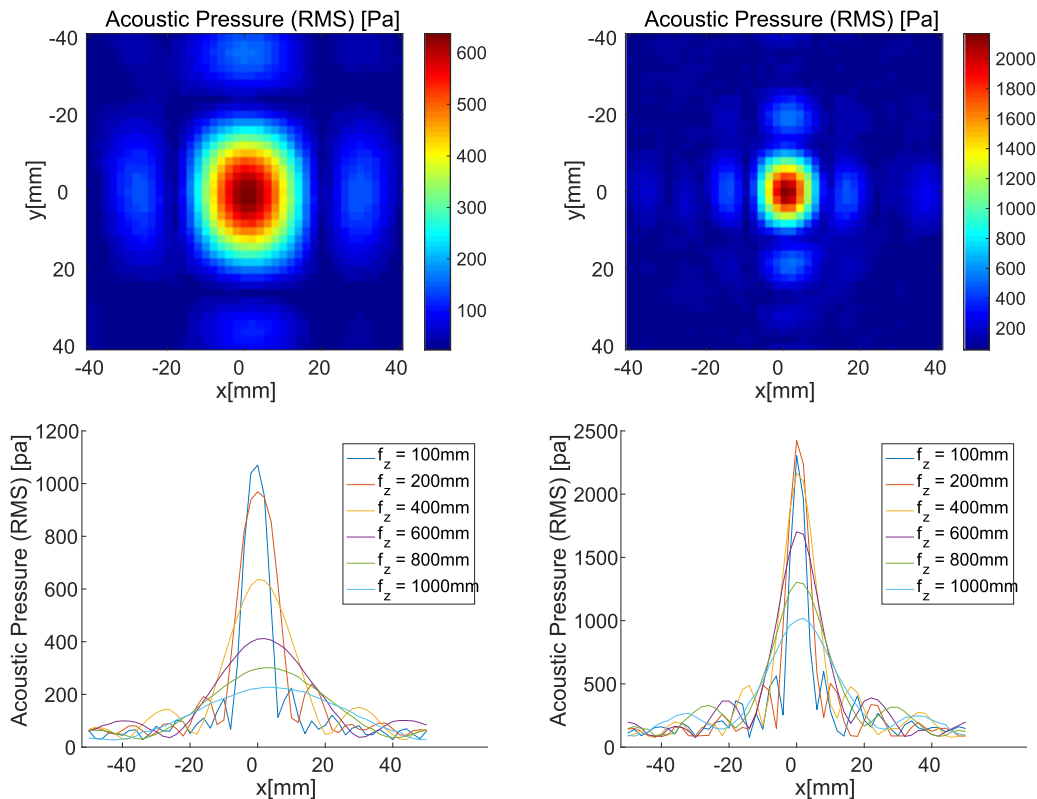


Fig. 11. Measured acoustic pressure field generated by one unit (left) and four units (right).

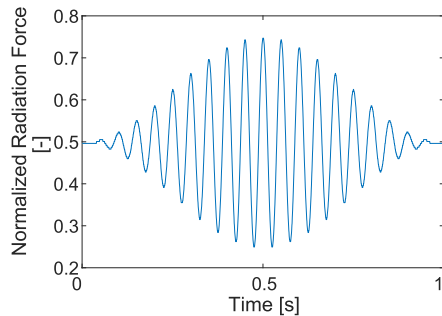


Fig. 12. Example of a sinusoidal wave packet presented to the participants. It is processed by the Hann window with an added DC offset.

500, and 1,000 Hz). Fig. 12 shows a waveform example of radiation force exerted on the participants' palms. Here, radiation force refers to the surface integral of the radiation pressure. Because the current system was not able to exert negative radiation force, we added a constant DC offset to the sinusoidal vibration, whose magnitude was a half of the maximum output. Attention should be paid to the beginning of the stimulation. When the sinusoidal vibration was abruptly displayed to the participants, it would cause discontinuous changes in the displayed force, which are sensitively detected. To remove this discontinuity in the waveform, we multiplied its sinusoidal part by the Hann window, such that its DC part and the vibrational part could be smoothly jointed. The Hann window had a period of 1 s. To sum up the above processes, the waveform $F(t)$ [N], where t [s] denotes time, is represented as

$$F(t) = \beta F_M h(t) \cos\{2\pi f(t - t_0/2)\} + \frac{1}{2} F_M, \quad (8)$$

$$h(t) = \frac{1}{2} - \frac{1}{2} \cos(2\pi t/t_0), \quad (9)$$

where $0 < t < t_0$, $t_0 = 1$ s, F_M [N] is the maximum radiation force, $\beta \in [0, 1]$ is the amplitude scaling factor, and f [Hz] is the central frequency of the sinusoidal vibration. The factor $h(t)$ is the Hann window. $F(t)$ reaches its highest value at $t = t_0/2$ and its temporal average is $F_M/2$, regardless of β when f is the frequency of displayed vibration. In the experiment, we varied the value of β in presenting vibration to the participants' palms. The focal depth was set to approximately 200 mm. As we previously mentioned, we verified that the generated radiation force around the focal region was approximately 20 mN with the maximum burst output of the system. Hence, the maximum instantaneous radiation force was 20 mN when $\beta = 1$ and the DC offset was 10 mN. In the experiment, the whole waveform was displayed twice in a row for every trial.

Participants sat in front of the framework on which an AUTD unit was mounted and placed one of their hands under the device, with the palm facing upward. They wore a headset playing pink noises masking audible noises generated by the ultrasound. Their hands were visually blocked by a curtain. These were for the sake of excluding nontactile stimuli from the participants' perception. They were instructed not to move their hands during the experiments.

In the experiment, we adopted the method of limit with three consecutive ascending and descending processes. A set of 200 linearly leveled stimuli was prepared (i.e.,

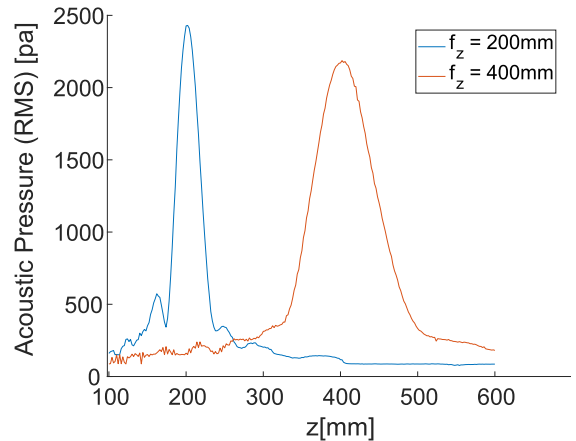


Fig. 13. Acoustic amplitudes along the depth axis at the focal center.

$\beta = 0.005, 0.01, \dots, 1$). We changed the value of β by 0.025 when $\beta > 0.025$ and by 0.005 when $\beta \leq 0.025$. Prior to every presented stimulation, the participants were auditorily informed that it was going to be presented. These alerts were given to the participants as a chime heard from the headset 0.5 s before the stimuli. The participants informed the experimenter as to whether or not they felt the vibrational stimuli by pressing a button they held. Because the focal intensity was dependent on the depth (Fig. 13), the participants had individually adjusted the focal depth before the experiments started.

6.2 Results

Fig. 14 shows the results. In the experiment, some of the participants did not respond to stimuli of maximum intensity at 5 Hz (four out of 12 participants did not respond) and 1,000 Hz (one out of 12 participants did not respond). In the literature, the minimum detectable displacement threshold of vibrotactile stimuli is approximately 200 Hz, where Pacinian corpuscles dominantly react to the exerted vibration [26]. In our experiments, the obtained threshold curve is shown to have a similar tendency along the frequency axis. The minimum radiation pressure threshold is approximately 200 Hz, where $\beta = 0.02$, resulting in a radiation force amplitude of 0.4 mN.

We also investigated the detection threshold of rectangular vibration with the same DC offset at the frequencies of 10, 50, and 100 Hz (Fig. 14). These stimuli were not processed with the Hann window. It is seen that the obtained thresholds are lower than for sinusoidal stimuli, especially at low frequencies. This is because of the harmonics contained within the rectangular pulses. The differences in threshold level indicate that our system was able to present Hann-windowed sinusoidal vibrations and rectangular pulses as distinctly different stimuli at the level of human perception.

The intrinsic difference between the experimental situations of the conventional experiments using factors and ours is whether the threshold is measured by displacement or pressure. These two physical quantities are different but linearly correlated in the range of elastic deformation. The present developed system scientifically clarified that the detection threshold of pressure is minimized at 200 Hz and the average pressure threshold to rectangular waves is

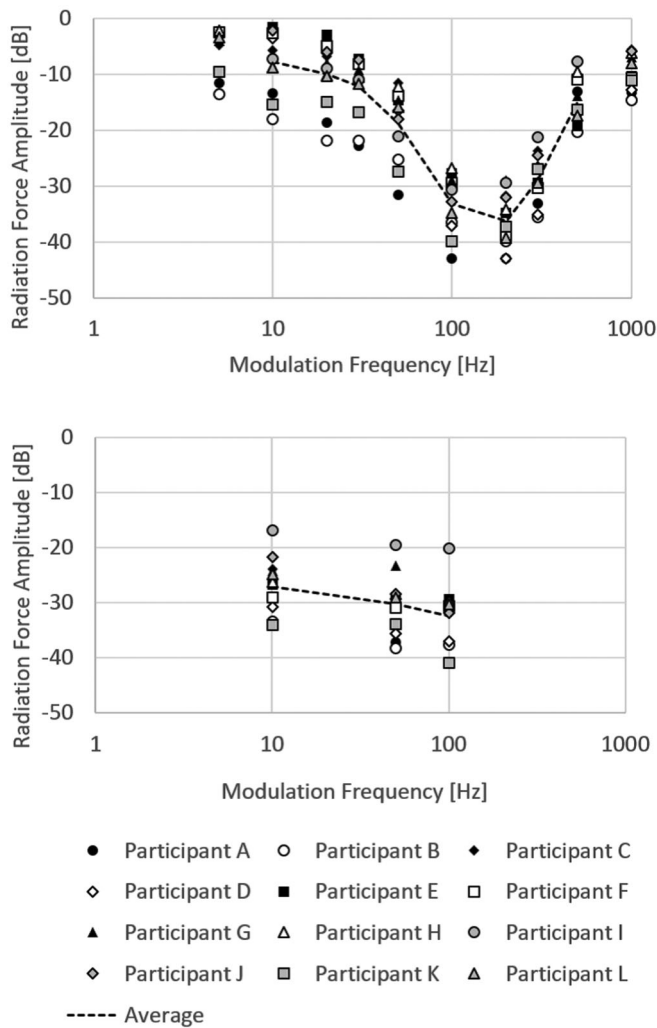


Fig. 14. Psychophysically determined detection threshold of the vibrotactile stimuli generated from focused acoustic radiation pressure for (upper figure) Hann-windowed sinusoidal vibration and (lower figure) rectangular vibration. 0 dB corresponds to the maximum instantaneous radiation force amplitude of 2 gf.

much lower than to sinusoidal waves in the low-frequency range, by approximately 20 dB at 10 Hz.

The individual difference ranged from 15 to 20 dB. Conceivable reasons for this include the difference of the finger shape, stiffness, and age of the participants [27].

7 DISCUSSION

The obtained spatial distribution of the acoustic radiation pressure demonstrates the focus-blurring effect magnified with smaller apertures, although these results seem incongruent with similar measurements conducted previously [28]. As mentioned above, the focal size is theoretically dependent on the aperture size. At least it can be claimed that the radiation pressure at the center of the focus increases as the number of the transducer increases.

The effects of combining multiple array units were experimentally demonstrated: one is the widened workspaces and the other is the broad frequency range of perceivable stimuli. In the case of the four-unit system, the maximum pressure at the focal center is halved when the focal depth was set to 800 mm. This “halved-amplitude depth” for the

case of a single unit was approximately 400 mm. The focal acoustic pressure with the focal depth of 1,000 mm was approximately 40 percent of that obtained with the focal depth of 200 mm. By squaring the ratio, the focal radiation pressure with the depth of 1,000 mm is estimated to be 16 percent of that with the depth of 200 mm (−15.9 dB). Because the vibrotactile detection threshold reaches its minimum (−35 dB) around 200 Hz with the focal depth of approximately 200 mm, it is expected that 200 Hz sinusoidal stimuli presented by the four-unit system are perceivable at a distance of 1,000 mm from the device.

As a component of human—computer interaction system, the most straightforward use of our technique would be to give a pinpoint vibrotactile feedback to users by the ultrasound arrays embedded in a large instrument. Owing to the technical principle of the system that uses converging ultrasound energy, our technique is difficult to implement in a miniature system that offers sufficiently perceivable stimuli. Therefore, our technique is more suitable for infrastructure systems with an unspecified number of users rather than handheld or wearable devices that are intended to be used by individuals.

Among all the ultrasound tactile technologies, the novelty of our work is the temporal evaluation of the time-variant ultrasound focus as a vibrotactile actuator. Because tactile sensation has degrees of freedom in both the temporal and spatial domains, a number of touch sensation are expected to be realized by properly designed spatiotemporal patterns of stimulation. Although this study simply focuses on a single-spot vibration, volumetric acoustic fields with spatially variant vibration patterns can be generated in principle. It should be an interesting research topic to investigate the types of sensation such stimuli can offer.

Prior to the experiments described above, we had preliminarily conducted a similar frequency perception experiment with only a single AUTD unit, in which several subjects claimed that they did not feel the vibrational sensation at frequencies lower than 60 Hz. The current four-unit system enables these low frequencies to be perceivable. This broadens the range of presentable sensation by the system.

It has been demonstrated that human skin and mechanoreceptors exhibit a self-demodulating effect of vibration applied to the skin surface [26], [29]. This research was conducted with a carrier frequency around several kilohertz. To the best of our knowledge there have been no examples work demonstrating the same effect with 40 kHz carrier sinusoidal waves. There is currently no definitive basis for determining whether the sensations felt by the subjects were purely the time-varying radiation pressure or self-demodulated 40 kHz-centered waveforms, as found in the above literature. We expect that the main factor is the radiation pressure, because the perceived stimulation feels comparable even in the case in which a light-weight ultrasound-reflective object is placed on the skin. The confirmation of this expectation remains for future work.

From the viewpoint of tactile texture, clarifying the type of tactile experience a modulated ultrasound focus can generate and how realistic they can be is our next primary research challenge.

8 CONCLUSION

In this paper, we described a multiunit system of ultrasound phased arrays for generating mid-air vibrotactile sensation whose waveform is arbitrarily designed. We verified that synchronized units composed a widened emission aperture, resulting in increased converged radiation pressure, which leads to a widened workspace. We described a PWM-based method to generate a desired waveform of radiation pressure from width-modulated rectangular pulse trains. The generated waveforms were measured via a standard microphone, which showed fairly good frequency and phase characteristics. We determined the detection threshold of the vibrotactile stimuli produced by focused ultrasound. The results showed that the threshold differed between sinusoidal and rectangular vibration. The obtained threshold curve for sinusoidal stimuli showed its lowest value around 200 Hz, as in the literature reporting similar experiments using contact vibrators.

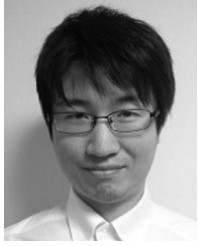
The paper is mainly dedicated to describing the system operation along with its physical effects on human skin. Therefore, more detailed psychophysical and physiological investigations on the tactile experiences offered by the system will be our next focus.

ACKNOWLEDGMENTS

This research is in part financially supported by JSPS Kakenhi 12J09694 and 16H06303. The authors thank Mr. Masafumi Takahashi for the early-stage prototyping of the system.

REFERENCES

- [1] Y. Suzuki and M. Kobayashi, "Air jet driven force feedback in virtual reality," *IEEE Comput. Graph. Appl.*, vol. 25, no. 1, pp. 44–47, Jan./Feb. 2005.
- [2] R. Sodhi, I. Poupyrev, M. Glisson, and A. Israr, "AIREAL: Interactive tactile experiences in free air," *ACM Trans. Graph.*, vol. 32, no. 4, 2013, Art. no. 134.
- [3] H. Lee, et al., "Mid-air tactile stimulation using laser-induced thermoelastic effects: The first study for indirect radiation," in *Proc. IEEE World Haptics Conf.*, 2015, pp. 374–380.
- [4] Y. Ochiai, K. Kumagai, T. Hoshi, J. Rekimoto, S. Hasegawa, and Y. Hayasaki, "Fairy lights in femtoseconds: Aerial and volumetric graphics rendered by focused femtosecond laser combined with computational holographic fields," in *Proc. ACM SIGGRAPH Emerging Technol.*, Aug. 2015, Art. no. 10.
- [5] T. Iwamoto, M. Tatezono, and H. Shinoda, "Non-contact method for producing tactile sensation using airborne ultrasound," in *Proc. Int. Conf. Human Haptic Sens. Touch Enabled Comput. Appl.*, 2008, pp. 504–513.
- [6] J. Awatani, "Studies on acoustic radiation pressure. I (general considerations)," *J. Acoustical Soc. America*, vol. 27, pp. 278–281, 1955.
- [7] T. Hoshi, M. Takahashi, K. Nakatsuma, and H. Shinoda, "Touchable holography," in *Proc. ACM SIGGRAPH Emerging Technol.*, 2009, Art. no. 23.
- [8] Y. Monnai, K. Hasegawa, M. Fujiwara, S. Inoue, and H. Shinoda, "HaptoMime: Mid-air haptic interactions with a floating virtual screen," in *Proc. ACM Annu. ACM Symp. User Interface Softw. Technol.*, 2014, pp. 663–667.
- [9] 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," in *Proc. ACM CHI Conf. Human Factors Comput. Syst.*, 2016, pp. 1980–1990.
- [10] S. Okamoto, M. Konyo, S. Saga, and S. Tadokoro, "Detectability and perceptual consequences of delayed feedback in a vibrotactile texture display," *IEEE Trans. Haptics*, vol. 2, no. 2, pp. 73–84, Apr.–Jun. 2009.
- [11] K. Hasegawa and H. Shinoda, "A method for distribution control of aerial ultrasound radiation pressure for remote vibrotactile display," in *Proc. SICE Annu. Conf.*, 2013, pp. 223–228.
- [12] L. R. Gavrilov, "The possibility of generating focal regions of complex configuration in application to the problems of stimulation of human receptor structures by focused ultrasound," *Acoustical Physics*, vol. 54, no. 2, pp. 269–278, 2008.
- [13] S. Inoue, Y. Makino, and H. Shinoda, "Active touch perception produced by airborne ultrasonic haptic hologram," in *Proc. IEEE World Haptics Conf.*, 2015, pp. 362–367.
- [14] T. Carter, S. A. Seah, B. Long, B. Drinkwater, and S. Subramanian, "UltraHaptics: Multi-point mid-air haptic feedback for touch surfaces," in *Proc. Annu. ACM Symp. User Interface Softw. Technol.*, 2013, pp. 505–514.
- [15] B. Long, S. A. Seah, T. Carter, and S. Subramanian, "Rendering volumetric haptic shapes in mid-air using ultrasound," *ACM Trans. Graph.*, vol. 33, no. 6, 2014, Art. no. 181.
- [16] S. Inoue, K. J. Kobayashi-Kirschvink, Y. Monnai, K. Hasegawa, Y. Makino, and H. Shinoda, "HORN the hapt-optic reconstruction," in *Proc. ACM SIGGRAPH Emerging Technol.*, 2014, Art. no. 11.
- [17] K. Hasegawa and H. Shinoda, "Aerial display of vibrotactile sensation with high spatial-temporal resolution using large-aperture airborne ultrasound phased array," in *Proc. IEEE World Haptics Conf.*, 2013, pp. 31–36.
- [18] H. Y. Yao and V. Hayward, "An experiment on length perception with a virtual rolling stone," in *Proc. Int. Conf. Human Haptic Sens. Touch Enabled Comput. Appl.*, 2006, pp. 275–278.
- [19] K. Minamizawa, M. Nakatani, Y. Kakehi, S. Mihara, and S. Tachi, "TECHTILE toolkit: A prototyping tool for designing haptic media," in *ACM SIGGRAPH 2012 Emerging Technol.*, Art. no. 22, New York, NY, USA, 2012, Art. no. 22.
- [20] S. Tsuchiya, M. Konyo, H. Yamada, T. Yamauchi, S. Okamoto, and S. Tadokoro, "VibTouch: Virtual active touch interface for hand-held devices," in *Proc. IEEE Int. Symp. Robot Human Interactive Commun.*, 2009, pp. 12–17.
- [21] J. M. Romano and K. J. Kuchenbecker, "Creating realistic virtual textures from contact acceleration data," *IEEE Trans. Haptics*, vol. 5, no. 2, pp. 109–119, Apr.–Jun. 2012.
- [22] M. Takahashi and H. Shinoda, "Large aperture airborne ultrasound tactile display using distributed array units," in *Proc. SICE Annu. Conf.*, 2010, pp. 359–362.
- [23] G. Korres and M. Eid, "Haptogram: Ultrasonic point-cloud tactile stimulation," *IEEE Access*, vol. 4, pp. 7758–7769, 2016.
- [24] T. Hoshi, M. Takahashi, T. Iwamoto, and H. Shinoda, "Noncontact tactile display based on radiation pressure of airborne ultrasound," *IEEE Trans. Haptics*, vol. 3, no. 3, pp. 155–165, Jul.–Sep. 2010.
- [25] K. Yoshino, K. Hasegawa, and H. Shinoda, "Measuring visio-tactile threshold for visio-tactile projector," in *Proc. SICE Annu. Conf.*, 2012, pp. 1996–2000.
- [26] P. J. J. Lamoreet, H. Muijser, and C. J. Keemink, "Envelope detection of amplitude-modulated high-frequency sinusoidal signals by skin mechanoreceptors," *J. Acoustical Soc. America*, vol. 79, pp. 1082–1085, 1986.
- [27] E. Perret and F. Regli, "Age and the perceptual threshold for vibratory stimuli," *Eur. Neurology*, vol. 4, pp. 65–76, 1970.
- [28] G. Korres and M. Eid, "Characterization of ultrasound tactile display," in *Proc. Int. Conf. Human Haptic Sens. Touch Enabled Comput. Appl.*, 2016, pp. 78–89.
- [29] Y. Makino, T. Maeno, and H. Shinoda, "Perceptual characteristic of multi-spectral vibrations beyond the human perceivable frequency range," in *Proc. IEEE World Haptics Conf.*, 2011, pp. 439–443.



Keisuke Hasegawa received the BE degree in information physics and the ME and PhD degrees in information science and technology from the University of Tokyo, Japan, in 2009, 2011, and 2014, respectively. He is a research associate in the Graduate School of Frontier Sciences, University of Tokyo, Japan. He was a research fellow with the Japan Society for the Promotion of Science (JSPS) from 2012 to 2014. He received the IEEE WHC Best Student Paper Award and the SICE Young Author's Award in

2013. His current research interests include haptic technologies to support our daily lives. He is a member of the IEEE.



Hiroyuki Shinoda received the BS degree in applied physics, the MS degree in information physics, and the PhD degree in engineering from the University of Tokyo, in 1988, 1990, and 1995, respectively. He is a professor in the Graduate School of Frontier Sciences, University of Tokyo, Japan. He was an associate professor from 1995 in the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. After a period at UC Berkeley as a visiting scholar in 1999, he was an associate professor with the University of Tokyo from 2000 to 2012. His research interests include information physics, tactile/haptic interfaces, haptic sensors and displays, two-dimensional communication, and electromagnetic/acoustic measurements. He was a board member of SICE in 2008 and 2009, and is currently a VRSJ board member and a member of the IEEJ, RSJ, JSME, and IEEE.