Rendering Ultrasound Pressure Distribution on Hand Surface in Real-Time

Atsushi Matsubayashi[✉], Yasutoshi Makino, and Hiroyuki Shinoda

The University of Tokyo, Tokyo, Japan
Matsubayashi@hapis.k.u-tokyo.ac.jp

Abstract. In this paper, we propose a method for rendering the pressure distribution on the skin surface of a hand in real-time using an ultrasonic phased array. Our method generates a polygon mesh model representing the hand shape by fitting a rigid template to the point cloud captured by depth sensors. Obtaining the entire hand shape as a mesh model enables to solve scattering problem to generate a precise distribution on the hand surface. Therefor, for example, our method can control the width of the distribution on the fingertip according to the size of the contact area with the virtual object. We have experimentally verified that considering the scattering on the mesh model contributes to accurate pressure pattern reproduction.

Keywords: Mid-air haptics · Scattering problem · Airborne ultrasound

1 Introduction

Ultrasound haptics is a technology to generate a tactile sensation on a human skin by creating a point with high sound pressure using an ultrasound phased array. Since a tactile presentation by airborne ultrasound was first demonstrated in 2008 [9], many attempts have been made to create haptic images in the air using this technology. In order to create a pressure distribution with the desired shape in the air, many methods solve the inverse problem using the relationship between the sound pressure of control points and the complex amplitudes of the transducers [4,6,10]. Since the intensity of a stimulus felt on the skin depends only on the amplitude of the sound pressure, the phase of the target sound pressure distribution can be set arbitrarily in the inverse problem. Properly setting this phase by solving eigenproblem [10] or phase retrieval problem [6] widens the range of reproducible haptic image. Some method, furthermore, present a stronger stimulus by tracking the hand and generating a pressure distribution only on the contact area of the hand touching the image, allowing the user to identify the shape of the haptic image more clearly [10,11]. However, since these above methods do not consider scattering on the hand surface, the pressure
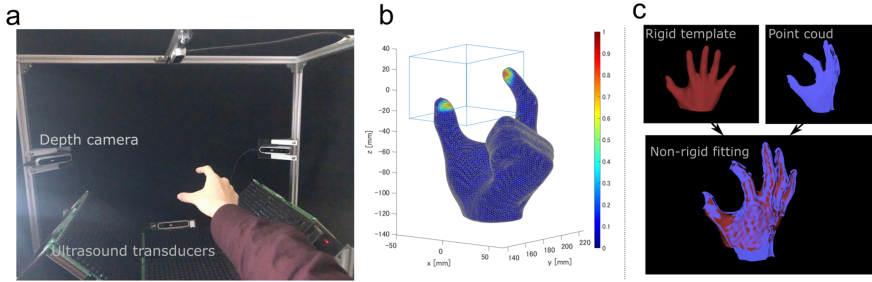
Fig. 1. a, b) The phase of the ultrasound transducers is determined according to the hand mesh model generated from the depth information acquired by depth cameras, and the desired pressure distribution is generated at the fingertip. c) By fitting the template rigid model to the point cloud, a non-rigid mesh model is dynamically generated in real-time.

distribution actually generated on the skin surface differs from target pressure distribution to be reproduced. Solving the scattering problem will enable more accurate reconstruction of the pressure pattern. Inoue et al. have proposed a method for generating a stronger focal point by considering the scattering on a polygon mesh model of a finger [7]. They demonstrated that a stronger ultrasound focus could be created by solving a scattering problem using a static mesh model, but it has not been possible to control the pressure distribution on the mesh model dynamically generated in real-time.

In this paper, we propose a method to render the pressure distribution in real-time on the polygon mesh model deforming according to the hand shape. In this method, the shape of the hand placed above a transducer array is aquired with multiple cameras as shown in Fig. 1 (a), and the hand polygon mesh model deforms non-rigidly to fit this shape. The scattering model formulates the relationship between a sound pressure pattern on the mesh model and phases of the ultrasound transducers in the style of the boundary element method. Based on this relationship, our method optimize both phases of transducers and phases of target pressure distribution to generate the desired pattern at any position on the hand surface. For example, as shown in Fig. 1 (b) our method can generate a distribution according to the size of the contacting region with a virtual object.

2   Method

2.1   Generation of Mesh Model

In order to dynamically generate a polygon mesh model of the hand, we used a mesh reconstruction technique similar to that proposed by Zollhöfer et al. [12]. Mesh reconstruction process consists of two phases. First, a rigid template is created by scanning the hand with a fixed form using multiple depth cameras. Then, as shown in Fig. 1 (c), a non-rigid mesh model is generated by fitting the rigid template to the point cloud obtained from the depth cameras. We used the

fitting method proposed by Dou et al. [5]. which has the advantage of being able to generate a mesh model that closely matches the skin surface of the actual hand compared to skeletal hand tracking methods used, for instance, in Leap Motion [2]. This is an important property for controlling the pressure distribution on the hand surface.

2.2   Scattering Model of Hand Surface

The relationship between the sound pressure on the faces of the mesh model and the phases of the ultrasound transducers is formulated in style of the boundary element method, similar to Inoue's adaptive focusing method [7]. The sound pressure $p(\boldsymbol{r}) \in \mathbb{C}$ scattered on the smooth surface Ω of a sound-hard rigid body is given by the following boundary integral Eq. [3].

$$\frac{1}{2}p(\boldsymbol{r}) = p_{inc}(\boldsymbol{r}) - \int_{\Omega} p(\boldsymbol{r})\frac{\partial g(\boldsymbol{r}, \boldsymbol{s})}{\partial \boldsymbol{n}}dS, \tag{1}$$

where g is the Helmholtz green function, and p_{inc} is the incident wave from the transducers. In our method, this is simplified by a spherical wave with directivity D_n as

$$p_{inc}(\boldsymbol{r}) = \sum_{n} D_n(\boldsymbol{r})\frac{e^{-jk\|\boldsymbol{r}-\boldsymbol{x}_n\|}}{\|\boldsymbol{r} - \boldsymbol{x}_n\|}a_n e^{\phi_n}, \tag{2}$$

where $a_n \in \mathbb{R}, \phi_n \in \mathbb{R}$ and $\boldsymbol{x}_n \in \mathbb{R}^3$ are the amplitude, phase and position of a transducer $n \in \{1, \cdots N\}$ respectively. To reduce computational cost of the optimization, we set the amplitude constant, i.e. $a_n = a$.

When the boundary surface is represented by a polygon mesh and the sound pressure and the gradient of the Green's function are approximated to be constant on each face of the mesh, the boundary integral equation (1) is discretized as follows:

$$B\left(p_1, \cdots, p_M\right)^{\mathrm{T}} = G\left(e^{\phi_1}, \cdots, e^{\phi_N}\right)^{\mathrm{T}}, \tag{3}$$

where

$$B = \begin{pmatrix} \frac{\partial g}{\partial n}(\boldsymbol{y}_1, \boldsymbol{y}_1)A_1 + \frac{1}{2} & \cdots & \frac{\partial g}{\partial n}(\boldsymbol{y}_1, \boldsymbol{y}_M)A_M \\ \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial n}(\boldsymbol{y}_M, \boldsymbol{y}_1)A_1 & \cdots & \frac{\partial g}{\partial n}(\boldsymbol{y}_M, \boldsymbol{y}_M)A_M + \frac{1}{2} \end{pmatrix}, \tag{4}$$

$$G = \begin{pmatrix} D_1(\boldsymbol{y}_1)\frac{ae^{-jk\|\boldsymbol{y}_1-\boldsymbol{x}_1\|}}{\|\boldsymbol{y}_1-\boldsymbol{x}_1\|} & \cdots & D_N(\boldsymbol{y}_1)\frac{ae^{-jk\|\boldsymbol{y}_1-\boldsymbol{x}_N\|}}{\|\boldsymbol{y}_1-\boldsymbol{x}_N\|} \\ \vdots & \ddots & \vdots \\ D_1(\boldsymbol{y}_M)\frac{ae^{-jk\|\boldsymbol{y}_M-\boldsymbol{x}_1\|}}{\|\boldsymbol{y}_M-\boldsymbol{x}_1\|} & \cdots & D_N(\boldsymbol{y}_M)\frac{ae^{-jk\|\boldsymbol{y}_M-\boldsymbol{x}_N\|}}{\|\boldsymbol{y}_M-\boldsymbol{x}_N\|} \end{pmatrix}, \tag{5}$$

and $p_m \in \mathbb{C}, \boldsymbol{y}_m \in \mathbb{R}^3$ and $A_m \in \mathbb{R}$ is the sound pressure, position, and area of a face $m \in \{1, \cdots, M\}$ of the mesh model respectively.

2.3   Optimizing Phases of Ultrasound Transducers

Given the target sound pressure amplitude $\boldsymbol{p}' = (p_1', \cdots, p_M')^{\mathrm{T}} \in \mathbb{R}^M$, we want to determine the phases of the transducers $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_N)^{\mathrm{T}} \in \mathbb{R}^N$ and phases of the sound pressure $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_M)^{\mathrm{T}} \in \mathbb{R}^M$ that minimize the least square error $\| (p_1' e^{\theta_1}, \cdots, p_M' e^{\theta_M})^{\mathrm{T}} - B^{-1} G (e^{\phi_1}, \cdots, e^{\phi_N})^{\mathrm{T}} \|_2^2$. However, calculating the inverse of B is very computationally expensive, so we solve the following optimization problem instead (Fig. 2).

$$\min_{\boldsymbol{\phi}, \boldsymbol{\theta}} \| B (p_1' e^{\theta_1}, \cdots, p_M' e^{\theta_M})^{\mathrm{T}} - G (e^{\phi_1}, \cdots, e^{\phi_N})^{\mathrm{T}} \|_2^2. \qquad (6)$$

We solve this problem iteratively using the Levenberg-Marquardt method. At each iteration, parameters $\boldsymbol{t} = (\phi_1, \cdots, \phi_N, \theta_1, \cdots, \theta_M)^{\mathrm{T}}$ is updated as follows:

$$\boldsymbol{t} \leftarrow \boldsymbol{t} - (J^{\mathrm{T}} J + \lambda I)^{-1} J^{\mathrm{T}} \boldsymbol{f}, \qquad (7)$$

where the residual vector $\boldsymbol{f} = B (p_1' e^{\theta_1}, \cdots, p_M' e^{\theta_M})^{\mathrm{T}} - G (e^{\phi_1}, \cdots, e^{\phi_N})^{\mathrm{T}}$, and J is the Jacobian of $\boldsymbol{f}$. If the number of parameters $M + N$ is large, the time taken for an iteration will be very long, but when pressure is generated only in a local part such as a fingertip, excluding zero pressure faces can save computation time.
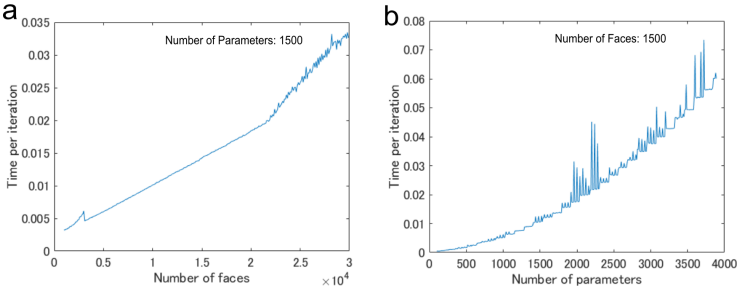

Fig. 2. a) The time per iteration against the number of faces. b) the time per iteration against the number of parameters

3   Implementation

The above algorithms were implemented with CUDA on two GeForce RTX 2080 Ti GPUs. One is used for mesh generation algorithm and the other is used for phase optimization. We measured the time taken for one iteration of the phase optimization in this environment. Figure 3 shows the time per iteration against the number of faces and parameters. The number of parameters is the sum of the number of transducers and the number of faces with non-zero pressure. In consideration of this result, possible resolution of the pressure distribution and
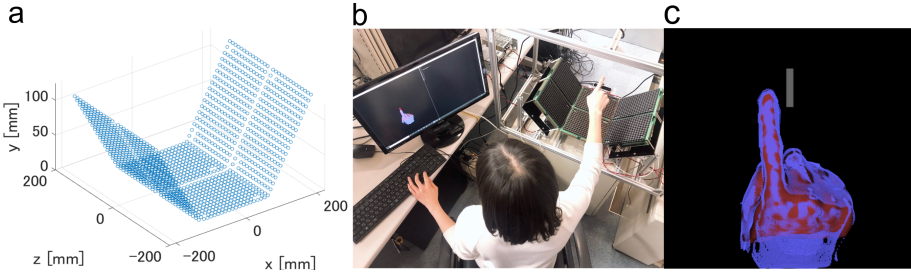
Fig. 3. a) The coordinate system and the arrangement of the ultrasound transducers in the experimental setup. b,c) The participants sat in front of the system and touched the box checking the position of the hand and box displayed on a LCD.

the limitations of the devices used, we set the number of faces to about 10,000 and the number of transducers to 1496. Therefore the time per iteration is about 10 ms, so we set the update frequency of the phase to 20 Hz with five iteration.

We constructed an experimental setup as shown in Fig. 1 (a). We installed Intel RealSense Depth Camera D415 [1] to measure the hand. The resolution of the depth image captured by each camera is 640 × 360, and the refresh rate as well as the update frequency of the mesh generation is 30 Hz. The architecture of the ultrasound transducer array unit is that proposed by Inoue et al [8]. The resonant frequency of the transducer is 40 kHz, and 200 Hz amplitude modulation is applied to make the tactile stimulus easier to perceive. Figure 3 (a) shows the coordinate system and arrangement of the ultrasound transducers in our setup.

4   Numerical Analysis

We performed numerical simulation to verify how close the distribution could be to the target in the experimental setup.

Figure 4 shows the simulation results of our method. In the target distribution (a1-a3), a constant pressure is applied to faces inside a box-shaped region. The width of the box is 5 mm, 8 mm and 11 mm in a1, a2 and a3 respectively. It can be seen that the distribution generated by our method (b1-b3) is close to the target and changes according to the width of the box. Figure 4 (c1-c3) shows the simulation result of the phase optimization performed without consideration of the scattering, which means replacing matrix B in our algorithm with the identity matrix. The result suggest an appropriate distribution cannot be generated without considering scattering. Our method can be applied to the case of touching with multiple fingers as well. The Fig. 4 (a4, b4 and c4) shows the simulation result of the case where a virtual box is grasped.
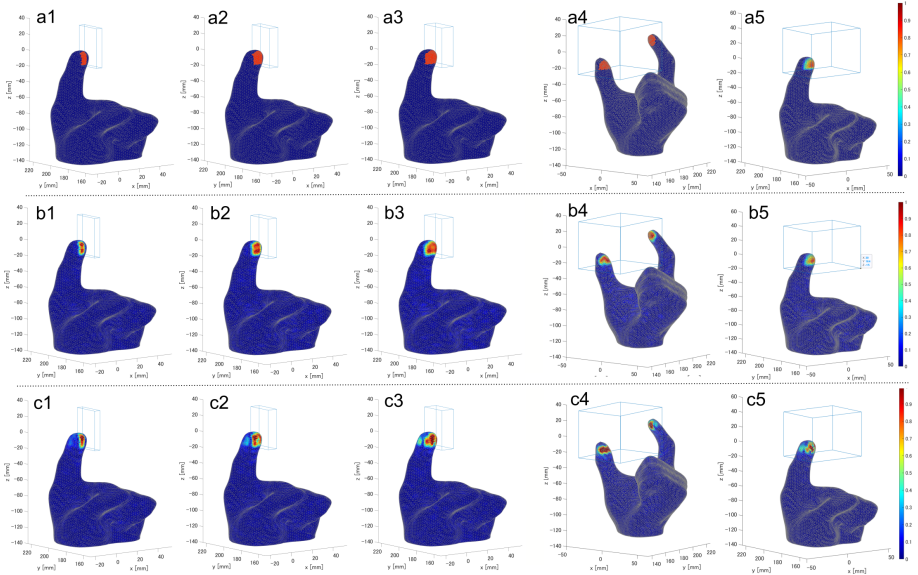
Fig. 4. Simulation results. a1-a5) Target pressure distribution. b1-b5) Distribution reproduced by our method. c1–c5) Distribution reproduced without scattering model.

When actually touching an object, the pressure on the contact area is not uniform, but greater toward the center. Figure 4 (a5, b5 and c5) shows the simulation result of a simplified model that a strong pressure is presented proportional to the penetration distance into the box. Although the shape is somewhat deformed, target distribution is reproduced. In these simulations, 30 iterations were performed in the phase optimization with the initial value as the zero vector. However, it has been empirically known that by setting the phases of the previous frame to the initial value, convergence can be sufficiently achieved in 5 to 10 iterations.

5 User Study

To verify if our method can generate a discernible difference in the pressure distribution, we conducted a user study. In this study, participants performed tasks of touching and identifying three types of distribution. We compared the accuracy of identification between the two methods. One is our method using the scattering model, and the other is the method without considering scattering.

Procedure. The participants sat in front of the system and placed his hand above the transducer array. For the stability of mesh generation, form of the hand was limited to only the index finger up throughout the experiment as shown in Fig. 3 (b). Then, the participant's hand was scanned and a rigid template was
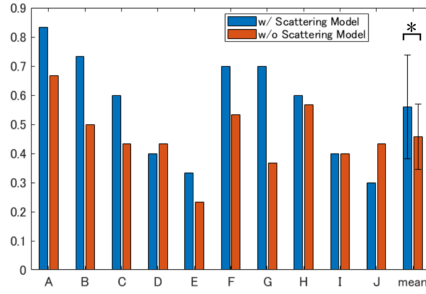
Fig. 5. Accuracy rates of the participants and the corresponding mean value.

created, which takes about 20 sec. After confirming that the mesh generation was working properly, the participants were asked to experience three different widths of pressure pattern for 15 sec each. As in the simulation, a uniform pressure is applied to the part that enters a box-shaped region. See the Fig. 4 (b1-b3) for the size of the region. The participants were not allowed to move their fingers horizontally to feel the width, but only to move vertically checking their hand and the box-shaped region displayed on a LCD as shown in the Fig. 3 (c). Then, the participants were asked to repeatedly perform the tasks to identify the width. After 15 s of touching, the tactile presentation was stopped, and participants answered one of three widths. Participants performed three task as a practice and then performed 10 tasks for each width (total 30 tasks). In either case, no answer was taught to the participants. The order of the tasks is randomized. The above process was done separately under two methods to avoid confusion between methods. Five of the ten participants performed the experiment with scattering model and the other five performed the experiment without scattering model.

Result and Discussion. Ten participants (eight males and two females), aged between 23 to 24, took part in the experiment. Figure 5 shows the accuracy rate of the participants. The participant A to E performed the experiment with scattering model first and the participant F to J performed the experiment without scattering model first. The mean value of the accuracy rates among participants was 0.56 with scattering model and 0.47 without scattering model, and the Wilcoxon signed-rank test yielded a significant difference ($p < 0.05$) between the two methods. The mean accuracy indicates that the differences in the distributions generated by our method are discernable to some extent, but not perfect. One of the reasons for this may be that an accurate mesh model could not be generated due to the error of the depth camera. However, the difference in accuracy between the two methods suggests that it is effective to consider scattering even in such a case. Also, in this experiment, we did not give instructions on the appropriate speed of touching. Since there is a delay between capturing the hand and presenting the tactile sensation, an appropriate

distribution cannot be presented for a fast movement of the finger. This may have led to a large differences among accuracy rates of participants. In particular, it is considered that participant E and J were greatly affected by the delay because their fingers shook during the experiment. We are required to verify how much delay there is and how it affects the result.

6   Conclusion

In this paper, we proposed and examined a method for rendering an ultrasound pressure distribution by solving the scattering problem. Although in the experiment, the pressure presentation was limited to the fingertip, our method can produce a pressure pattern on the entire hand surface, so there is still room for verification as to what kind of and how high the pressure distribution can be generated. At present, there is limitations on the temporal and spatial resolutions of the distribution due to computational cost, sensing accuracy, and transducer's resonant frequency. However, these problems will be solved with the advancement of the devices. One of our future work is the complete reproduction of the pressure distribution when touching a soft object using the presented approach.

References

1. Intel corporation (2018). https://www.intel.com/
2. Ultraleap ltd. (2020). https://www.ultraleap.com/
3. Bai, M.R., Ih, J.G., Benesty, J.: Acoustic Array Systems: Theory,implementation, and Application. Wiley, Hoboken (2013)
4. Carter, T., Seah, S.A., Long, B., Drinkwater, B., Subramanian, S.: UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In: Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, pp. 505–514. ACM (2013)
5. Dou, M., et al.: Fusion4D: real-time performance capture of challenging scenes. ACM Trans. Graph. (TOG) **35**(4), 1–13 (2016)
6. Inoue, S., Makino, Y., Shinoda, H.: Active touch perception produced by airborne ultrasonic haptic hologram. In: Proceedings of the World Haptics Conference, pp. 362–367. IEEE (2015)
7. Inoue, S., Makino, Y., Shinoda, H.: Mid-air ultrasonic pressure control on skin by adaptive focusing. In: Bello, F., Kajimoto, H., Visell, Y. (eds.) EuroHaptics 2016. LNCS, vol. 9774, pp. 68–77. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42321-0_7
8. Inoue, S., Makino, Y., Shinoda, H.: Scalable architecture for airborne ultrasound tactile display. In: Hasegawa, S., Konyo, M., Kyung, K.-U., Nojima, T., Kajimoto, H. (eds.) AsiaHaptics 2016. LNEE, vol. 432, pp. 99–103. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-4157-0_17
9. Iwamoto, T., Tatezono, M., Shinoda, H.: Non-contact method for producing tactile sensation using airborne ultrasound. In: Ferre, M. (ed.) EuroHaptics 2008. LNCS, vol. 5024, pp. 504–513. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-69057-3_64

10. Long, B., Seah, S.A., Carter, T., Subramanian, S.: Rendering volumetric haptic shapes in mid-air using ultrasound. ACM Trans. Graph. **33**(6), 181:1–181:10 (2014)
11. Matsubayashi, A., Oikawa, H., Mizutani, S., Makino, Y., Shinoda, H.: Display of haptic shape using ultrasound pressure distribution forming cross-sectional shape. In: 2019 IEEE World Haptics Conference (WHC), pp. 419–424. IEEE (2019)
12. Zollhöfer, M., et al.: Real-time non-rigid reconstruction using an RGB-D camera. ACM Trans. Graph. (ToG) **33**(4), 1–12 (2014)