IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 2 April 2024, accepted 27 May 2024, date of publication 5 June 2024, date of current version 21 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3409736

RESEARCH ARTICLE

Two-Dimensional Remote Actuation of Centimeter-Scale Objects With a Concave Bottom Using Airborne Ultrasound

GAKUTO ARAKAWA[1], RYUNOSUKE HIRAI[2], SHUN SUZUKI[1], TAKAAKI KAMIGAKI[1], YASUTOSHI MAKINO[1], AND HIROYUKI SHINODA[1], (Member, IEEE)

[1]Department of Complexity Science and Engineering, The University of Tokyo, Kashiwa-shi, Chiba 277-8561, Japan
[2]Department of Information Physics and Computing, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

Corresponding author: Gakuto Arakawa (arakawa@hapis.k.u-tokyo.ac.jp)

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 21H05301.

ABSTRACT We propose a method for two-dimensional remote position control of centimeter-scale objects using airborne ultrasound. To actuate objects 1–2 cm in size, previous studies utilized magnetic methods because they are too small to incorporate motors and batteries. In addition to magnets, airborne ultrasound, which produces a noncontact force, can actuate various objects from the outside. However, when the size of an object is larger than the ultrasound wavelength, methods using standing waves cannot be applied. We introduced an acoustically transparent screen and designed objects with concave bottoms so that the position of objects with a size larger than the wavelength could be controlled on the screen by an ultrasonic focus. Because our method relies only on the concave bottom of the object, its materials and visible upper parts can be designed according to the usage scenario. For example, food and transparent objects, which are difficult to equip with mechanical components, can be remotely actuated. It was confirmed that an object with a concave bottom diameter of 12 mm moved continuously at an average speed of 28.5 mm/s. Actuating multiple objects is also achieved by the time-division control of the focal point without depending on their mutual positional relationships.

INDEX TERMS Acoustic manipulation, airborne ultrasound, centimeter-scale object, multi object control, untethered actuation, two-dimensional position control.

I. INTRODUCTION

The size of a standalone small robot is primarily determined by the sizes of its motor and battery. Simple small robots for swarm robot applications, such as Kilobot [1], Sony Toio, and Zooids [2], have built-in motors and batteries and have sizes of approximately 3 cm per side. The lower limit of the robot size is approximately 3 cm per side when typical motors and batteries are used. Inspired by insects and animals in nature, other untethered drive methods have been proposed for moving robots smaller than 3 cm [3], [4]. In this study, we focused on remotely actuating smaller objects, 1–2 cm in size and made of diffrent materials, in a two-dimensional workspace.

Actuation methods using passive mechanisms without built-in motors and batteries have been studied at the scale of micrometers to centimeters [5]. In particular, attention has been focused on methods using magnets that can be driven in vivo. Soft materials allow deformation and movement [6], [7], [8] as well as designs and motions that mimic those of actual living organisms [4], [9]. Objects made of paper [10] or wood [11] can be actuated from the outside; however, because of the use of magnetic force, it is necessary to attach a magnet to the object or magnetize it.

The method of applying force through air using airborne ultrasound does not require the object to be actuated to have

The associate editor coordinating the review of this manuscript and approving it for publication was You Yang.

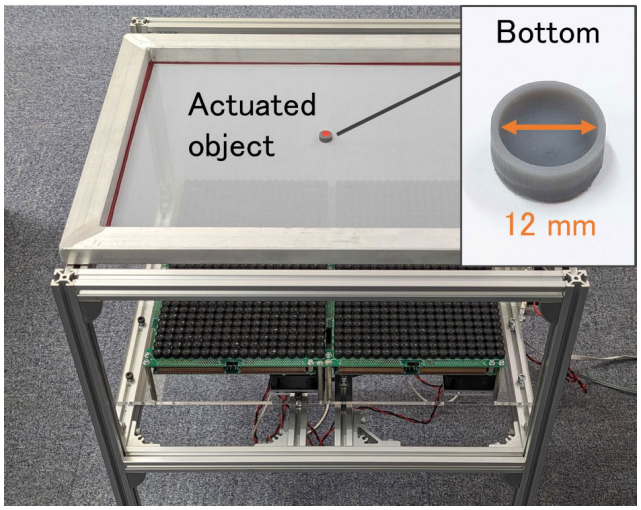

FIGURE 1. The proposed system. Airborne ultrasound phased arrays (AUPAs) and an acoustically transparent mesh screen are used to control the two-dimensional positions of objects on the screen. The bottoms of the objects have concave shapes with a diameter larger than the ultrasound wavelength (8.5 mm).

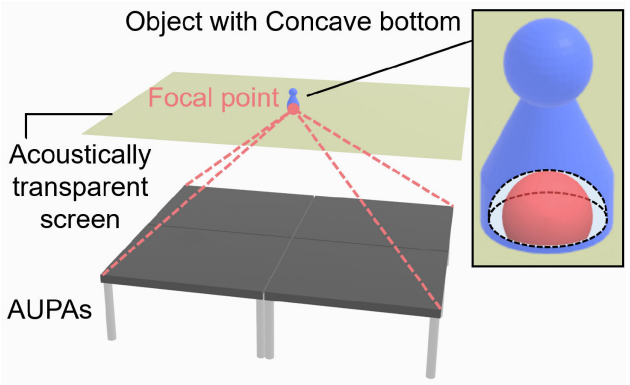

FIGURE 2. System overview. An ultrasound focal point is generated at the concave bottom of the object. The ultrasound is not blocked by the acoustically transparent screen or by other objects on the screen.

any specific physical properties other than large acoustic impedance. Typical examples include trapping spheres of a few millimeters in a standing-wave acoustic pressure node for levitation and manipulation [12], [13], rotation control [14], and optimal trajectory design [15], [16]. This technique can be applied to liquids, food [17], and insects [18]. An airborne ultrasound phased array (AUPA), which is used to control the acoustic field in a space, typically consists of 40 kHz transducers. The wavelength of ultrasound in air is approximately 8.5 mm at room temperature. Because standing wave methods can be applied to half-wavelength objects, they can be used to control objects in a three-dimensional space as long as they are smaller than 5 mm. Several studies have used different ultrasonic techniques to control spherical objects with diameters larger than 5 mm, for example, the three-dimensional position control of a 10-cm floating balloon injected with helium gas [19] and the levitation of a 3-cm polystyrene sphere in a steady position [20]. In addition, liquid droplets [21] or solid spheres have been moved on a support plane [22].

In this paper, we propose a method for controlling the two-dimensional positions of centimeter-sized objects with customizable shapes and materials on a plane using airborne ultrasound. As shown in Fig. 1, the proposed system consists of AUPAs, an acoustically transparent screen, and objects with concave bottoms. By directing the ultrasound from the AUPAs positioned below the screen toward the bottom of the object, the acoustic radiation force acting in the direction normal to the surface can be used as a driving force on the plane. In this case, the direction of the force received by the concave bottom is toward the center of the focal point; therefore, the object position can be controlled in a feedforward manner. This study focuses only on feedforward control for the basic verification of the proposed method; however, feedback control can be introduced in the future. Additionally, the ultrasound is not blocked below the screen even when there are objects on the screen other than the actuation target. Therefore, the system can easily be extended to multiple object actuations. The driving force of the object is derived from the ultrasound emitted at the concave bottom of the object, allowing the visible upper shape and overall material of the object to be customized. Thus, the proposed method can be used to remotely actuate objects that cannot be equipped with mechanical drive components, such as bite-sized candies, other foods, or objects that are entirely transparent. This method can be used to actuate a moving prey model in observational experiments with small animals such as insects and mice. Alternatively, it can be incorporated into interactive media for children, in which users can eat the target.

In the following sections, we verify the concept of the proposed actuation method according to the following four basic characteristics: i) Relationship between the height from the surface of the AUPAs to the workspace and the strength of the acoustic flow, a factor that prevents accurate control of objects. ii) Differences in the horizontal propulsive force extracted from the acoustic radiation force depending on the shape of the bottom. iii) Time-response characteristics of the movement of objects from rest. iv) Actuation of a single object and simultaneous actuation of multiple objects using a time-division method.

II. DESIGN AND CONFIGURATION

A. SYSTEM OVERVIEW

Fig. 2 presents an overview of the system for remote actuation of an object on a plane. The system comprises AUPAs and an acoustically transparent screen. The object has a concave surface at its bottom. The screen acts as a manipulation workspace and is installed horizontally, whereas the AUPAs are positioned under it.

As stated in a previous study [22], the propulsive force can also be applied to the objects by placing AUPAs at sides

of the workspace. In addition, AUPAs can be placed above the workspace. Compared to installing AUPAs at the sides or above, installing AUPAs below has the following advantages.

- The ultrasound is not blocked by any objects or animals on the workspace other than the target object.
- It can be easily observed from above.
- The direction of travel of the ultrasound is opposite to the gravitational acceleration, which reduces friction with the ground surface when the object is moved.
- Increasing the number of AUPAs is easy and there is no risk of falling.

The system is thus unaffected by the positional relationship of objects when operating multiple objects simultaneously, and it is easy to observe, record, and interact with objects in the workspace.

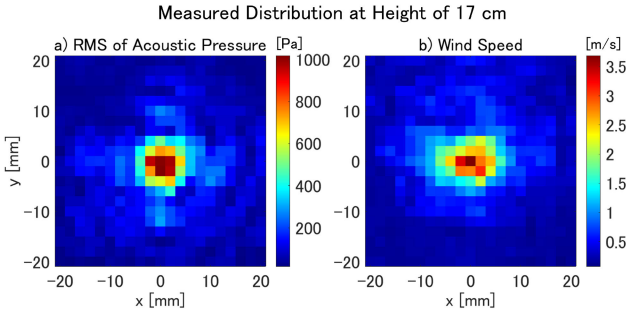

FIGURE 3. a) Acoustic pressure distribution and b) z-directional wind speed distribution measured near the focal point formed at a height of 17 cm. The acoustic pressure amplitude normalized by [0,1] was set to 0.1 in a) and 1 in b) for the four AUTD3 units.

B. DISTANCE BETWEEN AUPAS AND THE SCREEN

1) REASONS TO CONSIDER SCREEN HEIGHT

Although AUPAs can generate a focal point at an arbitrary height, this method requires consideration of the focal point height. The proposed method uses acoustic radiation pressure, a nonlinear phenomenon of ultrasound, as the driving force for an object. However, ultrasound simultaneously generates a flow in the medium itself known as acoustic flow, which is a different nonlinear phenomenon. The acoustic radiation pressure acts in the direction normal to the reflecting surface, whereas the acoustic flow acts in the direction of the flow; therefore, a low horizontal thrust can be extracted. Additionally, the acoustic flow has a slow start-up and low stability in the direction and magnitude of the flow, which interferes with stable control.

The dominant acoustic flow near the focal point is called Eckhardt flow, which is generated by the acoustic wave providing the force described below:

$$F_{\text{flow}} = -\frac{1}{\rho}\nabla\langle\mathcal{L}\rangle + \frac{\delta}{\rho^2 c^4}\langle p\rangle\nabla\frac{\partial p}{\partial t}, \tag{1}$$

where ρ is the density of the medium, c is the sound velocity, p is the acoustic pressure, t is time, $\mathcal{L}$ is the Lagrangian, δ is a coefficient related to the sound absorption, and $\langle\cdot\rangle$ is the time-averaging operator. This means that the flow is supplied with energy by the sound wave and thus accelerates as the sound wave proceeds. The theoretical determination of wind velocity and comprehensive analysis with radiation pressure require solving the Navier–Stokes equation, which is complex and beyond the scope of this study. Therefore, through measurements using actual equipment, we confirmed the existence of a range of distances where the effect of acoustic flow is negligibly small compared to the acoustic radiation pressure. The screen height from the surface of the AUPAs, H_{screen}, was selected from this range and used as the focal point height.

2) SELECTION OF THE SCREEN HEIGHT

In this study, the system utilized four AUTD3 [23] units as AUPAs and a polyester mesh screen as an acoustically transparent screen. The AUTD3 units emit ultrasound of 40 kHz, and are controlled by the AUTD3 software to generate an ultrasound focal point at any desired location in space. The four AUTD3 units with an overall aperture of 38.4 cm × 30.3 cm and 996 transducers were driven at a 24 Vp-p peak to peak voltage. The mesh screen had a thread diameter of 80 μm, and it was evenly distributed through a 100-mesh-per-inch grid. The resulting aperture ratio is approximately 49%. As determined by a standard microphone (Type 4138-A-015, Brüel & Kjær), this mesh screen efficiently transmitted ultrasound at a rate of approximately 95% in terms of acoustic pressure, irrespective of the distance from the surface of AUTD3 units. The mesh screen involved a 40 cm × 30 cm portion of the mesh and was installed centrally over the AUTD3 units. This mesh screen was used in the experiments described in Section IV.

In this study, H_{screen} was empirically set to 17 cm. Fig. 3 shows the measured distributions of acoustic pressure and wind speed near the focal point created at the height directly above the center of AUPAs. The acoustic pressure was measured using a standard microphone, and the wind speed was measured using an anemometer (Climomaster model 6531-21, KANOMAX). The microphone and anemometer were attached to a three-axis stage to scan the area near the focal point at 2-mm intervals while maintaining the height. Near the focal point, both the acoustic pressure and wind speed had a spread of 8–12 mm, which was similar to that of the wavelength; however, the wind speed had a slightly larger spread. This indicates that the localization of the acoustic flow is less than that of the acoustic pressure and that there is no advantage in actively using acoustic flow in terms of spatial resolution. When the focal point was formed at this height, the total force received near the focal point was 13 mN. The force received from wind pressure in the area of an 8-mm square near the center of the focal point was approximately 0.4 mN, as estimated from the measured airflow. Therefore, the ratio of the force originating from the acoustic flow to the total force received near the focal point was approximately 3%, which is sufficiently small.

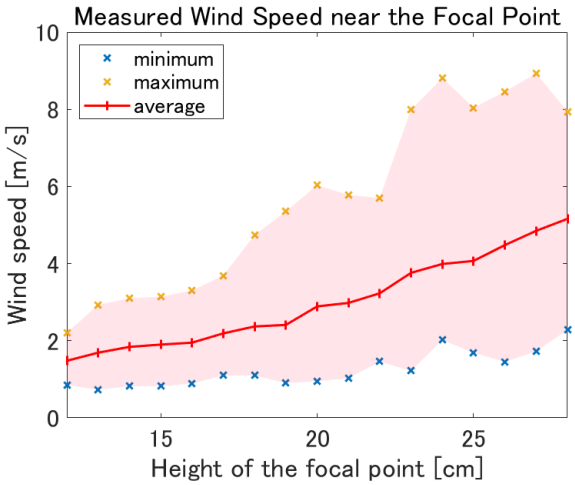

FIGURE 4. Minimum and maximum wind speeds measured near the focal point at various focal heights. The measurement area was a square-shaped area of 8 mm × 8 mm at the focal height, with 2-mm intervals. The calculated average values are plotted in red.

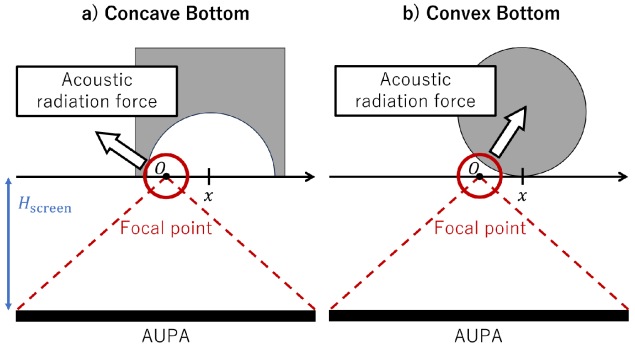

FIGURE 5. Difference in the direction of the acoustic radiation force received by object bottoms at position x from the focal point. With a) concave bottom, the force acts in the direction toward O, whereas with b) convex bottom, the force acts in the direction away from O. In this study, we consider only a).

To confirm the effect of height on acoustic flow, the wind speeds near the focal point were measured at different focal point formation heights. Fig. 4 shows the minimum and maximum wind speeds over an 8-mm square area measured at 2-mm intervals. The average values within the measured area increased with height. When the wind speed is v and air density is ρ, the wind pressure P_{wind} is given by $P_{\text{wind}}=\rho v^2/2$. Assuming $\rho = 1.293$ kg/m^3, the wind pressure calculated from the average wind speed was multiplied by the measurement area and converted into force, which was {0.22, 1.21, 4.74} mN over {15, 20, 25} cm. However, the focal acoustic pressure decreased because of the distance attenuation of the ultrasound as H_{screen} increased, such that the proportion of the wind pressure from the acoustic flow in the total pressure increased. In the setup of this study, in the range of $H_{\text{screen}} \leq 19$ cm, the force owing to the wind pressure near the focal point did not exceed 1 mN and had no significant influence on the actuation.

III. METHOD OF REMOTE ACTUATION

A. STRATEGY TO CONTROL OBJECTS

In this method, acoustic radiation force is used as the driving force for the object by applying ultrasound to the bottom of the object from below. The acoustic radiation force acts on the boundary surface when sound waves are reflected from surfaces with different acoustic impedance values. This force acts in the direction of the inward normal to the surface, as long as the object surface does not deform. Therefore, the shape of the bottom of the object can be designed to extract the propulsive force in the desired direction. The acoustic radiation pressure is $P_{\text{rad}} = (\alpha/\rho c^2)p^2$, where α is the reflection coefficient.

The bottom shapes of the objects are broadly classified into two categories: concave and convex. When the focal point on the screen is at origin O and the object is x mm from the origin, an acoustic radiation force acts on the object, as shown in Fig. 5. The direction of this force has horizontal and vertical components with different ratios depending on the bottom geometry. Fig. 5 a) shows that the ultrasound focal point attracts the object like a tractor beam when the bottom is concave. The object is pushed by ultrasound from inside the concave bottom and is stable when the center of the focal point and the object coincide. However, as shown in Fig. 5 b), an object with convex or flat bottom such as a sphere moves away from the focal point. In this case, the system requires feedback control combined with positional measurements. For stable two-dimensional position control of an object with a focal point that has a strong acoustic pressure, we used concave bottoms in this study.

As the ultrasound focal point had a circular distribution in the horizontal direction (Fig. 3), the symmetric shape of the bottom made the control strategy independent of the direction of movement. With a vertically axisymmetric concave bottom, the horizontal component of the force was directed toward the center of the focal point, independent of the displacement direction. Using a symmetrical concave bottom that covers the focal diameter, the object follows the focal point, allowing the feedforward control of the position of the object.

B. CONCAVE BOTTOM OF THE OBJECTS

In this study, concave shapes with horizontal symmetry were used as the object bottoms; that is, the horizontal cross-sections of the objects were concentric circles. Conical and semi-ellipsoidal concave bottoms were used as typical examples (Fig. 6a) and b)). The two concave shapes satisfy the following conditions, which ensure that no part of the shapes blocks the ultrasound from reaching its interior: 1) The radius of the concentric circles decreases monotonically from the bottom to the top, and 2) the concave surface is convex upward.

The parameters of the concave bottoms are the aperture radius r and depth of the center h. For a concave surface

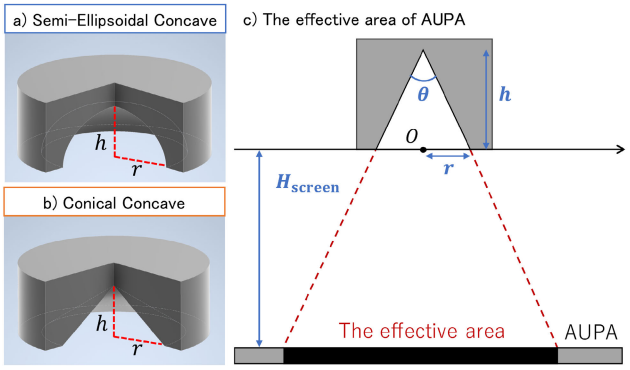

FIGURE 6. Objects with concave bottoms with aperture radius r and depth of the center h. We deal with two typical examples, a) semi-ellipsoidal concave and b) conical concave in this paper. c) The effective area of the AUPA becomes smaller when the apex angle θ of the concave bottom is small. Assuming linear propagation of ultrasound, the ultrasound emitted from transducers outside the area will be reflected at the sides and edges of the object.

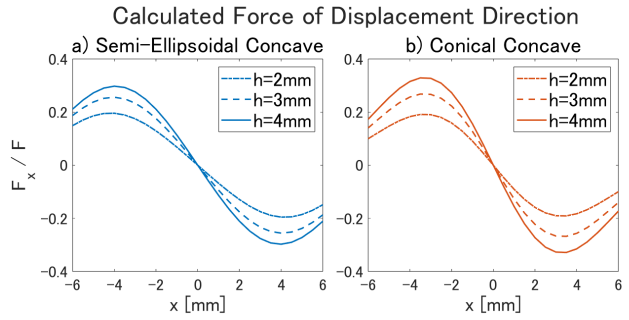

FIGURE 7. Numerical results for the horizontal component of the acoustic radiation force acting on objects with a) semi-ellipsoidal and b) conical concave bottoms. The horizontal axis shows the displacement relative to the focal point, and the vertical axis shows the x-directional force normalized by the total acoustic radiation force.

to cover the focal point, r must be greater than the focal radius. As the focal diameter is approximately equal to the wavelength λ, as shown in Fig. 3, r should satisfy $2r \geq \lambda = 8.5$ mm. The overall size of the object was approximately 1—2 cm, including the thickness of the edge of the object, which was the target of this study. To simplify the comparison of h, the ultrasound focal point in this study was formed at the screen height H_{screen}. The farther the concave bottom, the reflecting surface, is from the focal point, the less it is subjected to acoustic radiation pressure from the focal point. We set h to satisfy $h \leq \lambda/2 = 4.25$ mm, assuming that the vertical and horizontal diameters of the focal point are approximately λ.

C. NUMERICAL CALCULATION OF HORIZONTAL PROPULSIVE FORCE

The net acoustic radiation forces acting on the concave bottom of the object were numerically calculated using $H_{\text{screen}} = 17$ cm. Conical and semi-ellipsoidal concave shapes were used, with $r = 6$ mm and $h = 2$, 3, and 4 mm. r was empirically selected as a value that could be stably manipulated and corresponded to the experiment described below. When $r = 6$ mm and $h = 4$ mm, the vertex angle shown in Fig. 6 c) is $\theta = 109°$. Let the focal point be generated at origin O, at a distance H_{screen}directly above the center of the AUPAs. The diagonal length of the four AUTD3 units is 48.9 cm, which is approximately equal to the diameter of the effective transducer area ($= 2(H_{\text{screen}}+h)\tan\frac{\theta}{2} = 48.8$ cm). Thus, assuming rectilinear propagation of ultrasound, few transducers are ineffective when the object is at O.

Given l point sound sources, the acoustic radiation pressure $P_{\text{rad}}(\boldsymbol{r}) = (\alpha/\rho c^2)p(\boldsymbol{r})^2$ at position $\boldsymbol{r}$ at $t = 0$ is given by the acoustic pressure

$$p(\boldsymbol{r}) = \sum_{i=1}^{l} p_0 \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_i|} e^{jk(|\boldsymbol{r}-\boldsymbol{r}_i|-|\boldsymbol{r}_i|)}, \tag{2}$$

where, p_0 is a constant related to the amplitude. The initial phases $\phi_i = -k|\boldsymbol{r}_i|$ were set for each source located at $\boldsymbol{r}_i$ to focus on O. The net acoustic radiation force that a concave surface receives is

$$\boldsymbol{F} = (F_x, F_y, F_z) = \int_S P_{\text{rad}}(\boldsymbol{r})\boldsymbol{n}(\boldsymbol{r})dS \tag{3}$$

with the inward normal vector $\boldsymbol{n}(\boldsymbol{r})$ on the concave surface S. Note that the reflections of ultrasound on the concave surface were not considered. Using the longitudinal direction of the aperture of the AUPAs as the displacement direction (x-axis), calculations were performed in the range $|x| \leq 6$ mm.

Fig. 7 shows the magnitude of the x-directional component of the force on the object F_x when the displacement of the object is x. F_x is normalized by the magnitude of the acoustic radiation force F (which has a component in the three-dimensional direction) acting on the concave surface. Regardless of whether the surface is semi-ellipsoidal or conical, the forces are balanced at the origin and directed toward the origin as the displacement increases. The restoring force peaked at approximately $x = \pm 3$ mm for all shapes, and within $|x| \leq 2$ mm, it was similar to that of a linear spring. As x approached ± 6 mm, which corresponds to r, the force decreased as a part of the focal point protrusion outside the concave bottom.

IV. EXPERIMENTAL RESULTS

A. FORCE MEASUREMENT USING A LOAD CELL

The force acting along the displacement direction was measured using a load cell (LTS-50GA, Kyowa Sangyo) on objects with semi-ellipsoidal or conical concave bottoms. The parameters were $r = 6$ mm and $h = 4$ mm. The displacement of the object from the origin, that is, the focal point, was varied at intervals of 0.25 mm. The results are presented in Fig. 8, where x is the displacement and F_x is the x-directional force. Although there was a slight negative offset, the peak positions for both concave bottoms were approximately $x = \pm 3$ mm, which was similar to the numerical calculation. The force between the peaks is almost linear. Unlike in the numerical calculations, there was no difference in the

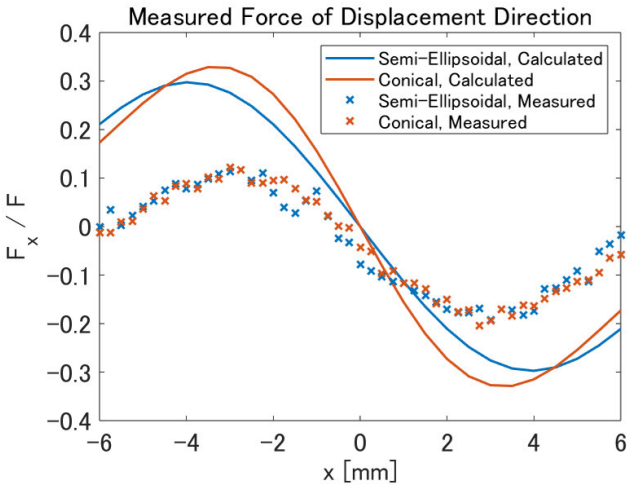

FIGURE 8. A load cell was connected to an object with a concave bottom to measure the force in the displacement direction relative to the focal point. Semi-ellipsoidal and conical surfaces with $h = 4$ mm were used. The measured values were normalized by the magnitude of the force near the focal point (13 mN). The solid line shows the numerical results.

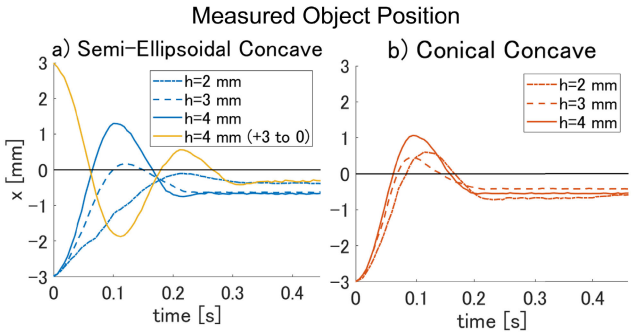

FIGURE 9. Time variations of the positions of the object initially placed at $x = -3$ mm as it moved to the origin (focal point). Measurements were obtained at different h when the bottom of the objects were a) semi-ellipsoidal concave and b) conical concave. For the condition $h = 4$ mm in a), we have also included the case where $x = +3$ mm is the initial position and moved to the origin.

force between the semi-ellipsoidal and conical bottoms. However, forces toward the focal point were obtained within $|x| \leq 6$ mm in both cases. Therefore, this force can be used to control objects with semi-ellipsoidal and conical concave bottoms in two dimensions.

B. TIME RESPONSE CHARACTERISTICS OF OBJECTS

The response times of the objects from resting to moving toward the focal point were measured. This allows for an approximate estimate of the feasible moving speeds of the objects.

In the experiment, the object was placed on the screen at $x = -3$ mm with respect to the origin. When an ultrasound focal point is generated at the origin, the object begins to move toward the origin. The time variation in the position of the object until it stabilized was measured using a high-speed camera (BU238MCF, Toshiba Teli) with a resolution of 1920 px × 1200 px and a frame rate of 165 fps. Objects with $r = 6$ mm and concave bottoms were produced using an SLA 3D printer (Form 3+, Formlabs) with Grey V4 resin (1.14 g/cm^3, Formlabs) to ensure uniform weights within the range of 782.5 ± 7.5 mg. The acoustic pressure amplitude of the ultrasound output from the four AUTD3 units was constant for all the objects. The amplitude was suppressed from the maximum amplitude to prevent the object from flipping over during the actuation. The total force near the focal point was 10 mN at this amplitude, whereas it is 13 mN at the maximum amplitude.

The results are shown in Fig. 9. The times taken to pass the target position for the first time were 0.10 and 0.067 s for the semi-ellipsoidal concave surface with $h = 3$ and 4 mm, and 0.085, 0.073, and 0.067 s for the conical concave surface with $h = 2$, 3, and 4 mm, respectively. The semi-ellipsoidal concave surface with $h = 2$ mm did not reach the target position, but the time taken to get closest was 0.22 s. Subsequently, the positions of the objects converged near the origin, where the center of the focal point was at approximately 0.2 s. For $h = 4$ mm, there was no significant difference between the semi-ellipsoidal and conical-concave shapes. However, for the semi-ellipsoidal concave structure, the response time increases significantly as h decreases. To observe the deviation from the target position after convergence, the root-mean-square errors with respect to the origin were calculated between 0.33 and 0.46 s. The maximum error was 0.46 mm with $h = 4$ mm for the semi-ellipsoidal concave surface and 0.47 mm with $h = 2$ mm for the conical concave surface.

C. TWO-DIMENSIONAL TRAJECTORY OF THE OBJECTS

1) SINGLE OBJECT

Experiments were conducted to move an object along a specific trajectory to confirm its deviation from the desired trajectory. The object to be moved had a semi-ellipsoidal concave bottom with $r = 6$ mm and $h = 4$ mm and weighed approximately 782.5 mg. We implemented a program to move an object on a screen using three types of two-dimensional trajectories: circular, square, and star-shaped. The path lengths were 188, 240, and 285 mm. For the circular trajectory, the step width of the focus update was constant along the trajectory. For the square trajectory, the step width was varied such that the rotation angle per unit time relative to the center was constant. In other words, the step width was largest at the corners. For the star-shaped trajectory with a long path length and acute angles at its apexes, the step width was determined to be small at the corners to avoid deviating from the trajectory when changing the direction. The focal point was controlled to move in a counterclockwise direction such that the time required for one round was 10 s. The focal point was updated every 50 ms, corresponding to the movement times of 0.94, 1.2, and 1.43 mm on average.

Initially, the object was placed on a path. The position of the moving object was measured for 1 min from the

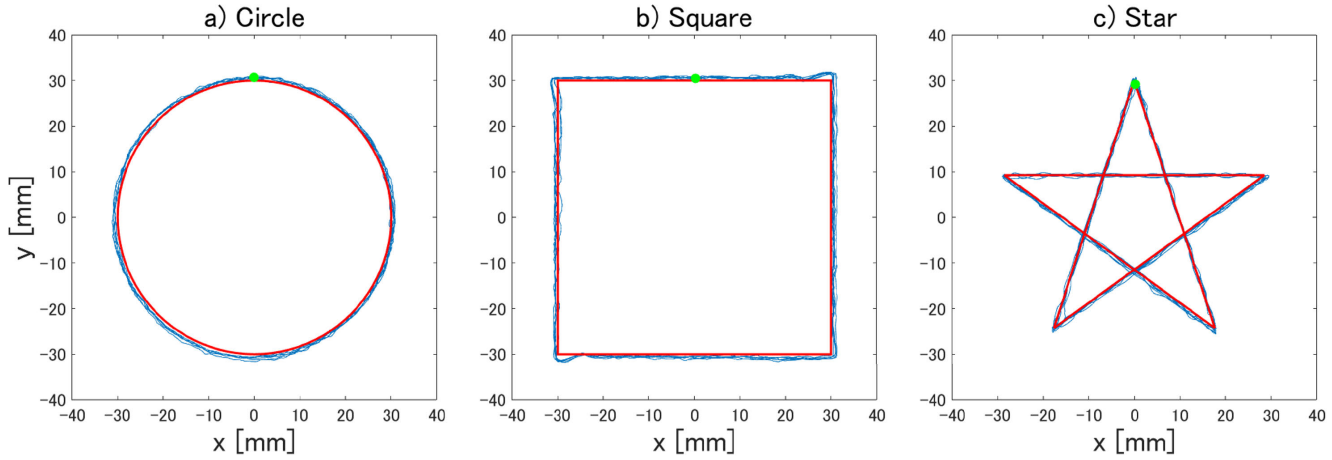

FIGURE 10. Trajectories of the single object actuated by the proposed method. The target trajectory is shown in red and the actual trajectory in blue. The object was observed until the end of the six rounds, each lasting 10 s. The speeds averaged over the path length are a) 18.8 mm/s, b) 24 mm/s, and c) 28.5 mm/s, respectively.

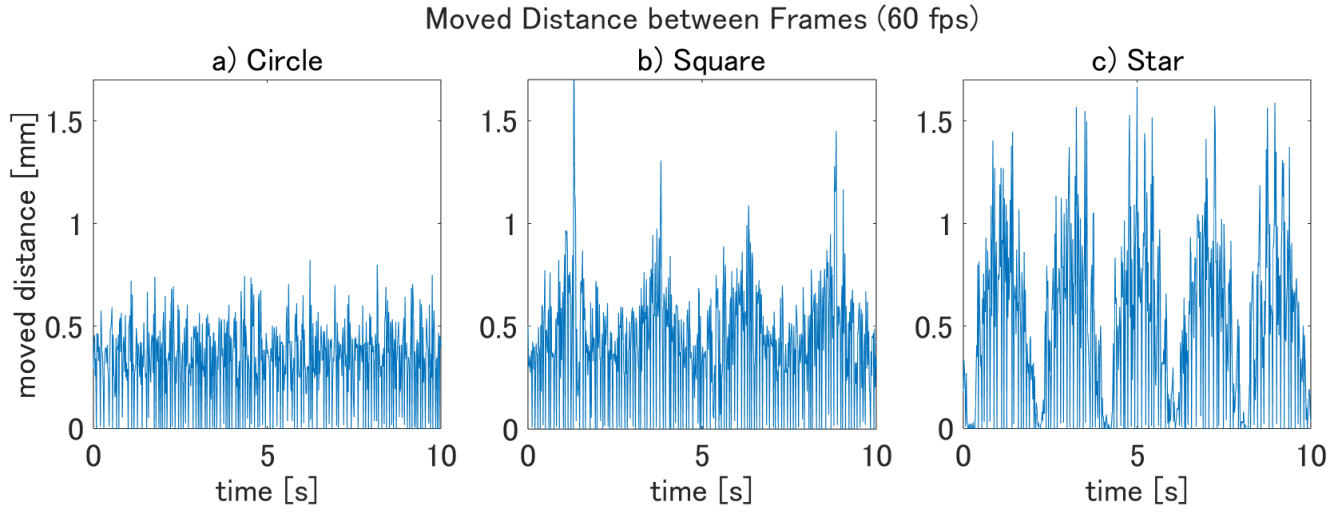

FIGURE 11. The distance an object moved between frames (1/60 s = 16.7 ms). The results are plotted from the time the object passed the green point in Fig. 10 for the first time (0 s) to the time moving one round (10 s). In a), no large peaks are observed because the step width of the focus update is constant; in b) and c), periodic peaks are observed because the step width was set empirically on the path.

beginning of the movement to the end of the sixth round using a 1920 px × 1200 px camera at 60 fps. The results are shown in Fig. 10. The object continued moving with no significant deviation from its prescribed trajectory. For a square trajectory, the inertia of the object caused a maximum deviation of 2 mm in the direction of travel toward the corners.

For each trajectory, the distance the object moved between frames ($\sqrt{\Delta x^2 + \Delta y^2}$) for a given 10 s is shown in Fig. 11. Because the video was captured at 60 fps, the time between frames was 16.7 ms. In all trajectories, the object was moving with finely repeated acceleration and deceleration. In a) the circular trajectory, the object moved at an approximately constant speed because the step width was constant. However, it is not perfectly constant because of the overshooting, as shown in Fig. 9. In b) the square and c) the star-shaped trajectories, the step width was varied along the path so that the moved distance peaked at the position with the largest step width. This peak appears at the corners in b) and at the center of the line segments in c).

2) MULTIPLE OBJECTS

Two or more objects can be actuated almost simultaneously using the time-division control of a focal point. Compared with methods that form multiple focal points simultaneously, this method has the advantage that the acoustic pressure per focal point is not reduced and the positional relationship between the objects is not affected in terms of acoustic field formation. Therefore, each object can move independently if they do not collide.

The trajectories of the two objects are shown in Fig. 12. The two objects had the same shape as those used in the

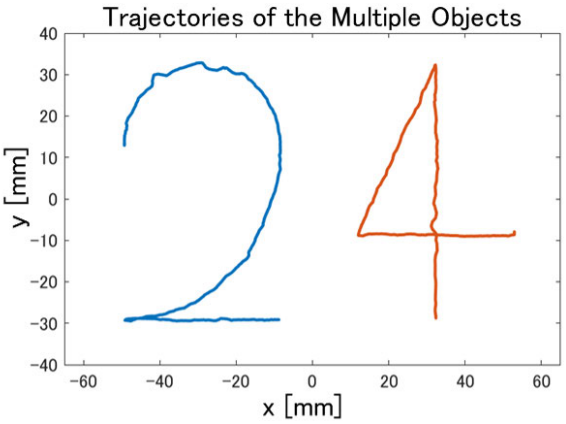

FIGURE 12. Trajectories of two objects simultaneously actuated by the proposed method. Each object was moved once from the starting point, which is the edge point with the larger y-coordinate to the endpoint, with a duration of 10 s. The left object moved along a trajectory shaped as "2" and the right object moved along a trajectory shaped as "4".

single-object experiment. Similar to a single object, the focal point was moved along a predefined trajectory and the object was observed using a camera. Each trajectory was a single stroke, moving only once; therefore, 10 s were required from the start to the end point. The edge point with the larger y-coordinate value was the starting point, and the edge point with the smaller y-coordinate value was the end point. The objects traced predefined trajectories; however, their positional deviations were greater than those of a single object.

V. DISCUSSION

A. DISCUSSION OF EXPERIMENTAL RESULTS

The force measurements, shown in Fig. 8, were lower than the numerical calculations for both the shapes. Additionally, there were no differences between the shapes, which may be due to the fact that the numerical calculations did not include the small forces acting on the sides of the object, or due to mounting errors of the object on the load cell.

Negative offsets were observed for the measured forces, as shown in Fig. 8. The object convergence positions shown in Fig. 9 were biased in the negative direction. Taking into account this offset, all the objects were considered to have converged on the target position. This offset may be due to a bias in the three-dimensional shape of the ultrasound focal point caused by individual differences in the AUPAs. Within the scope of this study, the error in the object position caused by this negative offset (less than 0.5 mm) is not considered to have a significant effect on the movement of objects on the order of several tens of millimeters, as shown in Fig. 10.

The results in Fig. 9 show that the time response characteristics differ depending on the concave shape and h. Each object responds like a damped vibration. Considering only the beginning of the motion ($x \leq 0$, $\dot{x} \geq 0$) for simplicity, the equation of motion of the object is given by:

$$m\ddot{x} = F_x(x) - \mu(mg - F_z(x)) \tag{4}$$

where m is the mass of the object, μ is the kinetic friction coefficient between the object and screen, and g is the gravitational acceleration. First, for both semi-ellipsoidal and conical surfaces, the rise time is longer for a smaller h. A smaller h decreases F_x instead of increasing F_z. The decrease in F_x has a larger effect because μ is small on the mesh screen, so that time response is considered to be delayed. In terms of response speed, h should be large. Because the focusing area of the focal point is approximately the diameter of the wavelength, $h = 4$ mm would be optimal at 40 kHz. Second, this delay is slighter for conical surfaces than for semi-ellipsoidal surfaces. The vertical component of the normal vector on surface $\boldsymbol{n}_z(\boldsymbol{r})$ is independent of the position of the conical surfaces, whereas it is large near the center and small at the periphery for semi-ellipsoidal surfaces. Under the measurement conditions of this study ($x = -3$ mm to 0), the characteristics of the conical surfaces were likely to be advantageous in reducing friction, resulting in a smaller delay.

Fig. 9 shows that the time taken to move 3 mm varies depending on the bottom shape of the object. Simple calculations suggest that an object with a semi-ellipsoidal concave bottom with $h = 2$, 3, and 4 mm can move at 14, 30, and 45 mm/s on average, whereas an object with a conical concave bottom with $h = 2$, 3, and 4 mm can move at 35, 41, and 45 mm/s on average, respectively. Moreover, Fig. 10 confirms that an object with a semi-ellipsoidal concave bottom with $h = 4$ mm can move stably at average speeds of 18.8, 24.0, and 28.5 mm/s over the path length in circular, square, and star-shaped trajectories, respectively. In Fig. 11 c), the object moved approximately 1.5 mm in one frame at peaks. This suggests that if an acceleration and deceleration interval is provided, stable operation can be achieved even if the object momentarily reaches 90 mm/s. The maximum achievable object speed is expected to vary with the trajectory, considering the acceleration of an object along a straight trajectory. When multiple objects are moved, as in the experiment shown in Fig. 12, this method uses a time-division method. This implies that multiple focal points were not formed simultaneously. Therefore, the position-update rate for each object decreases as the number of objects to be actuated increases. Referring to the experimental results of moving 3 mm in 0.067 s for one object (45 mm/s), it was theoretically calculated that the objects could move at 22.5 mm/s each for two objects and at 4.5 mm/s each for 10 objects.

B. LIMITATIONS

The weight of the object showing the response characteristics in Fig. 9 is limited by the upper output limit of the actual equipment. Referring to Eq. (4), $F_x/|F|$ and $F_z/|F|$ are determined by the distribution of the normal vectors of the concave surface, that is, its geometry and the acoustic pressure distribution. If the intensity of the acoustic radiation force $|F|$ can be arbitrarily adjusted to maintain a constant

$|F|/mg$, then the propulsive force will be independent of the weight. However, because the actual equipment has an upper limit on the force presented, the maximum weight of the objects that exhibit equivalent response characteristics is determined by the force limit. In the experiments, the acoustic radiation force near the focal point was 10 mN when moving an object approximately 782.5 mg. With four AUTD3 units, the maximum acoustic radiation force near the focal point at H_{screen} was 13 mN, so that the response characteristics for each shape are expected to be same to Fig. 9 up to 1.0 g. For heavier objects, the response is delayed due to the magnitude of inertia and frictional resistance. If $F_x(x)$ and $F_z(x)$ remain the same as in Eq. (4), and m increases to M, $\ddot{x} + \mu g$ decreases by a factor of m/M, resulting in a lower acceleration. Moreover, when

$$\mu_s Mg \geq F_x(x) + \mu_s F_z(x), \tag{5}$$

the object does not start moving, where μ_s is the static friction coefficient between the object and the screen. By contrast, when the weight is reduced, the presenting force must be reduced to satisfy

$$F_z(x) \leq mg \tag{6}$$

for every x so that the object does not move upward, and thus, greater acceleration is not obtained. The maximum speed at which the objects can be achieved depends on more advanced operational strategies, such as the focus-update rate, step width, trajectory design, or the introduction of feedback control. This is beyond the scope of this study and will be an issue to be investigated in the future.

In our configuration, the direction in which the objects can move depends on their positions on the screen. The further an object moves toward the outer edge of the screen, the fewer the transducers outside it. Therefore, objects positioned at the edges of the screen are, in principle, less subjected to force toward their centers. We confirmed that the object could move in any direction within an area of approximately 20 cm $\times$ 15 cm around the center of our setup. This area can be enlarged by widening the aperture of the AUPAs or using a reflector on the side. In the latter approach, the ultrasound reflector surrounds the sides of the space between the AUPAs and screen. Consequently, objects situated far from the center are subjected to forces directed toward the center through the reflectors. Therefore, the effectiveness of the method using reflectors must be verified separately.

The target objects rotated around the vertical axis during the translation in the current environment. Because the objects in this method have symmetrical shapes around the vertical axis, rotation is currently uncontrollable. This rotation is most likely due to the nonuniformity of the acoustic field and friction between the bottom of the targets and the screen. The rotation can be suppressed by lowering the focus-update step width and the update frequency. This is because the object does not rotate when stationary. Another possible method is to create multiple concave points on an object and move it while maintaining a constant posture with multiple focal points. Multiple focal points can be formed in a time-division manner, as used in the experiment. This method is expected to suppress the rotation and utilize the intended rotation.

The proposed method requires an acoustically transparent screen. When mesh screens are used, lowering the aperture ratio increases the load-bearing capacity, but decreases the transmission of sound waves. This renders the current layout unsuitable for situations in which heavy objects are placed or hand pressure is applied to the screen.

VI. CONCLUSION

In this paper, we present a method for actuating objects 1–2 cm in size using airborne ultrasound to control their two-dimensional position, regardless of their materials. We used AUPAs as the sound sources, an acoustically transparent screen as the workspace, and objects with concave bottoms as the actuated objects. The findings of this study are as follows. i) The distance between the four AUTD3 units and the screen should be less than 19 cm, considering the influence of acoustic flow. ii) Objects with a concave semi-ellipsoidal or conical bottom, with a bottom opening radius of 6 mm and a central depth of 4 mm, can extract the horizontal driving force from the acoustic radiation force. iii) In the experiments in this study, the object's response time was the fastest when moving a semi-ellipsoidal concave with a central depth of 4 mm. iv) Our study confirmed that it is possible to continuously actuate a single object and operate multiple objects simultaneously using the time-division method. In the experiments, an object weighing approximately 782.5 mg was moved at an average speed of 28.5 mm/s over its path length. The limitations of the current system include the fact that the weight of the object to be manipulated should be less than 2 g and that the rotation of an object during movement cannot be controlled. The proposed method does not require any specific physical properties, so that it can remotely actuate edible or entirely transparent objects.

REFERENCES

[1] M. Rubenstein, C. Ahler, and R. Nagpal, "Kilobot: A low cost scalable robot system for collective behaviors," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 3293–3298.

[2] M. Le Goc, L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer, "Zooids: Building blocks for swarm user interfaces," in *Proc. 29th Annu. Symp. User Interface Softw. Technol.* New York, NY, USA: Association for Computing Machinery, Oct. 2016, pp. 97–109.

[3] P. Bhushan and C. Tomlin, "An insect-scale self-sufficient rolling microrobot," *IEEE Robot. Autom. Lett.*, vol. 5, no. 1, pp. 167–172, Jan. 2020.

[4] V. K. Venkiteswaran, L. F. P. Samaniego, J. Sikorski, and S. Misra, "Bio-inspired terrestrial motion of magnetic soft millirobots," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1753–1759, Apr. 2019.

[5] L. Hines, K. Petersen, G. Z. Lum, and M. Sitti, "Soft actuators for small-scale robotics," *Adv. Mater.*, vol. 29, no. 13, Apr. 2017, Art. no. 1603483.

[6] W. Hu, G. Z. Lum, M. Mastrangeli, and M. Sitti, "Small-scale soft-bodied robot with multimodal locomotion," *Nature*, vol. 554, no. 7690, pp. 81–85, Feb. 2018.

[7] G. Z. Lum, Z. Ye, X. Dong, H. Marvi, O. Erin, W. Hu, and M. Sitti, "Shape-programmable magnetic soft matter," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 41, pp. E6007–E6015, 2016.

[8] A. Bhattacharjee, L. W. Rogowski, X. Zhang, and M. J. Kim, "Untethered soft millirobot with magnetic actuation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 3792–3798.
[9] C. Wang, V. R. Puranam, S. Misra, and V. K. Venkiteswaran, "A snake-inspired multi-segmented magnetic soft robot towards medical applications," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 5795–5802, Apr. 2022.
[10] C. J. Cai, X. Xiao, M. S. Kalairaj, I. J. J. Lee, A. K. Mugilvannan, B. S. Yeow, J. H. Tan, H. Huang, and H. Ren, "Diversified and untethered motion generation via crease patterning from magnetically actuated caterpillar-inspired origami robot," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1678–1688, Jun. 2021.
[11] L. Li, Z. Shang, J. Tang, J. Li, Z. Xiao, Y. Xie, J. Liu, Z. Li, S. Yang, Y. Liu, and W. Gan, "Wood robot with magnetic anisotropy for programmable locomotion," *Adv. Funct. Mater.*, vol. 33, no. 6, Feb. 2023, Art. no. 2207209.
[12] Y. Ochiai, T. Hoshi, and J. Rekimoto, "Pixie dust: Graphics generated by levitated and animated objects in computational acoustic-potential field," *ACM Trans. Graph.*, vol. 33, pp. 1–13, Jul. 2014.
[13] R. Morales, A. Marzo, S. Subramanian, and D. Martínez, "LeviProps: Animating levitated optimized fabric structures using holographic acoustic tweezers," in *Proc. 32nd Annu. ACM Symp. User Interface Softw. Technol.* New York, NY, USA: Association for Computing Machinery, Oct. 2019, pp. 651–661.
[14] A. Marzo, S. A. Seah, B. W. Drinkwater, D. R. Sahoo, B. Long, and S. Subramanian, "Holographic acoustic elements for manipulation of levitated objects," *Nature Commun.*, vol. 6, no. 1, p. 8661, Oct. 2015.
[15] V. Paneva, A. Fleig, D. M. Plasencia, T. Faulwasser, and J. Müller, "OptiTrap: Optimal trap trajectories for acoustic levitation displays," *ACM Trans. Graph.*, vol. 41, no. 5, pp. 1–14, Oct. 2022.
[16] T. Fushimi, A. Marzo, T. L. Hill, and B. W. Drinkwater, "Trajectory optimization of levitated particles in mid-air ultrasonic standing wave levitators," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2018, pp. 1–9.
[17] C. T. Vi, A. Marzo, D. Ablart, G. Memoli, S. Subramanian, B. Drinkwater, and M. Obrist, "TastyFloats: A contactless food delivery system," in *Proc. ACM Int. Conf. Interact. Surf. Spaces*. New York, NY, USA: Association for Computing Machinery, Oct. 2017, pp. 161–170.
[18] W. J. Xie, C. D. Cao, Y. J. Lü, Z. Y. Hong, and B. Wei, "Acoustic method for levitation of small living animals," *Appl. Phys. Lett.*, vol. 89, no. 21, Nov. 2006, Art. no. 214102.
[19] T. Furumoto, M. Fujiwara, Y. Makino, and H. Shinoda, "Bang-bang control with constant thrust of a spherical blimp propelled by ultrasound beam," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 4051–4057.
[20] S. Inoue, S. Mogami, T. Ichiyama, A. Noda, Y. Makino, and H. Shinoda, "Acoustical boundary hologram for macroscopic rigid-body levitation," *J. Acoust. Soc. Amer.*, vol. 145, no. 1, pp. 328–337, Jan. 2019.
[21] Y. Koroyasu, T.-V. Nguyen, S. Sasaguri, A. Marzo, I. Ezcurdia, Y. Nagata, T. Yamamoto, N. Nomura, T. Hoshi, Y. Ochiai, and T. Fushimi, "Microfluidic platform using focused ultrasound passing through hydrophobic meshes with jump availability," *PNAS Nexus*, vol. 2, no. 7, Jun. 2023, Art. no. pgad207.
[22] M. Marshall, T. Carter, J. Alexander, and S. Subramanian, "Ultra-tangibles: Creating movable tangible objects on interactive tables," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, May 2012, pp. 2185–2188.
[23] S. Suzuki, S. Inoue, M. Fujiwara, Y. Makino, and H. Shinoda, "AUTD3: Scalable airborne ultrasound tactile display," *IEEE Trans. Haptics*, vol. 14, no. 4, pp. 740–749, Oct. 2021.

GAKUTO ARAKAWA received the B.S. and M.S. degrees in information science and technology from The University of Tokyo, in 2018 and 2021, respectively, where he is currently pursuing the Ph.D. degree with the Department of Complexity Science and Engineering. His research interests include acoustic manipulation, human–computer interaction, and mid-air haptics.

RYUNOSUKE HIRAI received the B.S. degree in information science and technology from The University of Tokyo, in 2023, where he is currently pursuing the M.S. degree with the Department of Information Physics and Computing. His research interests include acoustic levitation, information physics, and application systems related to them.

SHUN SUZUKI received the B.S. degree in physics from Waseda University, Japan, in 2017, and the M.S. and Ph.D. degrees in complexity science and engineering from The University of Tokyo, Japan, in 2019 and 2022, respectively. He is currently a Project Assistant Professor with The University of Tokyo. His research interest includes mid-air haptics.

TAKAAKI KAMIGAKI received the B.S. and M.S. degrees in engineering from Kumamoto University, in 2014 and 2016, respectively, and the Ph.D. degree from The University of Tokyo, Japan, in 2019. He is currently a Project Assistant Professor with The University of Tokyo. His research interest includes mid-air haptics.

YASUTOSHI MAKINO received the bachelor's, master's, and Ph.D. degrees in information science and technology from The University of Tokyo, in 2002, 2004, and 2007, respectively. From 2009 to 2013, he was a Researcher with The University of Tokyo for two years and an Assistant Professor with Keio University. In 2013, he moved to The University of Tokyo, as a Lecturer, and has been an Associate Professor, since 2017. He is currently an Associate Professor with the Department of Complexity Science and Engineering, The University of Tokyo. His research interests include haptic interactive systems and human motion prediction using skeletal information.

HIROYUKI SHINODA (Member, IEEE) received the bachelor's degree in applied physics, the master's degree in information physics, and the Ph.D. degree in engineering from The University of Tokyo, in 1988, 1990, and 1995, respectively. Since 1995, he has been an Associate Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Japan. After a period at UC Berkeley as a Visiting Scholar, in 1999, he became an Associate Professor with The University of Tokyo, from 2000 to 2012. He is currently a Professor with the Graduate School of Frontier Sciences, The University of Tokyo. His research interests include information physics, haptics, mid-air haptics, two-dimensional communication, and their applications. He is also a member of IEEJ, RSJ, JSME, and ACM.