Sound-Field Creation for Haptic Reproduction

Atsushi Matsubayashi, Seki Inoue, Shun Suzuki, and Hiroyuki Shinoda

Abstract In ultrasound haptics, a tactile sensation is evoked on human skin that touches the high sound pressure area generated due to the interference of ultrasonic waves. Therefore, by controlling the distribution of sound pressure amplitude using a transducer array, it is possible to create tactile sensations at multiple points simultaneously. In this chapter, we first describe the relationship between the complex gains of the transducer array and the generated sound field. Then, we provide various algorithms to control the amplitude pattern based on this relationship.

1 Acoustic Field Reconstruction

Ultrasound haptics is a technology that uses an array of ultrasound transducers to present tactile sensations remotely. The tactile sensation is caused by a nonlinear phenomenon called acoustic radiation. Assuming that the particle velocity of ultrasound is sufficiently small, the acoustic radiation pressure can be approximated as a quantity proportional to the square of the sound pressure p calculated in the range of linear acoustics (Hasegawa et al. 2000). Therefore, controlling p over certain control points leads to the design of a spatial tactile pattern that spreads over them. In this section, we first describe the forward problem of determining the sound pressure field p created by an ultrasound transducer array and then outline the inverse problem of determining the gain of the transducer array that generates the desired sound pressure distribution.

1.1 Forward Problem of the Acoustic Field

The first step is to formulate the sound field produced by the ultrasound-phased array. The ultrasound transducers that make up the array emit ultrasound waves due

A. Matsubayashi (✉) · S. Inoue · S. Suzuki · H. Shinoda
University of Tokyo, Tokyo, Japan
e-mail: matsubayashi@hapis.k.u-tokyo.ac.jp

to the piston motion of their internal diaphragms. The sound field produced by an individual transducer is, in fact, affected by various factors, but let us first assume that it is a point source emitting spherical wave with frequency f_0 at the origin. In this case, the sound pressure field created by the transducer is as follows:

$$p(t, x, y, z) = \frac{q^{\text{amp}}}{\sqrt{x^2 + y^2 + z^2}} e^{j(-k\sqrt{x^2+y^2+z^2}+2\pi f_0 t + \theta)}, \tag{1}$$

where q^{amp}, θ are the intensity and phase of the point source, c is the speed of sound, and $k = 2\pi f_0/c$ is the wave number. This equation can be transformed into the product of the transfer function g, complex gain of the transducer q, and time-dependent term.

$$p(t, x, y, z) = g(x, y, z)q(q^{\text{amp}}, \theta)e^{j2\pi f_0 t}, \tag{2}$$

where

$$g(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}} e^{-jk\sqrt{x^2+y^2+z^2}} \tag{3}$$

$$q(q^{\text{amp}}, \theta) = q^{\text{amp}} e^{j\theta}. \tag{4}$$

Focusing on the f_0 frequency component of the sound field $\hat{p}(x, y, z)$, the relation between the complex gain and the sound field is expressed as follows:

$$\hat{p}(x, y, z) = g(x, y, z) \cdot q(q^{\text{amp}}, \theta). \tag{5}$$

To extend the transfer function to a general position, let $\mathbf{x}$ be the position of the sound source and $\mathbf{x}'$ be the observed position. In this case, the transfer function $g(\mathbf{x}, \mathbf{x}')$ is as follows:

$$g(\mathbf{x}, \mathbf{x}') = \frac{1}{|\mathbf{x} - \mathbf{x}'|} e^{-jk|\mathbf{x}-\mathbf{x}'|}. \tag{6}$$

Let us consider a case where there are multiple transducers. Suppose that there are N transducers, each with a gain of $q_1, q_2, \ldots q_N$, at $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N$. The sound pressure at point $\mathbf{y}$ can be expressed using vectors:

$$\hat{p}(\mathbf{y}) = \sum_i g(\mathbf{x}_i, \mathbf{y})q_i = \mathbf{g}^\top(\mathbf{y})\mathbf{q}. \tag{7}$$

where

$$\mathbf{g}(\mathbf{y}) = \begin{pmatrix} g(\mathbf{x}_1, \mathbf{y}) \\ g(\mathbf{x}_2, \mathbf{y}) \\ \vdots \\ g(\mathbf{x}_N, \mathbf{y}) \end{pmatrix}, \tag{8}$$

$$\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}. \tag{9}$$

This equation represents the continuous sound pressure created when the gain of the phased array transducer is given.

In practical situations pertaining to mid-air haptic rendering, we often consider controlling the sound pressure at discrete points in the observation space $\mathbf{y}$. When the discretized control points are represented by $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M$, the sound pressure at each point can be expressed as follows:

$$\hat{p}(\mathbf{y}_j) = \sum_i g(\mathbf{x}_i, \mathbf{y}_j)q_i = \mathbf{g}^\top(\mathbf{y}_j)\mathbf{q}. \tag{10}$$

Therefore, the sound pressure vector at control points $\mathbf{y}$ follows the matrix equation:

$$\hat{\mathbf{p}} = G\mathbf{q} \tag{11}$$

where

$$G = (\mathbf{g}(\mathbf{y}_1), \mathbf{g}(\mathbf{y}_2), \ldots, \mathbf{g}(\mathbf{y}_M))^\mathsf{T}. \tag{12}$$

This is the basic equation describing the forward problem of the phased array sound field. The vector of the transducer's gain $\mathbf{q}$ and the sound pressure on the control points $\hat{\mathbf{p}}$ are related by the transfer function matrix G.

As mentioned previously, matrix G represents the transfer function for the case where each transducer can be approximated as a point source. However, similar matrix equations can be obtained for other practical situations pertaining to mid-air haptic rendering. For example, if we consider the transducer as a piston disk of radius a attached to a baffle plate, the transfer function at the far field is as follows:

$$g(x, y, z) = \frac{i\omega\rho u a^2 J_1(ka\sin\theta)}{ka\sin\theta\sqrt{x^2 + y^2 + z^2}} e^{-jk\sqrt{x^2+y^2+z^2}}, \tag{13}$$

where J_1 is a Bessel function of the first order, u is the velocity of the disk, and θ represents the angle between (x, y, z) and the normal of the disk. This approximation has been adopted in several mid-air haptic rendering studies using common cylindrical transducers.

Another example is to consider the effects of scattering. Thus far, we have considered the forward problem in a free-sound field. When considering scattering on the surface of an object such as a hand, according to the style of the boundary element method, the sound field can be represented as follows (Inoue et al. 2016):

$$B\hat{\mathbf{p}} = G\mathbf{q}, \tag{14}$$

Matrix B represents the scattering effect. Even in this case, the same matrix equation can be obtained by designating $\bar{G} = B^{-1}G$:

$$\hat{\mathbf{p}} = \bar{G}\mathbf{q}. \tag{15}$$

Mid-air haptic rendering based on this scattered-field equation can increase the pressure at a single point on the fingertip (Inoue et al. 2016) or create an accurate pressure distribution on the hand surface (Matsubayashi et al. 2020).

1.2   Inverse Problem

The problem we face in controlling ultrasound-phased arrays is the reverse of this forward problem. We must first determine the acoustic radiation pressure at the control points that we want to present to the user and then find the transducer complex gain to output the pressure distribution.

Before discussing the sound pressure distribution control, let us first consider a simple case in which we want to maximize the sound pressure at a single point. The creation of this single focus is the most fundamental control and is practically significant. In this case, Eq. (7) can be written as

$$\text{argmax}_{\mathbf{q}}\hat{p}(\mathbf{y}) = \text{argmax}_{\mathbf{q}}\mathbf{g}^{\top}(\mathbf{y})\mathbf{q}. \tag{16}$$

If we increase the energy of $\mathbf{q}$, p may become infinitely large; however, this is not realistic as a physical transducer has nominal power. That is, there is an upper limit on the absolute value for each component of p. As units are meaningless in this discussion, we set the upper limit to 1. That is, $0 \le q^{\text{amp}} \le 1$. Understandably, the solution is when each element is driven at its maximum, and the phases are intensified at the focal point. One of the transducer's complex gains that achieves this is as follows:

$$\mathbf{q} = \begin{pmatrix} \frac{g(\mathbf{x}_1,\mathbf{y})^*}{|g(\mathbf{x}_1,\mathbf{y})|} \\ \vdots \\ \frac{g(\mathbf{x}_N,\mathbf{y})^*}{|g(\mathbf{x}_N,\mathbf{y})|} \end{pmatrix}. \tag{17}$$

where $*$ denotes the complex conjugate.

Next, we consider the case of pressure distribution control. If the complex gain on the control points $\mathbf{p}$ is given, it is reasonable to assume that $\mathbf{q}$ can be obtained using the generalized inverse matrix $G^{\dagger}$.

$$\mathbf{q} = G^{\dagger}\mathbf{p} \tag{18}$$

However, two problems specific to ultrasound mid-air haptic rendering arise here. One is the limitation of the transducer amplitude. As in the case with a single focus, it becomes necessary to drive the transducer at an amplitude lower than the transducer's maximum power. The other is the determination of the phase distribution of the sound field. Humans cannot distinguish the phase differences of ultrasonic waves. We can feel tactile sensations as vibrations below 1 kHz, which is considerably lower than the ultrasound frequency, and we can only distinguish the phase difference even smaller than that (Kuroki et al. 2016). Therefore, we are not interested in the phase of the generated sound pressure distribution, and only the sound pressure amplitude is our target of control. In other words, given the desired amplitude of the sound pressure $\mathbf{p}^{\mathrm{amp}}$, we need to find $\mathbf{q}$ such that

$$\begin{aligned}
\forall j \in \{1, \cdots, M\} : \quad & \{\mathbf{p}^{\mathrm{amp}}\}_j = |\{G\mathbf{q}\}_j| \\
\forall i \in \{1, \cdots, N\} : \quad & 0 \le |\{\mathbf{q}\}_i| \le 1,
\end{aligned} \tag{19}$$

where $\{\mathbf{x}\}_i$ means i component of $\mathbf{x}$.

To date, various approaches have been attempted to solve this complex problem. Long et al. were the first to solve this problem and rendering haptic shape using ultrasound (Long et al. 2014). They proposed a method to determine q after solving an eigenvalue problem to find the sound pressure phase such that the focal points at control points strengthen each other. A little later, Inoue et al. (2015) proposed a method to determine the sound pressure phase first, similar to Long et al. by relaxing the phase optimization problem to a semidefinite programming. In a different approach, a fast method for finding q has been proposed by applying the Gerchberg–Saxton algorithm used in optics (Hertzberg et al. 2010; Inoue et al. 2015; Marzo and Drinkwater 2019). This algorithm has been improved to a faster and more accurate version, GS-PAT, by Plasencia et al. (2020). Matsubayashi et al. (2020) solved the least-squares problem using Levenberg–Marquardt method, which is slower than the other methods but achieves a more accurate reproduction of the sound pressure field. All of the above methods consider q as a continuous quantity. However, in practical applications, the amplitude and phase of the transducers are input in discrete quantities. Suzuki et al. (2021) focused on this and proposed a very fast method to determine q by combinatorial optimization. In the next section, we review these methods.

2 Overview of Various Algorithms

This section outlines the algorithms that have been proposed for generating ultrasonic amplitude distributions and their respective applications.

2.1 Eigenmethod

As described in the previous section, we are only interested in the amplitude of sound pressure on the control points. The eigenmethod proposed by Long et al. in 2014 determines a good candidate for phase distribution that can achieve the target amplitude distribution (Long et al. 2014). This method first considers the transducer gain vector $\bar{\mathbf{q}}_j$, which creates a focus of sound pressure p_j^{amp} on the control point j:

$$\bar{\mathbf{q}}_j = p_j^{\text{amp}} \begin{pmatrix} \frac{G_{j,1}^*}{\sum_{i=1}^{N} |G_{j,i}|^2} \\ \vdots \\ \frac{G_{j,N}^*}{\sum_{i=1}^{N} |G_{j,i}|^2} \end{pmatrix} \tag{20}$$

This vector is a minimum-norm solution. $G\bar{\mathbf{q}}_j$ represents the pressure distribution at the control points when a focus is generated at the control point j. Therefore, the following matrix R represents the interaction between focal points:

$$R = G\left(\bar{\mathbf{q}}_1, \cdots, \bar{\mathbf{q}}_M\right). \tag{21}$$

For phase of the focal points $\mathbf{t}$ ($|t_i| \leq 1$) such that $R\mathbf{t}$ is large, the focal points have a constructive effect on each other; thus, the energy efficiency from the transducers is high. This implies that it is easier to achieve the target amplitude distribution under the upper limit of the transducer array. Finding such a $\mathbf{t}$ is equivalent to solving the following eigenvalue problem:

$$R\mathbf{x} = \lambda\mathbf{x} \tag{22}$$

The eigenmethod seeks the eigenvector corresponding to the largest eigenvalue of R and uses its phase for the inverse problem. When phase $\mathbf{t}$ is obtained, the eigenmethod solves the matrix equation with weighted Tikhonov regularization to obtain the transducer gain $\mathbf{q}$.

$$\begin{pmatrix} G \\ \sigma_1^{\gamma} \cdots 0 \\ \vdots \ddots \vdots \\ 0 \cdots \sigma_N^{\gamma} \end{pmatrix} \mathbf{q} = \begin{pmatrix} \text{diag}(\mathbf{p}^{amp})\mathbf{t} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{23}$$

$$\sigma_i = \sqrt{\left| \sum_{j=1}^{M} \frac{G_{j,i} p_j^{\text{amp}}}{M} \right|} \tag{24}$$

where $\text{diag}(\mathbf{p}_{\text{amp}})$ is a diagonal matrix with $\mathbf{p}_{\text{amp}}$ as a diagonal element. γ is a regularization parameter. The solution $\mathbf{q}$ is then truncated so that it does not exceed the

Fig. 1 When a tracked hand touches a virtual object, multiple focal points are generated at the intersection (Long et al. 2014)

maximum output of the transducers ($q_i \in [0, 1]$). The computational complexity of solving this linear equation is $O(N^3)$, which is the bottleneck of this method. Therefore, it should be noted that the computation time of this method becomes extensive when the number of transducers is considerable. The entire process is presented in Algorithm 1.

Algorithm 1 Eigen Method

Input: $\mathbf{p}^{\text{amp}}$ and γ
1: $\mathbf{x} \leftarrow$ Largest eigenvector of R
2: **for** $j = 1$ to M **do**
3: $\quad t_j \leftarrow \frac{x_j}{|x_j|}$
4: **end for**
5: **for** $i = 1$ to N **do**
6: $\quad \sigma_i \leftarrow \sqrt{\left| \sum_{j=1}^{M} \frac{G_{j,i} \, p_j^{\text{amp}}}{M} \right|}$
7: **end for**
8: Solve $\begin{pmatrix} G \\ \begin{pmatrix} \sigma_1^{\gamma} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N^{\gamma} \end{pmatrix} \end{pmatrix} \mathbf{q} = \begin{pmatrix} \text{diag}(\mathbf{p}^{amp})\mathbf{t} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$
9: $\mathbf{q} \leftarrow \text{truncate}(\mathbf{q}, [0, 1])$
Output: $\mathbf{q}$

Algorithm 2 Corrected Eigen Method

Input: $\mathbf{p}^{\text{amp}}$ and γ
1: $\mathbf{x} \leftarrow$ Largest eigenvector of R
2: **for** $j = 1$ to M **do**
3: $\quad t_j \leftarrow \frac{x_j}{|x_j|}$
4: **end for**
5: **for** $i = 1$ to N **do**
6: $\quad \sigma_i \leftarrow \sqrt{\left| \sum_{j=1}^{M} \frac{G_{j,i} p_j^{\text{amp}}}{M} \right|}$
7: **end for**
8: Solve $\begin{pmatrix} G \\ \sigma_1^{\gamma} \cdots 0 \\ \vdots \ddots \vdots \\ 0 \cdots \sigma_N^{\gamma} \end{pmatrix} \mathbf{q} = \begin{pmatrix} \text{diag}(\mathbf{p}^{amp})\mathbf{t} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$
9: $\mathbf{p} \leftarrow G\mathbf{q}$
10: $\mathbf{q} \leftarrow \frac{M}{\sum_j \frac{p_i}{p_i^{\text{amp}}}} \mathbf{q}$
11: $\mathbf{q} \leftarrow \text{truncate}(\mathbf{q}, [0, 1])$
Output: $\mathbf{q}$

Using this algorithm, Long et al. configured a system to present the cross-sectional shape of an object. The object shape is rendered by generating many focal points in real time at the intersection of the hand and the object, as shown in Fig. 1. They demonstrated that this system enabled users to recognize the shape of objects only from haptic sensations. This system is the first reported example of generating a sound pressure distribution with a specific shape for mid-air haptic rendering by determining the optimal sound pressure phase. However, it has been shown that the eigenmethod produces a distribution with slightly less pressure than the given sound pressure amplitude $\mathbf{p}^{\text{amp}}$. Plasencia et al. modified the eigenmethod by multiplying the obtained transducer amplitudes by a correction factor (Plasencia et al. 2020) (shown in Algorithm 2). Comparisons between certain algorithms in their paper are described in Sect. 2.4.

2.2 Semidefinite Relaxation

In 2015, Inoue et al. proposed a shape rendering method based on the generation of sound pressure distribution. Similar to the eigenmethod, this method first obtains the distribution of the optimal sound pressure phase. This method considers the following minimization problem.

$$\begin{aligned} \text{minimize} \quad & \|G\mathbf{q} - \text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t}\|_2^2 \\ \text{subject to} \quad & |t_i| = 1, \quad i \in \{1, \cdots, N\} \end{aligned} \quad (25)$$

Then, it assumes that, when the phase distribution $\mathbf{t}$ is obtained, the transducer gain is determined by $\mathbf{q} = G^-\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t}$, where G^- is the Tikhonov regularization matrix $G^- = (G^*G + \lambda I)G^*$. In this case, the objective function can be transformed as follows:

$$\|G\mathbf{q} - \text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t}\|_2^2 \tag{26}$$

$$= \|(GG^- - I)\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t}\|_2^2 \tag{27}$$

$$= \mathbf{t}^*\text{diag}(\mathbf{p}^{\text{amp}})(GG^- - I)^*(GG^- - I)\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t}. \tag{28}$$

Denoting $M = \text{diag}(\mathbf{p}^{\text{amp}})(GG^- - I)^*(GG^- - I)\text{diag}(\mathbf{p}^{\text{amp}})$, the minimization problem can be expressed as follows:

$$\begin{aligned} \text{minimize} \quad & \mathbf{t}^*M\mathbf{t} \\ \text{subject to} \quad & |t_i| = 1, \quad i \in \{1, \cdots, N\} \end{aligned} \tag{29}$$

To make the problem easier to solve, it can be rewritten as an equivalent problem.

$$\begin{aligned} \text{minimize} \quad & \text{Tr}(TM) \\ \text{subject to} \quad & \text{diag}(T) = 1, \ T \succeq 0, \text{rank}(T) = 1, \end{aligned} \tag{30}$$

where $T = \mathbf{t}\mathbf{t}^*$. If the rank constraint is removed, we obtain a semidefinite programming (SDP) problem, and the global optimal solution can be easily obtained. To solve this SDP problem efficiently, Inoue et al. employed the block-coordinate descent method. The approximate solution of the original problem $\mathbf{t}$ was obtained as the phase of the eigenvector corresponding to the largest eigenvalue of the solution of the relaxed problem T. After phase determination, the transducer amplitude was obtained by solving the Tikhonov regularized linear equation. Similar to the eigen-method, the computational complexity of this part is $O(N^3)$, which is the bottleneck of this method. The details are shown in Algorithm 3.

Using this algorithm, Inoue et al. constructed a system called HORN, which presents volumetric tactile objects in air (Inoue et al. 2015). This system uses a large number of ultrasonic transducers to present a sound pressure pattern of sufficient intensity with some shape. As shown in Fig. 2, this allows the user to interact with a virtual object without any delays and time resolution losses caused by hand tracking.

Algorithm 3 Semidefinite Relaxation i^c denotes the index set $i^c = \{1, \cdots, N\}\backslash\{i\}$, and X_{i^c,i^c} denotes the matrix X minus i rows and i columns.

Input: $\mathbf{p}^{\mathrm{amp}}$, $\mathbf{t}_0$ and $\mu > 0$ (small valued parameter)
 $T \Leftarrow \mathbf{t}_0\mathbf{t}_0^*$
 for $k = 1$ to K **do**
  Pick $i \in \{1, \cdots, N\}$
  $\mathbf{h} \Leftarrow T_{i^c,i^c}M_{i^c,i}$
  $\gamma \Leftarrow \mathbf{h}^*M_{i^c,i}$
  if $\gamma > 0$ **then**
   $T_{i^c,i} \Leftarrow -\sqrt{\frac{1-\mu}{\gamma}}\mathbf{h}$
   $T_{i,i^c} \Leftarrow -\sqrt{\frac{1-\mu}{\gamma}}\mathbf{h}^*$
  else
   $T_{i^c,i} \Leftarrow \mathbf{0}$
   $T_{i,i^c} \Leftarrow \mathbf{0}$
  end if
 end for
 $\hat{\mathbf{t}} \leftarrow$ Largest eigenvector of T
 for $i = 1$ to M **do**
  $t_i \Leftarrow \hat{t}_i/|\hat{t}_i|$
 end for
 $\mathbf{q} \Leftarrow G^-\mathrm{diag}(\mathbf{p}^{\mathrm{amp}})\mathbf{t}$
 $\mathbf{q} \leftarrow \mathrm{truncate}(\mathbf{q}, [0, 1])$
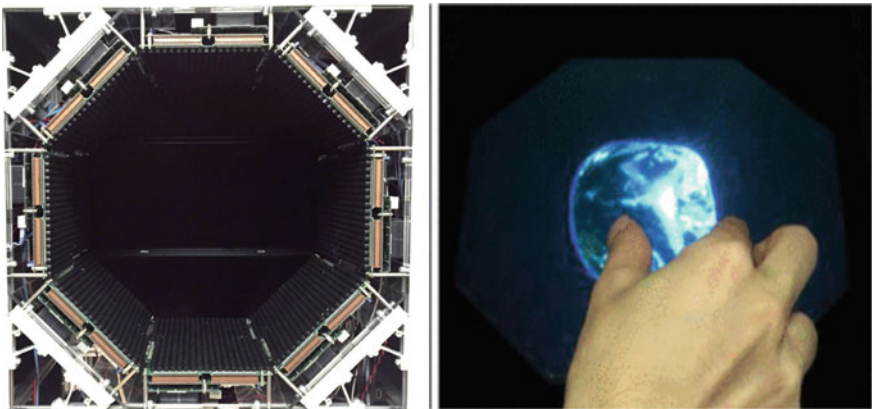**Output:** $\mathbf{q}$

Fig. 2 HORN generates static volumetric haptic image using a large number of transducers surrounding workspace

2.3 Gerchberg–Saxton Algorithm

The methods of first finding the optimum sound pressure phase and then solving the linear equation, as described in the previous sections, are computationally expensive when a large number of transducers are used. In contrast, the methods for iteratively obtaining a solution at high speeds have been studied. One of them is the Gerchberg–Saxton algorithm (GSA) (Gerchberg 1972). The GSA is known as an effective tool to solve phase recovery problems and has attracted attention in a wide range of fields, such as electromicroscopy, computer holography, and astronomy. Phase recovery is the problem of estimating the phase of a physical quantity when only the intensity of the quantity is observed. Our problem can be considered to be a type of phase recovery problem. In terms of acoustic pressure amplitude control using an ultrasound-phased array, which is the subject of this chapter, GSA-based methods have been proposed in the context of hyperthermia (Hertzberg et al. 2010), mid-air haptic rendering (Inoue et al. 2015), and acoustic manipulation (Marzo and Drinkwater 2019). The details of the algorithm differ between these methods, and the method proposed by Marzo and Drinkwater (2019) is presented here.

In this method, each ultrasound transducer is assumed to change only its phase. In one iteration, constraints are placed on each of the two vectors $\mathbf{q}$, $\mathbf{p}$ while alternately propagating between them as follows:

1. Propagate forward: $\mathbf{p} \Leftarrow G\mathbf{q}$

2. Impose constraint: $p_i \Leftarrow p_i^{\mathrm{amp}} \frac{p_i}{|p_i|}$

3. Propagate backward: $\mathbf{q} \Leftarrow G^*\mathbf{p}$

4. Impose constraint: $q_i \Leftarrow \frac{q_i}{|q_i|}$

Algorithm 4 Gerchberg-Saxton Algorithm

Input: $\mathbf{p}^{\mathrm{amp}}$ and $\mathbf{p}^0$ such that $|p_j^{(0)}| = p_j^{\mathrm{amp}}$
1: $\mathbf{p} \Leftarrow \mathbf{p}^0$
2: **for** $k = 0$ to $K - 1$ **do**
3: $\quad \hat{\mathbf{q}} \Leftarrow G^*\mathbf{p}$
4: $\quad$ **for** $i = 1$ to N **do**
5: $\quad\quad q_i \Leftarrow \frac{\hat{q}_i}{|\hat{q}_i|}$
6: $\quad$ **end for**
7: $\quad \hat{\mathbf{p}} \Leftarrow G\mathbf{q}$
8: $\quad$ **for** $j = 1$ to M **do**
9: $\quad\quad p_j \Leftarrow p_j^{\mathrm{amp}} \frac{\hat{p}_j}{|\hat{p}_j|}$
10: $\quad$ **end for**
11: **end for**
Output: q

The computational complexity in one iteration is $O(NM)$. In their study, they state that approximately 100 iterations are sufficient. Considering that $N \gg M$ in several applications, this method is fast and was proposed to take advantage of its high speed to perform acoustic manipulation of multiple micro-objects in three dimensions. However, because acoustic manipulation and mid-air haptic rendering have various similarities in terms of the requirements of the target sound field, this algorithm is also useful for mid-air haptic rendering.

2.4 GS-PAT

In 2020, Plasencia et al. proposed an algorithm that is more accurate and faster than the GSA, that is, GS-PAT. The GSA uses G^* as the back-propagation matrix, which is inaccurate in terms of the amplitudes of the transducer. Conversely, GS-PAT uses the normalized matrix F and updates $\mathbf{p}$ in the following manner:

1. $\hat{\mathbf{p}} \Leftarrow GF\mathbf{p}$

$$= G \begin{pmatrix} \frac{G_{0,0}^*}{\sum_{i=1}^N |G_{0,i}|^2} & \cdots & \frac{G_{M,0}^*}{\sum_{i=1}^N |G_{M,i}|^2} \\ \vdots & \ddots & \vdots \\ \frac{G_{0,N}^*}{\sum_{i=1}^N |G_{0,i}|^2} & \cdots & \frac{G_{M,N}^*}{\sum_{i=1}^N |G_{M,i}|^2} \end{pmatrix} \mathbf{p}.$$

2. $p_i \Leftarrow p_i^{\text{amp}} \dfrac{\hat{p}_i}{|\hat{p}_i|}, \qquad i \in \{1, \cdots, N\}.$

The matrix F is composed of vectors that are minimum-norm solutions when each control point is generated independently, which is identical to the one used in the eigenmethod described in Sect. 2.1. In GS-PAT, the matrix $R = GF$ is calculated in advance to reduce the computational complexity of the matrix-vector product in iterations. When the number of iterations of the algorithm is K, The computational complexity of the GSA is $O(KNM)$, whereas that of GS-PAT is $O(KM^2 + NM^2)$. Considering $M < K < N$ in various practical situations pertaining to multi-point mid-air haptic rendering and multi-object manipulations, GS-PAT has an advantage over the GSA in terms of computation time. Using a middle-end GPU (NVIDIA GTX 1660) with $N = 512$, $M = 32$, $K = 100$, GS-PAT was experimentally shown to be capable of 17000 optimizations per second.

Plasencia et al. also conducted simulations to compare the accuracy of the eigenmethod, the GSA, and GS-PAT for multi-point amplitude control. The results showed that GS-PAT performed well as the other algorithms and was considerably faster. However, when the number of focal points was large ($M \geq 16$), GS-PAT was found to be less accurate than the eigenmethod.
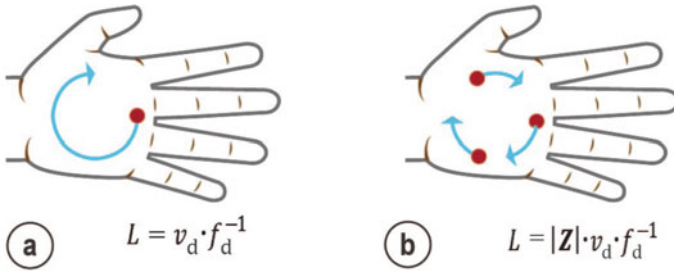
Fig. 3 **a** Spatio-temporal modulation, **b** independent control of focal movement speed and refresh rate of haptic stimuli by multi-focus generation (Plasencia et al. 2020)

The fast multi-point pressure rendering achieved by this algorithm has the potential to enable more diverse mid-air haptics sensations. In recent years, it has been reported that, by moving a single focal point at a high speed, it is possible to provide more intense tactile stimulation and recognize the shape of the trajectory (Frier et al. 2018; Takahashi et al. 2018) (see Chap. "Modulation Methods for Ultrasound Midair Haptics," for details). However, how the speed and frequency of the tactile stimuli affect human perception remains to be clarified. Fast multi-focus generation enables independent control of the focus movement speed and the refresh rate of the haptic stimulus in these modulation methods. For example, rotating three focal points on a circle can present haptic sensations at three times the refresh rate of a single point rotated at the same speed (see Fig. 3). This algorithm is expected to reveal a variety of human perceptual characteristics.

Algorithm 5 GS-PAT

Input: $\mathbf{p}^{\mathrm{amp}}$ and $\mathbf{p}^0$ such that $|p_j^{(0)}| = p_j^{\mathrm{amp}}$
1: $\mathbf{p} \Leftarrow \mathbf{p}^0$
2: $R \Leftarrow GF$
3: **for** $k = 0$ to $K - 1$ **do**
4:    $\hat{\mathbf{p}} \Leftarrow R\mathbf{p}$
5:    **for** $i = 1$ to M **do**
6:       $p_i \Leftarrow p_i^{\mathrm{amp}} \frac{\hat{p}_i}{|\hat{p}_i|}$
7:    **end for**
8: **end for**
9: $p_i^{(\Omega)} \Leftarrow (p_i^{\mathrm{amp}})^2 \frac{\hat{p}_i}{|\hat{p}_i|^2}$
10: $\mathbf{q} = F\mathbf{p}^{\Omega}$
11: $\mathbf{q} \leftarrow \mathrm{truncate}(\mathbf{q}, [0, 1])$
Output: q

2.5 *Levenberg–Marquardt Algorithm*

Basically, there is a tradeoff between accuracy and speed in sound pressure amplitude control. In 2019, Matsubayashi et al. proposed a method to control the sound pressure amplitude accurately, although it is slower than previous methods. They used the Levenberg–Marquardt algorithm (LMA), which is known as an effective solution method for unconstrained nonlinear least-squares problems, to control the sound pressure amplitude in a scattered sound field. The scattering of ultrasonic waves on the surface of the hand is not a problem for macroscopic tactile rendering (e.g., different intensities of tactile rendering for each of the five fingers). However, when more detailed pressure reproduction is required, such as when we want to control the shape of the pressure generated on the fingertips, we need to consider its effect. They described the optimization for controlling the scattering sound field as the following least-squares problem.

$$
\begin{aligned}
&\text{minimize} \quad \|\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t} - B^{-1}G\mathbf{q}\|_2^2 \\
&\text{subject to} \quad |t_i| = 1, |q_i| \le 1,
\end{aligned} \tag{31}
$$

If we fix the transducer amplitude to the maximum value and omit the computationally expensive calculation of B^{-1}, the following problem is obtained:

$$
\begin{aligned}
&\text{minimize} \quad \|B\,\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t} - G\mathbf{q}\|_2^2 \\
&\text{subject to} \quad |t_i| = 1, |q_i| = 1
\end{aligned} \tag{32}
$$

Furthermore, by setting $\mathbf{q} = [e^{j\theta_1}, \cdots, e^{j\theta_N}]^T$ and $\mathbf{t} = [e^{j\theta_{N+1}}, \cdots, e^{j\theta_{N+M}}]^T$, the problem is simplified to an unconstrained least-squares problem for the phases of sound pressure and transducers. $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_{M+N}]^T$.

Without going into detail, the update step of $\boldsymbol{\theta}$ in the LMA is calculated as follows:

$$
\mathbf{h} = -[J(\boldsymbol{\theta})^T J(\boldsymbol{\theta}) + \lambda I]^{-1} J(\boldsymbol{\theta})^T \mathbf{f}(\boldsymbol{\theta}), \tag{33}
$$

where

$$
\mathbf{f}(\boldsymbol{\theta}) = \begin{pmatrix} \text{Re}[B\,\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t} - G\mathbf{q}] \\ \text{Im}[B\,\text{diag}(\mathbf{p}^{\text{amp}})\mathbf{t} - G\mathbf{q}]. \end{pmatrix}, \tag{34}
$$

and $J(\boldsymbol{\theta})$ is a Jacobian matrix of $\mathbf{f}(\boldsymbol{\theta})$. Only if the value of the objective function after the step $\|\mathbf{f}(\boldsymbol{\theta} + \mathbf{h})\|_2^2$ decreases, the phases will be updated $\boldsymbol{\theta} \Leftarrow \boldsymbol{\theta} + \mathbf{h}$. In addition, λ is a damping factor that contributes to the convergence stability of the rhythm and is updated according to the behavior of the objective function after the step. When λ is large, the behavior of the LMA is similar to the steepest descent method, and when λ is small, the LMA converges in a manner similar to the Gauss–Newton method.
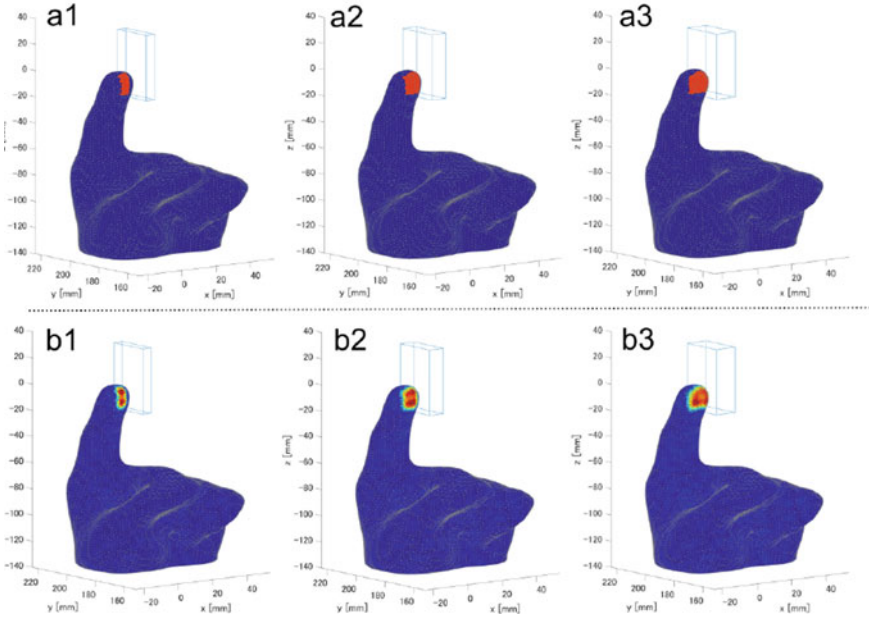
Fig. 4 Pressure amplitude control of the finger surface: (a1-a3) Target amplitude distribution. (b1-b3) Simulation results of the distribution reproduced by LMA (Matsubayashi et al. 2020)

Algorithm 6 describes the details of this process in which the update method for λ follows (Madsen et al. 2004).

Matsubayashi et al. proposed a method to dynamically generate a mesh model of the hand surface and render the sound pressure amplitude on the surface in real time using the LMA described above. They generated pressure amplitude distributions with different widths at the fingertips (see Fig. 4) and verified that these distributions were discriminable via a user study. This is the first reported study of accurately controlling the sound pressure amplitude distribution on the skin surface by considering the scattering on the hand surface.

LMA can be applied not only to the scattered sound field but also to the free field by replacing matrix B with the identity matrix. Sakiyama et al. used LMA in a free-field condition to reproduce the pressure distribution measured by a microphone array to render textures such as fingers, brushes, and towels.

LMA can generate an accurate amplitude distribution, but it is computationally expensive. The bottleneck of this method is that linear Eq. 33 need to be solved for step calculations, whose computational complexity is represented by $O((N+M)^3)$. However, if there is a control point with zero target amplitude, the phase at that point is not considered, and the computational complexity is reduced. The number of zero-amplitude control points can be quite large in situations where we want to prevent the sound pressure from being affected by scattering effects, other than the tactile point position. If the number of control points with nonzero amplitude is M', the

---

Algorithm 6 Levenberg-Marquardt Algorithm

Input: $\mathbf{p}^{\mathrm{amp}}$, $\boldsymbol{\theta}_0$, and $\gamma > 0$ (small valued parameter)

$\quad \boldsymbol{\theta} \Leftarrow \boldsymbol{\theta}_0$
$\quad \nu \Leftarrow 2$
$\quad A \Leftarrow J(\boldsymbol{\theta})^T J(\boldsymbol{\theta})$
$\quad \mathbf{g} \Leftarrow J(\boldsymbol{\theta})^T \mathbf{f}(\boldsymbol{\theta})$
$\quad \lambda \Leftarrow \gamma * \max A_{ii}$
$\quad$ **for** $k = 1$ to K **do**
$\quad\quad$ solve $(A + \lambda I)\mathbf{h} = -\mathbf{g}$
$\quad\quad \sigma \Leftarrow (F(\boldsymbol{\theta} + \mathbf{h}) - F(\boldsymbol{\theta}))/(\mathbf{h}^T(\mathbf{g} - \lambda \mathbf{h}))$
$\quad\quad$ **if** $\sigma > 0$ **then**
$\quad\quad\quad \boldsymbol{\theta} \Leftarrow \boldsymbol{\theta} + \mathbf{h}$
$\quad\quad\quad \lambda \Leftarrow \lambda * \max\left(\frac{1}{2}, 1 - (4\sigma - 1)^3\right)$
$\quad\quad\quad A \Leftarrow J(\boldsymbol{\theta})^T J(\boldsymbol{\theta})$
$\quad\quad\quad \mathbf{g} \Leftarrow J(\boldsymbol{\theta})^T \mathbf{f}(\boldsymbol{\theta})$
$\quad\quad\quad \nu \Leftarrow 2$
$\quad\quad$ **else**
$\quad\quad\quad \lambda \Leftarrow \lambda * \nu$
$\quad\quad\quad \nu \Leftarrow 2 * \nu$
$\quad\quad$ **end if**
$\quad$ **end for**
Output: $\boldsymbol{\theta}$

---

computational complexity is $O((N + M')^3)$. Using a high-end GPU (NVIDIA GeForce RTX 2080 Ti), it has been shown experimentally that each iteration takes approximately 10 ms when $N + M' = 1500$, $M = 10,000$. In this case, five iterations for one optimization would result in a haptic refresh rate 20 Hz.

2.6 Combinatorial Optimization

In the methods described so far, the amplitude and phase of the transducers were taken as continuous quantities. However, in practical use, they are input as discrete quantities. Furthermore, it has been found that even small amounts of gain information input to each transducer can produce accurate distributions. For example, it was reported that about eight values (3-bit) in phase are sufficient to reproduce some sound field patterns (Morales et al. 2021). Therefore, the inverse problem can be formulated as combinatorial optimization. This is much simpler than the problems we have considered so far, and allows for fast optimization.

In 2021, Suzuki et al. proposed a method to solve the inverse problem by discretizing the gain and applying a greedy algorithm (Suzuki et al. 2021). In this method, the amplitude q^{amp} and phase θ of the transducer are divided into I and J respectively;

Algorithm 7 Combinatorial optimization

Input: $\mathbf{p}^{\text{amp}}$

  for $i = 1$ to N **do**

    $q_i \Leftarrow q_*^{\text{amp}} e^{-j\theta_*}$ s.t. $q_*^{\text{amp}}, \theta_* = \arg\min E_i(q_i^{\text{amp}}, \theta_i)$

  end for

Output: q

$$q^{\text{amp}} \in \left\{ \frac{1}{I}, \frac{2}{I}, \ldots, 1 \right\}, \tag{35}$$

$$\theta \in \left\{ 0, 2\pi \frac{1}{J}, \ldots, 2\pi \frac{J-1}{J} \right\}. \tag{36}$$

In this method, the gain $q_i = q_i^{\text{amp}} e^{j\theta_i}$ is determined one by one for all transducers by searching all combinations of these discretized gain sets. Algorithm 7 describes this method. The decision to the gain of the i-th transducer is made by minimizing the difference between the target sound pressure and the sound pressure when the 1 to i-th transducers are driven. Therefore, objective function E_i to decide the gain of the i-th transducer is set as

$$E_i(q_i^{\text{amp}}, \theta_i) = \sum_{j=1}^{M} \left| p_j^{\text{amp}} - \left| \sum_{k=1}^{i} g(\mathbf{x}_k, \mathbf{y}_j) q_k^{\text{amp}} e^{j\theta_k} \right| \right|^2. \tag{37}$$

To minimize the objective function $E_i(q_i^{\text{amp}}, \theta_i)$, the brute-force search is used, which means that the objective function is computed for all combinations and the smallest one is selected. The computational complexity of the whole process is $O(IJMN)$, but since this method can generate distributions with sufficient accuracy using small values of I and J, the computation time is short in various practical situations.

Suzuki et al. have implemented the previously described methods (the eigenmethod, the semidefinite relaxation, GS-PAT, the LMA, and this method) on a CPU and performed experiments to compare them. The results show that this method has the shorter computation time than the other method when $I = 1$, $J = 16$. They also performed simulations to generate multiple foci with identical amplitudes. Their results showed that the method could reconstruct the sound pressure given at the control point with at least 80% accuracy which is better than eigenmethod and semidefinite relaxation. LMA gave the most accurate results, but the longest computation time. GS-PAT (Matsubayashi et al. 2020) showed an intermediate performance between this method and the LMA in terms of both accuracy and time.

3 Summary

In this chapter, we introduced the basic equations for controlling the ultrasonic sound field and the various algorithms for solving them. The requirements for sound field control in mid-air haptics rendering are accuracy and speed. The use of GS-PAT or the combinatorial optimization is suitable for fast movement of stimulus points, such as when generating a focus that perfectly follows a fast-moving finger. Note, however, that in such cases, even if the algorithm is fast enough, the effects of hardware and sound transmission delays need to be compensated for. On the other hand, if you want to generate an accurate sound pressure distribution, LMA is superior. Especially in the situation where tactile sensation is generated on a specific part of the hand, taking into account the effect of scattering on the surface of the hand, this algorithm is suitable because the computation time is not much affected by the number of zero pressure points. These and other ultrasonic sound field control algorithms proposed so far, and the applications realized by them have been introduced above, but this is only an overview. Please refer to the respective literature for details.

One of the major advantages of ultrasonic haptic technology is the ability to freely control the pressure applied to the skin surface. This is also important in investigating the sense of touch in the human body. Investigating how the spatial patterns of pressure and the waveforms of the vibrations applied to it affect human perception is the basis for future mid-air haptic rendering technology. It has the potential to define the necessary elements for haptic design and furthermore to create new tactile sensations that do not exist in the real world. For this purpose, further development of algorithms to control the ultrasonic field is expected.

References

Frier W, Ablart D, Chilles J, Long B, Giordano M, Obrist M, Subramanian S (2018) Using spatiotemporal modulation to draw tactile patterns in mid-air. In: International conference on human haptic sensing and touch enabled computer applications. Springer, pp 270–281

Gerchberg RW (1972) A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik 35:237–246

Hasegawa T, Kido T, Iizuka T, Matsuoka C (2000) A general theory of Rayleigh and Langevin radiation pressures. Acoustical Sci Technol 21(3):145–152

Hertzberg Y, Naor O, Volovick A, Shoham S (2010) Towards multifocal ultrasonic neural stimulation: pattern generation algorithms. J Neural Eng 7(5):056002

Inoue S, Makino Y, Shinoda H (2015) Active touch perception produced by airborne ultrasonic haptic hologram. In: 2015 IEEE World Haptics Conference (WHC). IEEE, pp 362–367

Inoue S, Makino Y, Shinoda H (2016) Mid-air ultrasonic pressure control on skin by adaptive focusing. In: International conference on human haptic sensing and touch enabled computer applications. Springer, Heidelberg, pp 68–77

Kuroki S, Watanabe J, Nishida S (2016) Neural timing signal for precise tactile timing judgments. J Neurophysiol 115(3):1620–1629 PMID: 26843600

Long B, Seah SA, Carter T, Subramanian S (2014) Rendering volumetric haptic shapes in mid-air using ultrasound. ACM Trans Graphics (TOG) 33(6):1–10

Madsen K, Nielsen H, Tingleff O (2004) Methods for non-linear least squares problems, 2nd edn, p 60

Marzo A, Drinkwater BW (2019) Holographic acoustic tweezers. Proc Natl Acad Sci 116(1):84–89

Matsubayashi A, Makino Y, Shinoda H (2020) Rendering ultrasound pressure distribution on hand surface in real-time. In: International conference on human haptic sensing and touch enabled computer applications. Springer, Heidelberg, pp 407–415

Morales R, Ezcurdia I, Irisarri J, Andrade MA, Marzo A (2021) Generating airborne ultrasonic amplitude patterns using an open hardware phased array. Appl Sci 11(7):2981

Plasencia DM, Hirayama R, Montano-Murillo R, Subramanian S (2020) Gs-pat: high-speed multi-point sound-fields for phased arrays of transducers. ACM Trans Graphics (TOG) 39(4):138-1

Suzuki S, Fujiwara M, Makino Y, Shinoda H (2021) Radiation pressure field reconstruction for ultrasound midair haptics by greedy algorithm with brute-force search. IEEE Trans Haptics

Takahashi R, Hasegawa K, Shinoda H (2018) Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli. In: International conference on human haptic sensing and touch enabled computer applications. Springer, Heidelberg, pp 276–288