

Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound

Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, and Hiroyuki Shinoda

Abstract—This paper describes a tactile display which provides unrestricted tactile feedback in air without any mechanical contact. It controls ultrasound and produces a stress field in a 3D space. The principle is based on a nonlinear phenomenon of ultrasound: Acoustic radiation pressure. The fabricated prototype consists of 324 airborne ultrasound transducers, and the phase and intensity of each transducer are controlled individually to generate a focal point. The DC output force at the focal point is 16 mN and the diameter of the focal point is 20 mm. The prototype produces vibrations up to 1 kHz. An interaction system including the prototype is also introduced, which enables users to see and touch virtual objects.

Index Terms—Emerging technologies, haptic I/O, tactile displays, 3D displays, virtual reality.

1 INTRODUCTION

THERE has been wide interest in midair displays and such displays have been seen in many SF movies [1]. People in the movies manually interact with touch screens or virtual objects floating in front of them. Recently, glasses-based 3DTVs are ready to be launched onto the market, and meanwhile novel technologies are being developed to render images hovering in air without special glasses. FogScreen [2] and Heliodisplay [3] use a thin layer of fog as a projection screen. HoloVision [4] provides floating images from an LCD by utilizing a concave mirror. Integral photography reproduces 3D images through a lens array [5]. SeeReal Technologies is working on a real-time, computer-generated 3D holography [6] through the use of an eye-tracking technology for reduction of calculation amount. Furthermore, by applying the camera-based and markerless hand tracking techniques demonstrated in Holovizio [7] or GrImage [8], we will be able to handle the projected images with our hands. In that situation, tactile feedback will be the next demand. If tactile feedback is provided additionally, the usability of the interaction systems will be highly improved.

Requirements for a tactile display to be combined with floating images are: 1) avoiding the interference that the tactile display hides the visual images and 2) allowing users to move their arms, hands, and fingers freely. It is most preferable that users are not required to touch on or hold any devices, unlike general tactile displays [9]. For example, we expect the situation that a user explores with his/her

hands within a workspace, as illustrated in Fig. 1. Each side length of the cubic space is several hundred millimeters (e.g., 300 mm in Fig. 1). The issue is to provide tactile feedback anywhere within that 3D free space.

There are three types of conventional strategies for tactile feedback in free space. The first is attaching tactile devices on user's fingers and/or palms. Employed devices are, for example, vibrotactile stimulators (CyberTouch [10]), motor-driven belts (GhostGlove [11]), or pin-array units (SaLT [12]). In this strategy, the skin and the device are always in contact and that leads to undesired touch feelings. The second is controlling the positions of tactile devices so that they contact with the skin only when tactile feedback is required. In the master-slave system, shown in [13], the encounter-type force feedback is realized by the exoskeleton master hand. The detailed tactile feedback for each finger is provided by the electrotactile display attached on the finger part of the master hand. The drawback of this strategy is that it requires bulky robot arms. The last is providing tactile feedback from a distance without any direct contact. For example, air jets are utilized in [14] to realize noncontact force feedback. Fans or air cannons are used in theme parks to amaze visitors. Although they are effective for rough "force" feedback, their spatial and temporal properties are quite limited and they cannot provide detailed "tactile" feedback. The reachable distance of an air jet is determined by the diameter and the velocity of the jet stream, which results in the trade-off between the spatial resolution of the pressure on the skin and the distance from the device to the skin.

Our proposal is to use ultrasound for tactile stimulation. Under the limitation that the maximal generated force is small (e.g., 16 mN in the current prototype), people are free from any bothering tactile devices. It is unnecessary for users to wear tactile devices or to touch some special devices installed in a room. We can stimulate arbitrary parts of the user with short time delays. It is even possible to give tactile stimulation to the people moving freely in a room if the tactile display is placed on the ceiling, for example.

The usage of airborne ultrasound for a tactile display was first introduced in [15] and the feasibility was examined using the prototype consisting of 91 ultrasound transducers

• T. Hoshi is with the Graduate School of Science and Technology, Kumamoto University, Room 613, Research Bldg. of Faculty of Engineering I, 2-39-1, Kurokami, Kumamoto-shi, Kumamoto, Japan. E-mail: star@alab.t.u-tokyo.ac.jp.

• M. Takahashi, T. Iwamoto, and H. Shinoda are with the Graduate School of Information Science and Technology, The University of Tokyo, Room 133, Eng. Bldg. 6, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan. E-mail: {masafumi, iwa, shino}@alab.t.u-tokyo.ac.jp.

Manuscript received 31 Aug. 2009; revised 9 Dec. 2009; accepted 12 Jan. 2010; published online 29 Jan. 2010.

Recommended for acceptance by L. Jones, M. Harders, and Y. Yokokohji. For information on obtaining reprints of this article, please send e-mail to: toh@computer.org, and reference IEEECS Log Number THSI-2009-08-0066. Digital Object Identifier no. 10.1109/ToH.2010.4.

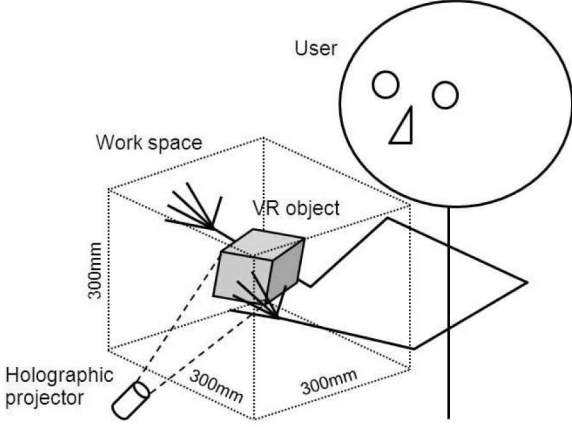


Fig. 1. Target situation. Some kind of tactile feedback is desired when a user touches a floating VR object with hands.

that generates a fixed focal point. In this paper, we add movements of the focal point and interactions with humans using an upgraded version consisting of 324 transducers (Fig. 2) [16].

The following paper outlines, first, the principles of the ultrasound-based tactile display. Second, the structure and performance of the current prototype are described in Section 3, and the experiments and results are presented in Section 4. Third, a developed interaction system is shown in Section 5, which is a combination of our tactile display and an aerial imaging display. After some issues are discussed in Section 6, finally, Section 7 concludes this paper.

2 PRINCIPLES

2.1 Acoustic Radiation Pressure

Our tactile stimulation is based on a nonlinear phenomenon of ultrasound: Acoustic radiation pressure [17], [18]. Assuming a plane wave, the acoustic radiation pressure P [Pa] is described as

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

where E [J/m³] is the energy density of ultrasound, I [W/m²] is the sound intensity, c [m/s] is the sound speed, p [Pa] is the RMS sound pressure of ultrasound, and ρ [kg/m³] is the density of medium. α is the constant ranging from 1 to 2 depending on the amplitude reflection coefficient R at an object surface; $\alpha \equiv 1 + R^2$. If the object surface perfectly reflects the incident ultrasound, $\alpha = 2$, while if it absorbs the entire incident ultrasound, $\alpha = 1$. In case that ultrasound beam is reflected vertically at the object surface, the surface is subjected to the constant vertical force in the direction of the incident beam. Equation (1) suggests that the spatial distribution of the radiation pressure P can be controlled by synthesizing the spatial distribution of the ultrasound p .

Air is a medium of ultrasound in this paper, while water was used as a medium in our previous study [19]. It is useful to notice the following two advantages of the air case. First, since the radiation pressure P is inversely proportional to the sound speed c for a constant sound intensity I , the power theoretically required for a given stimulation

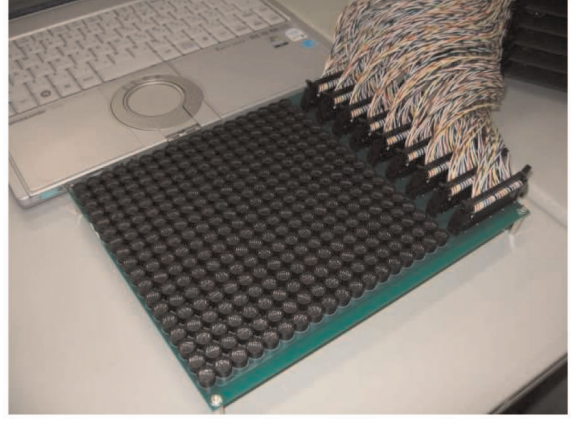


Fig. 2. An 18 × 18 array of airborne ultrasound transducers.

force in air ($c = 340$ m/s) is 4.4 times as small as that in water ($c = 1,500$ m/s). The second difference is related to the reflection coefficient. The characteristic acoustic impedance of skin (soft tissue) Z_s and that of air Z_a are 1.63×10^6 and 0.0004×10^6 N · s/m³, respectively [20]. In this case, the reflection coefficient R is determined to be

$$R = \left| \frac{Z_s - Z_a}{Z_s + Z_a} \right| \approx 0.9995. \quad (2)$$

Since 99.9 percent ($=R^2 \times 100$) of the incident acoustic energy is reflected at the skin surface, airborne ultrasound can be directly applied onto skin with acceptable invasion. In the water medium, an ultrasound reflective film is needed on the skin surface [19] to avoid considerable power absorption by tissue.

The effective range and the spatial resolution of the proposed method are in the relation of trade-off. As the frequency of ultrasound becomes higher, the diameter of the focal point of ultrasound becomes smaller. From the viewpoint of spatial resolution, the smaller diameter is preferable. However, air is a lossy medium and its attenuation coefficient β [Np/m] of a plane sound wave varies according to the frequency. The energy density E at the distance z [m] is described as

$$E = E_0 e^{-2\beta z}, \quad (3)$$

where E_0 is the energy density at the transducer surface (i.e., $z = 0$ mm). Based on [21], it is assumed that the value of the attenuation coefficient β at 40 kHz is 1.15×10^{-1} Np/m (i.e., 100 dB/100m), and β is proportional to the square of the frequency. Fig. 3 shows the relationship between the frequency of ultrasound and the energy loss ratio at $z = 200$ mm. When the frequency is 40 kHz, the energy loss is 4 percent. However, if the frequency becomes four times as large, 50 percent of the emitted acoustic energy is lost. Our prototype utilizes 40 kHz ultrasound because the attenuation is relatively small and 40 kHz ultrasound transducers are commercially available. Then, if we limit the range to 400 mm, the radiation pressure maintains more than 90 percent.

2.2 Phased Array Focusing (Spatial Control)

To produce the radiation pressure perceivable by human skins, we use the Phased Array Focusing technique. The focal point of ultrasound is generated by controlling the

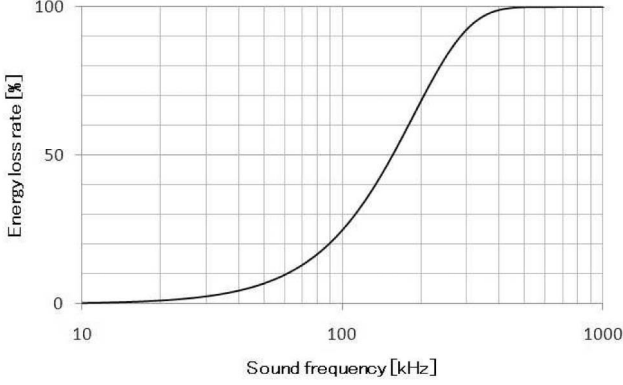


Fig. 3. Relationship between sound frequency and energy loss rate at 200 mm distance.

phase delays of multiple transducers. That technology is widely used in sonography, optics, astronomy, and various radar systems. It is also employed to generate a localized audible sound spot in a 3D space [22]. The radiation pressure at the focal point is proportional to the square of the number of transducers according to (1) when the sound pressure is assumed to be a linear sum of the contributions from the transducers. For example, the RMS sound pressure $p_{0.3}$ at the distance of 300 mm radiated from a single typical transducer (T4010A1 [23], Nippon Ceramic Co. Ltd.) is 121.5 dB SPL (=24 Pa) at the normal-rated voltage (10 Vrms), and the resulting radiation pressure is 8.3×10^{-3} Pa. If we drive 100 transducers so that the phases of all the ultrasounds coincide at a point, the radiation pressure as large as 83 Pa will be generated there.

Here, we theoretically derive the resulting sound pressure field on the focal plane generated by the current prototype consisting of 324 transducers. First, the specifications of the transducer array are listed. The diameters of the transducer housing and the diaphragm are $d = 10$ mm and 8 mm, respectively. The resonant frequency is 40 kHz. While the transducer has directivity (its half-amplitude full angle is 100 deg), we assume a spherical sound wave for simplicity. We use $N \times N$ ($=18 \times 18$) pieces of the transducers and arrange them into a 180×180 mm² rectangle.

Next, the sound pressure field is formulated. We take the coordinate system shown in Fig. 4. Let r [m] be the focal length. The RMS sound pressure p_r from each transducer on the focal plane is in inverse proportion to r . The phase delay is assumed to be adequately determined so that the focal point is generated at (x_c, y_c, r) . Then, assuming that the RMS sound pressure from each transducer is constant p_r anywhere on the focal plane, the resulting sound pressure field $p(x_0, y_0)$ on it is written as

$$p(x_0, y_0) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \sqrt{2} p_r e^{-jk r'} e^{j(k r'' - \omega t)} \quad (4)$$

$$\approx \sqrt{2} p_r N^2 \frac{\text{sinc}\left(\frac{N d \nu_x}{2}, \frac{N d \nu_y}{2}\right)}{\text{sinc}\left(\frac{d \nu_x}{2}, \frac{d \nu_y}{2}\right)} e^{j\{\varphi(x_0, y_0) - \omega t\}},$$

where $r' \equiv \{(x_m - x_c)^2 + (y_n - y_c)^2 + r^2\}^{1/2} \approx r + \{(x_m - x_c)^2 + (y_n - y_c)^2\}/2r$ is the distance from the m th row and n th column transducer to the focal point, and $r'' \equiv \{(x_m - x_0)^2 + (y_n - y_0)^2 + r^2\}^{1/2} \approx r + \{(x_m - x_0)^2 + (y_n - y_0)^2\}/2r$ is the distance from the transducer to the

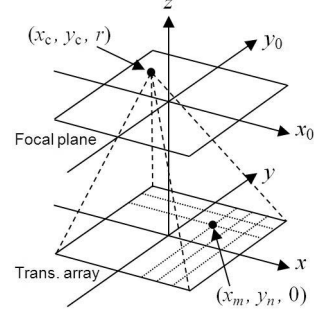


Fig. 4. Coordinate system.

arbitrary point on the focal plane. In going from the first line to the second line of (4), the Fresnel approximation [24] is applied to r' and r'' , whose validity condition is fulfilled when $r \geq 200$ mm in our case. (x_m, y_n) is written as $(md + \xi, nd + \xi)$, where ξ ($= -85$ mm) is an offset. $\exp(-jk r')$ is the phase control factor to focus the ultrasound like a lens [24] and $\sqrt{2} p_r \exp\{j(k r'' - \omega t)\}$ is the spherical wave from each transducer. j is the imaginary unit and k [rad/m] is the wavenumber. The function sinc is defined as $\text{sinc}(x, y) \equiv \sin(x) \sin(y)/xy$. Transformation of variables is done as

$$\nu_x \equiv (k/r)(x_0 - x_c), \quad \nu_y \equiv (k/r)(y_0 - y_c).$$

$$\varphi(x_0, y_0) \equiv (k/2r)(x_0^2 + y_0^2 - x_c^2 - y_c^2)$$

$$- \{\xi + (N-1)d/2\}(\nu_x + \nu_y)$$

is a phase delay.

Equation (4) determines the size of the focal point and the total force. The diameter w_f [m] is given by $4\pi r/kNd$ and the depth w_d [m] is geometrically derived to be $2rw_f/Nd$ (Fig. 5). For example, $w_f = 19$ mm and $w_d = 42$ mm when $r = 200$ mm. Fig. 6 shows the theoretical distribution of radiation pressure normalized so that the maximum value is equal to 1.0. The main lobe is accompanied by the four side lobes whose normalized amplitudes are about 0.05. The total force F [N] within the focal point is given by

$$F = \frac{\alpha}{\rho c^2} \int_{-w_f/2}^{w_f/2} \int_{-w_f/2}^{w_f/2} \frac{|p(x_0, y_0)|^2}{2} dx_0 dy_0 \quad (5)$$

according to (1). For example, $F = 143$ mN when $r = 200$ mm, $p_{0.2} = 36$ Pa, and $\alpha = 2$.

2.3 Modulation (Temporal Control)

From the viewpoint of human tactile properties, the bandwidth up to 1 kHz is sufficient for tactile displays [25]. Our ultrasound-based method has a possibility to produce various kinds of texture feelings because the carrier frequency, 40 kHz, is much higher than the required modulation bandwidth, 1 kHz. A parameter to be concerned with is the actual bandwidth of the transducer which usually utilizes resonance in order to obtain impedance matching between the air and the device. According to the specifications of the transducer we use, the half-power bandwidth is about 2 kHz. It covers the required bandwidth.

The Pulse Width Modulation (PWM) technique is used in our prototype instead of modulating the amplitude of

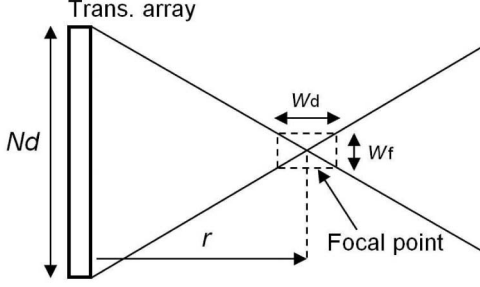


Fig. 5. Simplified diagram of focal point.

driving voltages, for the easiness of electrical circuits. The input voltage $V(t)$ [V] is a rectangular wave written as

$$V(t) = \begin{cases} V_0, & (nT \leq t < nT + D), \\ 0, & (nT + D \leq t < nT + T), \end{cases} \quad (6)$$

where n , D [s], and T [s] are an integer, the pulse width, and the cycle time ($25 \mu\text{s}$ for 40 kHz), respectively. We control the pulse width D . The amplitude of a 40-kHz component a_1 [V] of $V(t)$ is written as

$$a_1 = \frac{2}{\pi} V_0 \left| \sin \pi \frac{D}{T} \right|. \quad (7)$$

Equation (7) means that the 50-percent duty ratio (i.e., $D = T/2$) gives the maximum carrier amplitude. The output sound pressure is proportional to a_1 since the driven device is a resonator, and the resulting acoustic radiation pressure is accordingly proportional to $|\sin(\pi D/T)|^2$.

2.4 Safety of Ultrasound

There are two types of limits of ultrasound intensity regarding the safety for human bodies. One comes from the heat damage to tissue under skin. In the field of ultrasonography, a safe level of tissue ultrasound exposure has been defined as 100 mW/cm^2 [26]. Since 0.1 percent of ultrasound power is absorbed by skin, the ultrasound up to 100 W/cm^2 can be applied onto the skin surface, and that

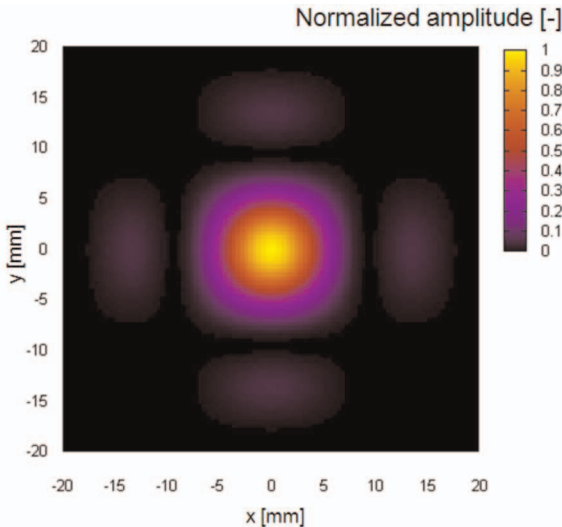


Fig. 6. Simulated distribution of focused radiation pressure.

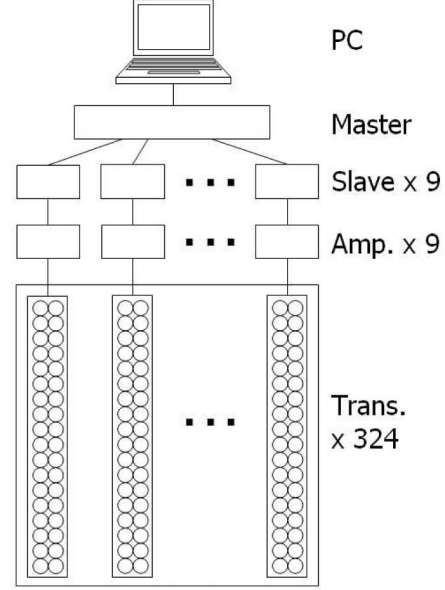


Fig. 7. Block diagram of prototype.

value leads to the total force of 428 mN. That maximum force is strong enough for many applications.

The other comes from the impingement on human ear. The limit recommended in [27] is 110 dB SPL. The sound pressure at the distance of 300 mm from a single transducer is 121.5 dB SPL and it is larger than the recommended limit. Besides, hundreds of transducers are used in the proposed tactile display. Until the safety is confirmed, users should take care not to place their ears near the transducer array or wear headphones to protect their ears as well as to hear stereophonic sounds in a virtual world.

3 PROTOTYPE DEVICE

3.1 Hardware Description

Fig. 2 shows the array of the airborne ultrasound transducers, whose specifications are described in Section 2.2. Fig. 7 shows the block diagram of the system. It consists of a laptop PC with a digital I/O card, a master circuit, nine slave circuits, nine amplifier circuits, and the transducer array. The master circuit has an FPGA and a 25.6-MHz oscillator which acts as the system clock, and each slave circuit also has two FPGAs. The master circuit receives a command or datum from the PC through the digital I/O, and broadcasts it to the slaves. Each slave drives 36 transducers individually through the 36-ch amplifier circuit. The driving signal into the transducer is a 24-Vp-p, 40-kHz rectangular wave whose DC component is cut by an HPF.

The system operates as follows: When the system is turned on, the precalculated lookup table of phase delays and amplitudes are downloaded to the slaves. After that, according to the command (including the target position of the focal point) sent from the PC, the slaves drive the ultrasound transducers at the corresponding phases and amplitudes based on the table.

Here, we describe how to control the phase and amplitude in the prototype. One cycle of 40-kHz rectangular wave is divided into 16 segments (i.e., $1.5625 \mu\text{s}$). The phase is controlled by the position of a HIGH (=24 V)

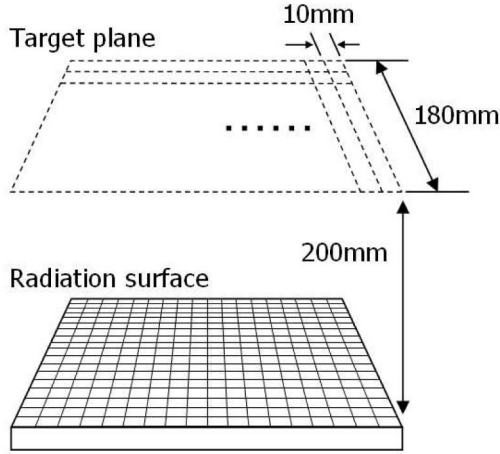


Fig. 8. Description of target plane.

period within the 16 segments, and the amplitude by the duration of the HIGH period. That is, the phase and amplitude are quantized in 4 and 3 bit, respectively.

The objective spatial pattern is a single-peak pattern in this research. To reduce the total data amount, a target area is limited to a $180 \times 180 \text{ mm}^2$ horizontal plane at 200 mm above the radiation surface. The target area is divided into $10 \times 10 \text{ mm}^2$ subareas. The center positions of the subareas are selectable as the focal position (Fig. 8). That is, the focal point can move among 18×18 discrete positions. For each position and each transducer, the corresponding phase is determined based on the distance between them. The amplitudes of all the transducers are tentatively fixed at the maximum value (i.e., the 50 percent duty ratio). Fig. 9 shows a demonstration of the scanning movement of the focal point. Paper strips are flipped up when the focal point comes.

The pressure is controlled not only spatially but also temporally. It can be modulated by a rectangular wave whose frequency is ranging from 1 to 1,000 Hz and the duty ratio is 50 percent. The frequency is quantized in 5 bit so that the selectable values are equally spaced on the logarithmic frequency scale.

The time-averaged power consumption of the whole system is about 70 W (measured); 35 W for the master and slaves, 35 W for the amplifiers and transducers. The latter value is approximately consistent with the value theoretically estimated based on the specifications of the transducer.

3.2 Total Force

The total force was measured using an electronic balance. The transducer array was fixed just above the electronic balance placed on a table. The radiation surface of the array was faced toward the electronic balance. (i.e., the radiation surface was upside down.) The ultrasound was radiated without modulation (i.e., DC). The focal point was fixed at 200 mm above the radiation surface and the distance between the radiation surface and the electronic balance was also fixed at 200 mm. When the amplitude of the input signal was 24 Vp-p, the measured force was 16 mN, which is about 11 percent of the theoretical value predicted in Section 2.2. The possible factors of the difference are:

1. differences of decay according to the positions of the transducers,
2. the directivity of the transducer,

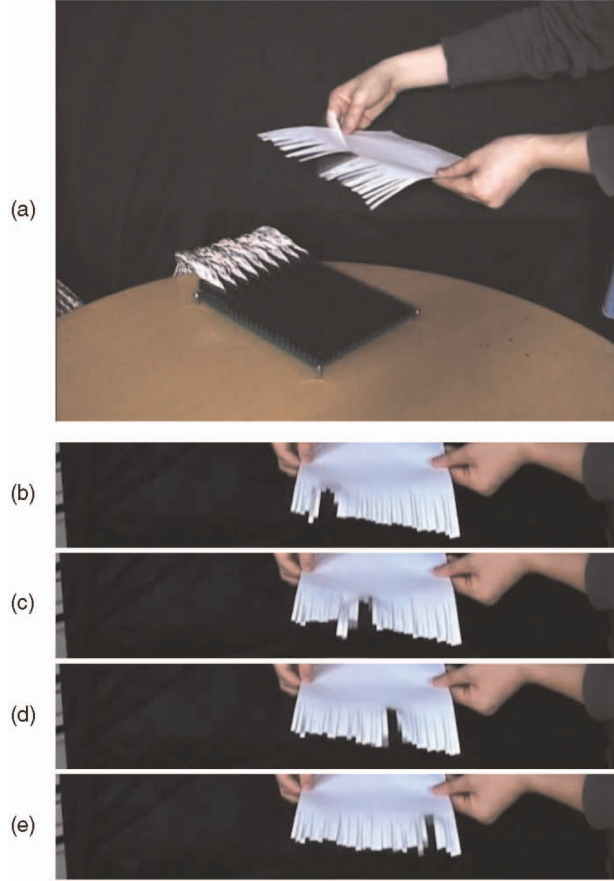


Fig. 9. Left-to-right scanning movement of focal point. (a) is an overview and (b), (c), (d), and (e) are close-ups taken every 0.1s.

3. incident angles of ultrasound waves,
4. differences of input-to-output delays between the individual transducers,
5. discretization errors of the phase delays,
6. a radiation efficiency originating from the mutual radiation impedance [29], and
7. other nonlinear effects of air.

Those factors would also affect the spatial distribution of the acoustic radiation pressure.

3.3 Spatial Distribution

In order to measure the spatial distribution of the acoustic radiation pressure, we used a setup shown in Fig. 10. A pressure sensor probe was attached to an XYZ stage whose resolution was 0.1 mm. The output signal of the sensor was amplified after its DC and high-frequency components were cut by a BPF, that is, the modulated acoustic radiation pressure was emphasized. The aperture of the sensor was 2 mm. The position of the focal point was set to $(x, y, z) = (0 \text{ mm}, 0 \text{ mm}, 200 \text{ mm})$. Data were acquired at every 2, 1, and 5 mm around the focal point for Figs. 11a, 11b, and 11c, respectively. The modulation frequency was 100 Hz, and the amplitude of the 100-Hz component was extracted from the stored data through the FFT. A digital oscilloscope was employed to sample data and to conduct the FFT calculation simultaneously. The amplitude turned out to be time varying, and hence, we

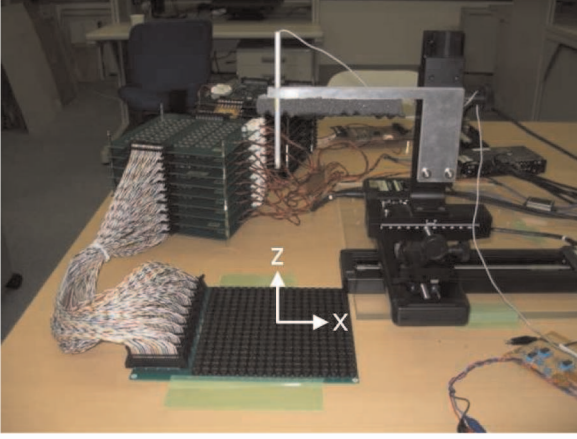


Fig. 10. Experimental setup.

recorded maximum and minimum values during 1-minute observation at each point.

Fig. 11 shows the spatial distribution of the acoustic radiation pressure (100-Hz component) around the focal point; a and b are the results of XY and X scanning on the focal plane, respectively, and c is the result of Z scanning along the central axis. The amplitude is normalized. Fig. 11a shows the maximum value within 1-minute observing time at each point; b and c show the mean values as dots and the maximum and minimum values as both ends of vertical bars. The following are seen in Fig. 11. One focal point is surely generated and its diameter is about 20 mm, which is consistent with the theory described in Section 2.2. There are the side lobes around the main lobe as the theory predicted, but they are not symmetric. The depth of the focal point is about 100 mm if we fix a threshold at 0.2, which is 2.4 times as large as the theoretical value. The radiation pressure around the focal point is unstable due to an air flow (discussed in Section 6.1).

3.4 Temporal Properties

The temporal properties of the prototype were examined with the setup described in Section 3.3. The pressure sensor was placed just at the focal point. The position of the focal point was set to $(x, y, z) = (0 \text{ mm}, 0 \text{ mm}, 200 \text{ mm})$. The driving signal was a 40-kHz rectangular wave modulated by a rectangular wave whose frequency was variable. In observing waveforms shown in Figs. 12a and 12b, the modulation frequency was fixed at 100 Hz. The waveform of ultrasound was measured by a pressure sensor whose frequency response was flat within the audible range and fell to -10 dB at 40 kHz. The waveform of radiation pressure was extracted from the measured signal by cutting its 40 kHz and DC components with an LPF and an HPF, respectively. Then, the changes of amplitude and phase of the radiation pressure waveform caused by the filters were compensated.

Figs. 12a and 12b show the waveforms of the modulated ultrasound and the resulting radiation pressure. The vertical axis represents the amplitude normalized by the maximum value. The horizontal axis represents time. The following are seen in Figs. 12a and 12b. The waveform of the radiation pressure changes according to the envelope of the ultrasound, as expected. While the rising and falling

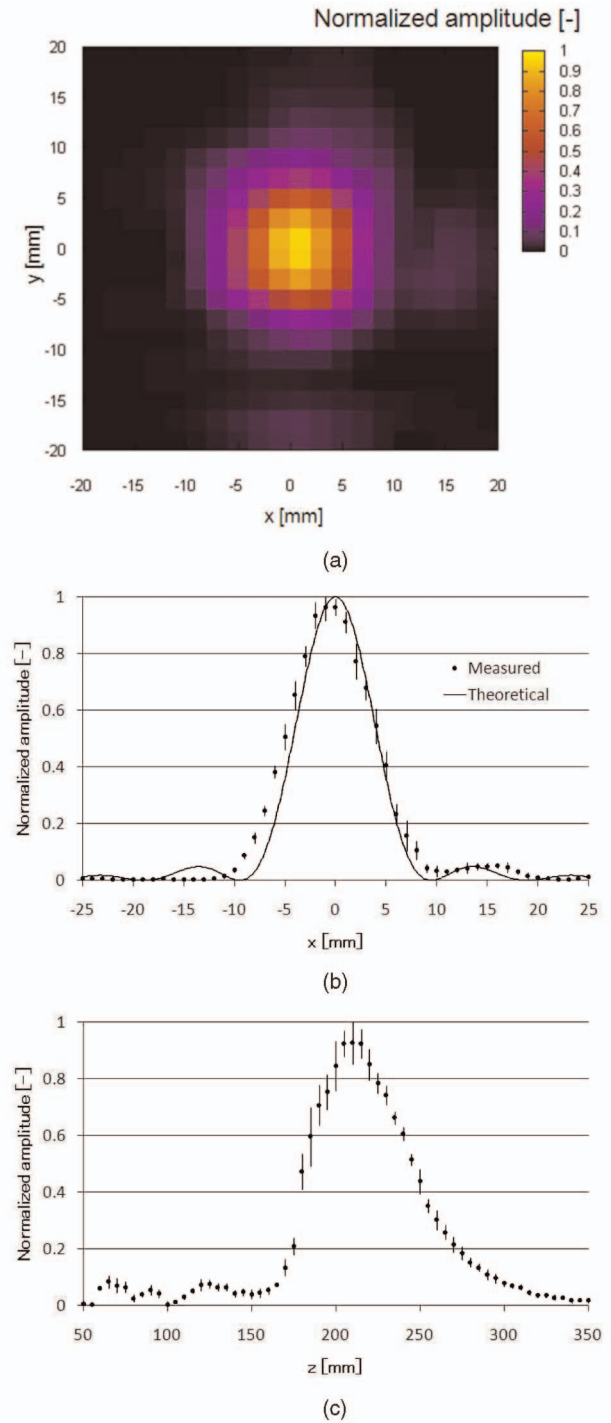


Fig. 11. Spatial distribution of measured radiation pressure. (a) Scanning along XY axes ($z = 200 \text{ mm}$). (b) Scanning along X axis ($y = 0 \text{ mm}$ and $z = 200 \text{ mm}$). (c) Scanning along Z axis ($x = 0 \text{ mm}$ and $y = 0 \text{ mm}$).

times of the ultrasound are 0.5 ms, those of the radiation pressure are nearly 1 ms.

Fig. 12c shows the frequency characteristics of the radiation pressure of the prototype. The horizontal axis represents the modulation frequency. The vertical axis represents the amplitude of radiation pressure (the maximum value during 1-minute observation) on a decibel scale. The maximum amplitude is the 0-dB reference. While the carrier (ultrasound) frequency was fixed at 40 kHz, the sound power of ultrasound was modulated by each frequency. The

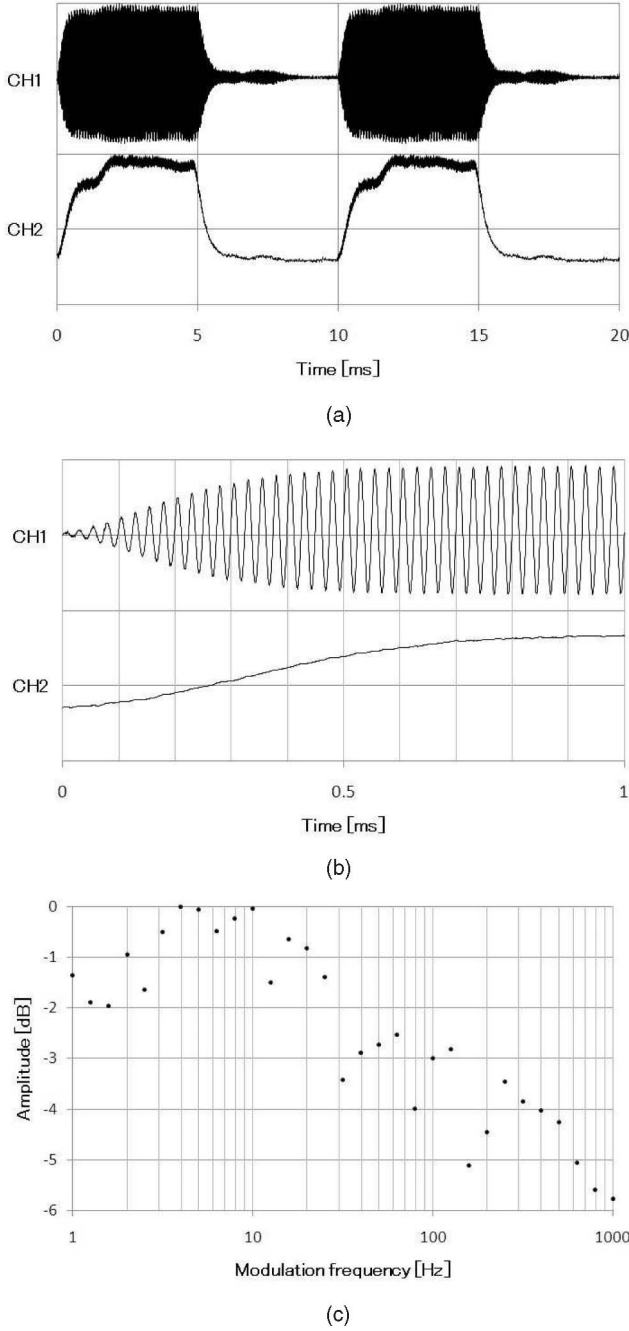


Fig. 12. Temporal properties of measured radiation pressure. (a) 100-Hz modulated waveforms measured at focal point; ultrasound (CH1) and radiation pressure (CH2). The amplitudes are normalized by their maximum values. (b) Closeup of (a) from 0 to 1 ms. (c) Frequency characteristics.

maximum and minimum amplitudes are 0 dB and -5.8 dB at 10 Hz and 1 kHz, respectively (i.e., the amplitude at 10 Hz is nearly twice as large as that at 1 kHz). That indicates that the proposed tactile display covers the bandwidth required for tactile displays.

4 EXPERIMENTS AND RESULTS

4.1 Direction Discrimination Test

A short experiment was conducted in order to confirm whether users can feel effective tactile sensation. Six

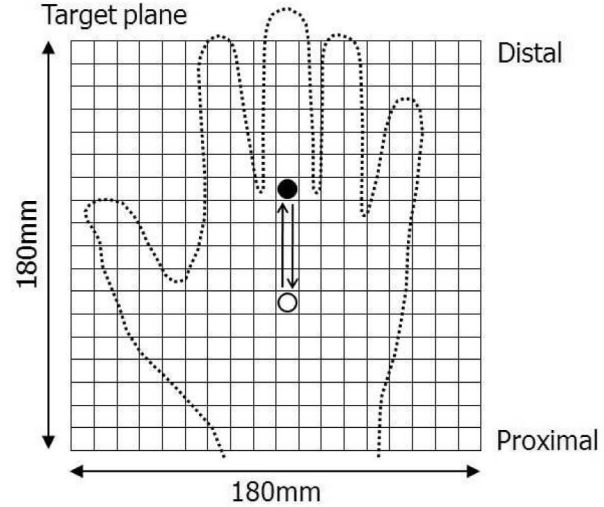


Fig. 13. Trajectory of focal point in direction discrimination test. It moves from the black dot to the white one or the white to the black.

volunteers (age range between 23 and 29 years, all male and right-handed) took part in it. The subject held his right hand at 200 mm above the radiation surface of the transducer array (i.e., at the target plane). Two types of linear movements (distal-to-proximal and proximal-to-distal, as shown in Fig. 13) were presented in a random order on his palm, and he was asked to answer the direction. Each direction was presented 20 times per person. The focal point moved by 5 cm at a speed of 5 cm/s. The modulation frequency was 100 Hz. During the entire trials, he wore headphones and heard a white noise, which prevented him from hearing the audible sound radiated from the tactile display (discussed in Section 6.2).

As a result, all of the subjects could discriminate the direction with 100 percent accuracy. This result indicates that they could feel tactile sensation reliably enough to perform that task.

4.2 Positioning Test

Another experiment was conducted in order to examine how precisely users can localize the position of the focal point. Eight volunteers (age range between 23 and 29 years, all male and right-handed) took part in it. At first, the focal point was presented at the center of the target plane. The subject was asked to remember the position on his palm where he felt the focal point. After that, the focal point was moved to another position, which was randomly selected from the eight positions shown in Fig. 14. He was asked to follow the focal point to feel it at the same position on his palm. The start and goal positions of his hand were recorded. He repeated the trial 16 times without pretraining (i.e., the first try was recorded as the first data). The focal point was moved at a speed of 3 cm/s. The modulation frequency was 200 Hz. The audible sound gave no hint for him to infer the motion or position of the focal point.

The position of the subject's hand was sensed as follows: A retroreflective marker was attached on the root of his middle finger. IR LEDs illuminated it and two IR cameras (Wii Remote [30], Nintendo Co. Ltd.) saw it. Based on triangulation, its 3D position was determined. Calibration was done by hand and its accuracy was less than 1 mm.

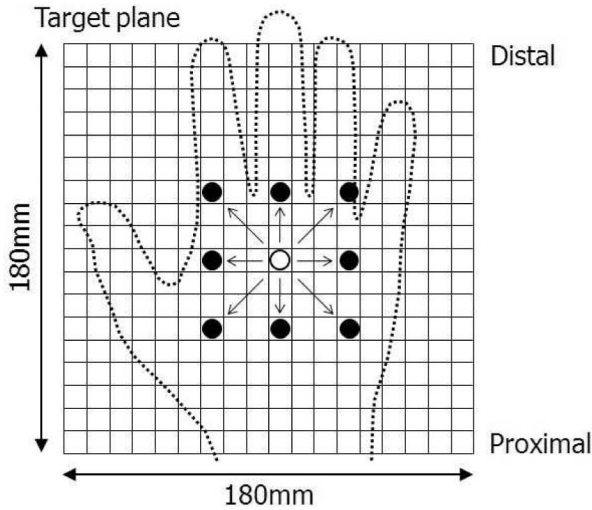


Fig. 14. Moves of focal point in positioning test. It moves from the white dot to one of the eight black dots.

The results are shown in Fig. 15. The positioning error is the difference between the displacements of the hand and the focal point on the XY plane. Fig. 15 shows the mean value as a black dot and the maximum and minimum values as both ends of a vertical bar for each subject. While the errors of Subjects E and G went beyond 30 mm three times and once, respectively, in the early stages of their trials, they achieved the tasks well after that. The mean value and standard deviation among the trials of all the subjects are 8.9 mm and 7.4 mm, respectively. That inversely indicates that people can stabilize their hands within 16.3 mm from a fixed focal point. For example, it can work as a home position for touch-typing on a virtual floating keyboard.

4.3 Various Movements

The tactile display has been demonstrated in several international conferences and Japanese domestic exhibitions, and more than 2,000 attendees tried the display. Five typical movements of the focal point,

1. line,
2. circle,
3. spiral, and
4. random,

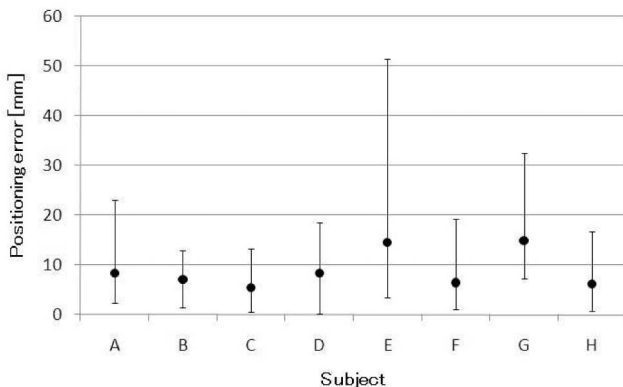


Fig. 15. Positioning errors of each subject.

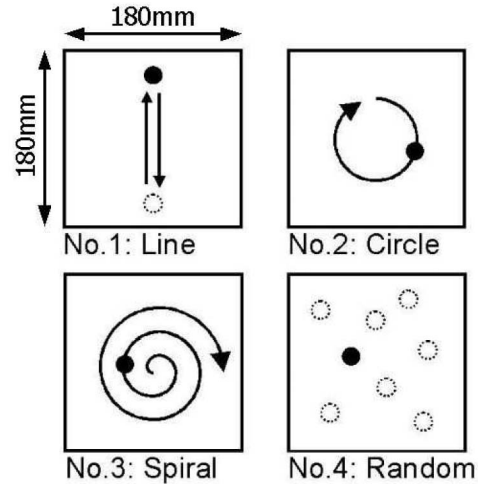


Fig. 16. Presented movements of focal point.

with 200 Hz modulation were presented on a user's palm. Fig. 16 shows the symbols that indicate these movements. All the attendees that answered our questions recognized the movements 1, 2, and 3 easily and correctly. That means the tactile display can surely produce the focal point along the intended trajectory. Besides, they said that the movement 4 evoked an interesting, tickling sensation like "an electrostatic" or "a fur." That means the tactile display has possibilities to reproduce not only tactile cues but also tactile feelings.

4.4 Reports on Tactile Feelings

The experimental subjects and conference attendees expressed the tactile feelings of the 100-Hz or 200-Hz modulated stimulation as follows: Some of them described the feeling as a kind of "electrostimulation," which might be because the modulation waveform was a rectangle. Others explained it as "a stream of air" because an air flow was also generated around the focal point (discussed in Section 6.1). A few people compared it to a tickling sensation rubbed with "a small soft-haired blush." Overall, all of them said "the pressure distribution is surely localized."

5 MULTIMODAL SYSTEM

An interaction system was developed by employing the airborne ultrasound tactile display (Fig. 17a). The system provides images floating in air. According to the position of a user's hand, the images are changed and tactile feedback is produced simultaneously. The user can feel floating images (e.g., bounding balls, raindrops, small creatures walking around, and so on) both visually and tactually (Figs. 17b and 17c). The demo movie is available in [31]. This system was demonstrated in SIGGRAPH 2009 Emerging Technologies for 5 days and nearly 1,000 attendees tried and enjoyed it [32].

The detailed descriptions of the components (the tactile display, the floating image display, and the hand tracker) are as follows: The tactile feedback is limited to the 200 Hz modulated stimulation, tentatively. In this system, by extending data size, the focal point can be moved along not only X and Y axes but also Z axis between 150 mm and 300 mm from the transducer array. The images are projected by HoloVision (Holo17T [4], Provision Interactive

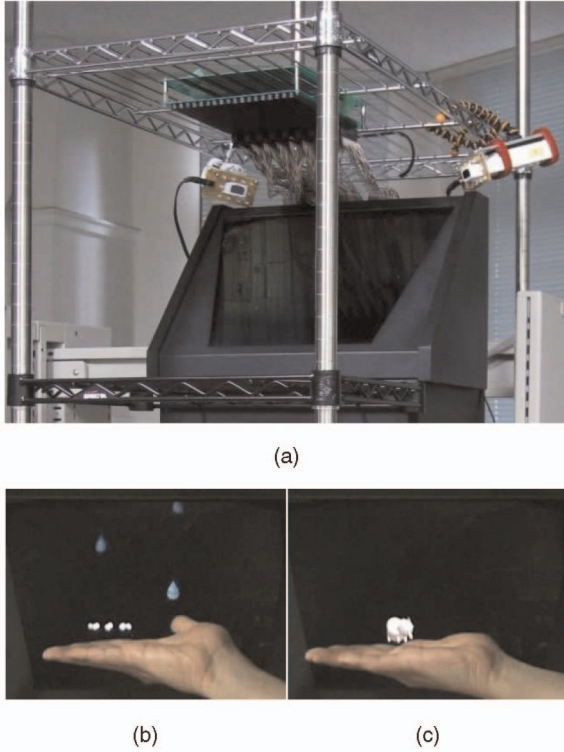


Fig. 17. (a) Interaction system consisting of an aerial imaging system, a noncontact tactile display, and an IR-based handtracking system. The ultrasound is radiated from above and the user feels as if a rain drop hits his palm in (b) or a small creature is walking around on his palm in (c). (a) System overview. (b) Raindrops. (c) Small creature.

Technologies, Inc.) which provides floating images from an LCD by utilizing a concave mirror. The projected images are 2D and float at 300 mm away from it. The position of the user's hand is sensed by the same IR-based method mentioned in Section 4.2. Although the single-point tracking is not very satisfactory, users can feel virtual objects.

6 DISCUSSIONS

6.1 Air Flow

The focal point is accompanied by an air flow localized around it, as mentioned in Sections 3.3 and 4.4. It is considered to be generated by the gradient of pressure. Assuming an ideal gas, it is explained by the Euler equation of fluid dynamics:

$$\frac{D\mathbf{u}}{Dt} = -\frac{1}{\rho} \nabla p_g, \quad (8)$$

where D/Dt is the Lagrangian Derivative operator, \mathbf{u} [m/s] is the particle velocity, ∇ is the del operator, and p_g is the pressure of gas. Equation (8) means the particle velocity is proportional to the gradient of the temporal integration of the pressure. Note that the time-varying component of p_g is the sound pressure p , and the time average of p leads to the acoustic radiation pressure P [18]. Therefore, the air flow felt on the skin increases in proportion to the impulse (the integral of the force with respect to time) applied on the skin, in a rough estimation. We reported in [28] that shortening the duration of radiation pressure was effective to reduce the air flow while the intensity of the feeling of impact

was maintained. Evaluating the more detailed effect of the air flow is one of the future works.

6.2 Audible Sound

The users hear some audible sound when they use the proposed tactile display. It is as loud as it can be heard within several meters from the display and some of users say it is annoying. Headphones for ear protection recommended in Section 2.4 are also helpful to prevent the users from hearing it.

There are two sources of the audible sound. One is the envelope of the ultrasound. If 100 Hz modulation is used, the 100 Hz audible sound is produced due to the nonlinearity of air, which is a phenomenon utilized in a directive loudspeaker [33]. The other is the discontinuity of the phases of the driving signals when the position of the focal point is discontinuously changed, which results in a crackle noise.

6.3 Arbitrary Pressure Distribution

The proposed tactile display has the ability to produce not only a single focal point but also an arbitrary pressure pattern. There are two approaches. One is the scanning or time-division method, which is discussed in [19]. In that method, the switching time of phase delays limits the number of focal points within one frame period which is preferred to be 1 ms for tactile stimulation. Since the rising time is 1 ms in the current system, as shown in Section 3.4, it results in one point per frame at the maximum.

The other is a simultaneous method, in which the sound energy is distributed by interference. Note that the radiation pressure at each point of this case is weaker than that of a single focal point. The driving pattern corresponding to a desired pressure pattern is determined by a solution of an inverse problem. An example of formulations is as follows: The definitions of mathematical symbols are the same as in Section 2.2. The same phase control factor $\exp(-jkr')$ is used. Let $g(x_m, y_n)$ [Pa] be a sound pressure on the surface of the m th row and n th column transducer. Then, the resulting sound pressure pattern is written as

$$\begin{aligned} p(x_0, y_0) &= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \frac{g(x_m, y_n)}{r} e^{-jkr'} e^{j(kr'' - \omega t)} \\ &\approx e^{j\left\{\frac{k}{2r}(x_0^2 + y_0^2 - x_c^2 - y_c^2) - \omega t\right\}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \frac{g(x, y)}{r} \times \text{rect}\left(\frac{x}{Nd}, \frac{y}{Nd}\right) \right. \\ &\quad \times \left. \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - x_m, y - y_n) \right\} e^{-j(\nu_x x + \nu_y y)} dx dy \\ &= e^{j\left\{\frac{k}{2r}(x_0^2 + y_0^2 - x_c^2 - y_c^2) - \omega t\right\}} \frac{N^2}{4\pi^2 r} G(\nu_x, \nu_y) * \text{sinc}\left(\frac{Nd}{2}\nu_x, \frac{Nd}{2}\nu_y\right) \\ &\quad * \left\{ \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} (-1)^{m+n} \delta\left(\nu_x - m\frac{2\pi}{d}, \nu_y - n\frac{2\pi}{d}\right) \right\}, \end{aligned} \quad (9)$$

where

$$\text{rect}(x, y) \equiv \begin{cases} 1, & (|x| \leq 1/2 \text{ and } |y| \leq 1/2), \\ 0, & (\text{otherwise}), \end{cases} \quad (10)$$

and $\delta(x, y)$ is the Dirac's delta function and the asterisk $*$ means convolution.

As shown in the third line of (9), G , the Fourier transform of g , is blurred by sinc (according to the array size) and

copied by the series of δ (according to the interval between the transducers). The interval between the copies is $2\pi r/kd$ ($=170$ mm when $r = 200$ mm), and therefore, G should be localized within a smaller region than that interval to avoid spatial aliasing (i.e., overlapping each other). The above results indicate that if one wants to produce the pattern of sound pressure $p(x_0, y_0) = (N^2/4\pi^2 r)G(\nu_x(x_0), \nu_y(y_0))$ in the neighborhood of the focal point, neglecting the phase and the blur effect, the corresponding driving pattern g is determined by the inverse Fourier transform of G .

6.4 Tactile Rendering

Our tactile display has advantages in rendering tactile feelings (e.g., textures, furs, particles, and so on). First, it does not provide an undesired tactile feeling originating from the contact between the device and the skin because it is noncontact. Second, its temporal bandwidth is broad enough to produce tactile feelings although the spatial bandwidth (or resolution) is limited. Realizing such a broad bandwidth is one of the important issues for tactile displays [34]. Last, it is free from the "contact problem" unlike vibrating pins that sometimes lose contact with the skin [35] even when a user does not move his/her hand. That is because our tactile display directly controls pressure which has some focal depth, instead of displacements of pins.

Although the temporal pattern of the modulation signal is currently limited to a rectangular wave for simplicity of the circuits, we expect that it is possible to reproduce a class of realistic tactile feelings by using adequate modulation signals. Exploring such signals is one of the future works.

7 CONCLUSION

In this paper, a tactile display using airborne ultrasound was presented. It enables users to feel virtual objects without mechanical touches on any display devices. It utilized a nonlinear effect of ultrasound: Acoustic radiation pressure. Controlling phase delays to generate a focal point was equivalent to placing a lens, and it was formulated based on the Fresnel (or near-field) diffraction. After that, the current prototype was introduced. It could produce 16 mN at the focal point. The spatial resolution was 20 mm. It could produce vibrations up to 1 kHz. Although the produced force was weak for users to feel a constant force, it was sufficient for vibratory sensation. Two experiments were conducted. As a result, it was confirmed that 1) users could feel a localized focal point clearly enough to discriminate its moving direction and 2) they could stabilize their hands within 16.3 mm from it. A trial to combine the ultrasound-based tactile display and a midair image display was also shown.

The future works are roughly divided into three issues. The first one is to intensify the total force by enlarging the device area so that more transducers are driven. The second issue is to control the pressure distribution and the force direction. Producing various spatial distributions is still a future work. Multiarrays to generate multidirectional forces are needed for an application in which a user touches VR objects from arbitrary angles. The last one is the issues about how to make signals in tactile rendering. Since the spatial resolution is limited, the selective stimulation [36] or multiprimitive [37] methodology based on the properties of the human skin or tactile perception will be required for producing realistic tactile feelings.

ACKNOWLEDGMENTS

This work was partly supported by the Japan Society for the Promotion of Science (JSPS) Research Fellowship for Young Scientists (19-1708) and JSPS Grant-in-Aid for Scientific Research (18206046). The authors were provided with HoloVision by Provision Interactive Technologies, Inc. during SIGGRAPH 2009.

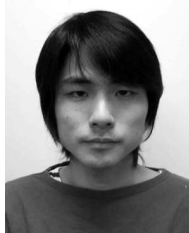
REFERENCES

- [1] I. Rakkolainen, "How Feasible Are Star Wars Mid-Air Displays?" *Proc. 11th Int'l Conf. Information Visualization (IV '07)*, pp. 935-942, 2007.
- [2] FogScreen, <http://www.fogscreen.com/>, 2010.
- [3] Heliodydisplay, <http://www.io2technology.com/>, 2010.
- [4] HoloVision, <http://www.provision3dmedia.com/>, 2010.
- [5] H. Liao, M. Iwahara, Y. Katayama, N. Hata, and T. Dohi, "Three-Dimensional Display with a Long Viewing Distance by Use of Integral Photography," *Optics Letters*, vol. 30, no. 6, pp. 613-615, 2005.
- [6] Real-Time Holography, <http://www.seereal.com/>, 2010.
- [7] T. Rodriguez, A.C. de Leon, B. Uzzan, N. Livet, E. Boyer, F. Geffray, T. Balogh, Z. Megyesi, and A. Barsi, "Holographic and Action Capture Techniques," *Proc. ACM SIGGRAPH '07*, 2007.
- [8] J. Allard, C. Menier, B. Raffin, E. Boyer, and F. Faure, "Grimace: Markerless 3D Interactions," *Proc. ACM SIGGRAPH '07*, 2007.
- [9] M.B. Khoudja, M. Hafez, J.M. Alexandre, and A. Kheddar, "Tactile Interfaces: A State-of-the-Art Survey," *Proc. 35th Int'l Symp. Robotics (ISR '04)*, 2004.
- [10] CyberTouch, <http://www.est-kl.com/products/data-gloves/cyberglove-systems/cybertouch.html>, 2010.
- [11] K. Minamizawa, S. Kamuro, S. Fukamachi, N. Kawakami, and S. Tachi, "GhostGlove: Haptic Existence of the Virtual World," *Proc. ACM SIGGRAPH '08*, 2008.
- [12] S.-C. Kim, C.-H. Kim, T.-H. Yang, G.-H. Yang, S.-C. Kang, and D.-S. Kwon, "SaLT: Small and Lightweight Tactile Display Using Ultrasonic Actuators," *Proc. 17th IEEE Int'l Symp. Robot and Human Interactive Comm. (RO-MAN '08)*, pp. 430-435, 2008.
- [13] K. Sato, K. Minamizawa, N. Kawakami, and S. Tachi, "Haptic Telexistence," *Proc. ACM SIGGRAPH '07*, 2007.
- [14] Y. Suzuki and M. Kobayashi, "Air Jet Driven Force Feedback in Virtual Reality," *IEEE Computer Graphics and Applications*, vol. 25, no. 1, pp. 44-47, Jan./Feb. 2005.
- [15] T. Iwamoto, M. Tatezono, and H. Shinoda, "Non-Contact Method for Producing Tactile Sensation Using Airborne Ultrasound," *Proc. EuroHaptics 2008*, pp. 504-513, 2008.
- [16] T. Hoshi, T. Iwamoto, and H. Shinoda, "Non-Contact Tactile Sensation Synthesized by Ultrasound Transducers," *Proc. Third Joint Eurohaptics Conf. and Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems (World Haptics '09)*, pp. 256-260, 2009.
- [17] J. Awatani, "Studies on Acoustic Radiation Pressure. I. (General Considerations)," *J. Acoustical Soc. Am.*, vol. 27, pp. 278-281, 1955.
- [18] T. Hasegawa, T. Kido, T. Iizuka, and C. Matsuoka, "A General Theory of Rayleigh and Langevin Radiation Pressures," *Acoustical Science and Technology*, vol. 21, no. 3, pp. 145-152, 2000.
- [19] T. Iwamoto and H. Shinoda, "Two-dimensional Scanning Tactile Display Using Ultrasound Radiation Pressure," *Proc. Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems (IEEE Haptics Symp. '06)*, pp. 57-61, 2006.
- [20] T. Togawa, T. Tamura, and P.A. Öberg, *Biomedical Transducers and Instruments*. CRC Press, 1997.
- [21] H.E. Bass, L.C. Sutherland, A.J. Zuckerwar, D.T. Blackstock, and D.M. Hester, "Atmospheric Absorption of Sound: Further Developments," *J. Acoustical Soc. Am.*, vol. 97, pp. 680-683, 1995.
- [22] K. Shinagawa, Y. Amemiya, H. Takemura, S. Kagami, and H. Mizoguchi, "Three Dimensional Simulation and Measurement of Sound Pressure Distribution Generated by 120 ch Plane Loud-speaker Array," *Proc. 2007 IEEE Int'l Conf. Systems, Man and Cybernetics (SMC '07)*, pp. 278-283, 2007.
- [23] Ultrasound Transducer, [http://www.nicera.co.jp/pro/ut/pdf/T4010A1\(ENG\).pdf](http://www.nicera.co.jp/pro/ut/pdf/T4010A1(ENG).pdf), 2010.
- [24] F. Träger, *Springer Handbook of Lasers and Optics*. Springer, 2007.

- [25] A.B. Vallbo and R.S. Johansson, "Properties of Cutaneous Mechanoreceptors in the Human Hand Related to Touch Sensation," *Human Neurobiology*, vol. 3, pp. 3-14, 1984.
- [26] R.K. Creasy, R. Resnik, and J.D. Iams, *Maternal-Fetal Medicine*. Saunders, 1999.
- [27] C.Q. Howard, C.H. Hansen, and A.C. Zander, "A Review of Current Ultrasound Exposure Limits," *The J. Occupational Health and Safety of Australia and New Zealand*, vol. 21, no. 3, pp. 253-257, 2005.
- [28] T. Hoshi, D. Abe, and H. Shinoda, "Adding Tactile Reaction to Hologram," *Proc. 18th IEEE Int'l Symp. Robot and Human Interactive Communication (IEEE RO-MAN '09)*, pp. 7-11, 2009.
- [29] C. Audoly, "Some Aspects of Acoustic Interactions in Sonar Trans-Ducer Arrays," *J. Acoustical Soc. Am.*, vol. 89, pp. 1428-1433, 1991.
- [30] Wii Remote, <http://www.nintendo.com/wii/>, 2010.
- [31] Touchable Holography, <http://www.youtube.com/watch?v=Y-P1zZAcPuw>, 2010.
- [32] T. Hoshi, M. Takahashi, K. Nakatsuma, and H. Shinoda, "Touchable Holography," *Proc. ACM SIGGRAPH '09*, 2009.
- [33] M. Yoneyama, J. Fujimoto, Y. Kawamo, and S. Sasabe, "The Audio Spotlight: An Application of Nonlinear Interaction of Sound Waves to a New Type of Loudspeaker Design," *J. Acoustical Soc. Am.*, vol. 73, pp. 1532-1536, 1983.
- [34] J. Pasquero and V. Hayward, "STReSS: A Practical Tactile Display System with One Millimeter Spatial Resolution and 700 Hz Refresh Rate," *Proc. Eurohaptics 2003*, pp. 94-110, 2003.
- [35] J.Z. Wu, R.G. Dong, A.W. Schopper, and W.P. Smutz, "Analysis of Skin Deformation Profiles During Sinusoidal Vibration of Fingert," *Annals of Biomedical Eng.*, vol. 31, pp. 867-878, 2003.
- [36] N. Asamura, N. Yokoyama, and H. Shinoda, "Selectively Stimulating Skin Receptors for Tactile Display," *IEEE Computer Graphics and Applications*, vol. 18, no. 6, pp. 32-37, Nov. 1998.
- [37] Y. Makino, N. Asamura, and H. Shinoda, "Multi Primitive Tactile Display Based on Suction Pressure Control," *Proc. 12th IEEE Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems (Haptic Symp. '04)*, pp. 90-96, 2004.



Masafumi Takahashi received the BE degree from the University of Tokyo, Japan, in 2009. Currently, he is working toward the ME degree in the Graduate School of Information Science and Technology, The University of Tokyo, Japan. He is interested in tactile displays. He is a member of the VRSJ.



Takayuki Iwamoto received the BE, ME, and PhD (information science and technology) degrees from the University of Tokyo, Japan, in 2001, 2003, and 2006, respectively. He was a research fellow at JSPS from April 2003 to March 2006. Currently, he works at Canon, Inc. He won the Best Paper Award at ICAT 2001, the Best Student Paper Award at World Haptics 2005, and the Young Author's Award at SICE 2005. He is a member of the SICE and the VRSJ.



Hiroyuki Shinoda received the BS degree in applied physics, the MS degree in information physics, and the PhD degree in electrical engineering from the University of Tokyo, Japan, in 1988, 1990, and 1995, respectively. He was a lecturer from 1995 and an associate professor from 1997, both in the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Japan. He was a visiting researcher at UC Berkeley in 1999. From 2000, he is an associate professor in the Graduate School of Information Science and Technology, The University of Tokyo, Japan. His research interests include information physics, tactile/haptic interfaces, sensor systems and devices, sensor networks, 2D communication, human interfaces, and optical/acoustic measurement. He and his students have won eight international conference awards and a lot of Japanese academic awards. He is a board member of the SICE and a member of the IEEE, the RSJ, the VRSJ, and the JSME.



Takayuki Hoshi received the BE, ME, and PhD (information science and technology) degrees from the University of Tokyo, Japan, in 2003, 2005, and 2008, respectively. He was a research fellow at JSPS from April 2007 to March 2009. Currently, he is an assistant professor in the Graduate School of Science and Technology, Kumamoto University, Japan. He won the SICE International Award at SICE 2007. He is interested in tactile sensors and displays. He is a member of the SICE, the RSJ, and the VRSJ.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.