Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 11, No. 1, pp. 37–56, January 2015

LOCALIZATION OF ACOUSTIC SOURCE VIA OPTIMAL BEAMFOMER DESIGN

Zhiguo Feng, Ka Fai Cedric Yiu*,
Randolph Chi Kin Leung and Sven Nordholm

Abstract: In localizing acoustic sources in an enclosed area, the effect of reverberation remains a difficult obstacle. With various reflections within the area, the acoustic energy map of the enclosure is becoming rather complex and spiky, so finding the correct location is often tricky and erroneous. In this paper, a new framework of an acoustic source localization system is proposed. First of all, based on a room model, the frequency response between a point in the room and a microphone can be calculated. The idea is to discretize the room and to form a map of beamformers for all discrete points. The beamformer design is formulated as a functional optimization problem, and the optimized frequency response function can be sought directly. We show that, under some mild conditions, any source location will be the global maximizer of a response function. Once the beamforming map is constructed, the implementation in finding the source location can be done in real time. For illustration, we demonstrate how the proposed method localizes the source with three numerical examples.

Key words: *localization, beamformer design, reverberation, semi-definite programming*

Mathematics Subject Classification: *68U99, 90C90*

1 Introduction

The localization problem of a wideband acoustic source arises in many fields, such as teleconferencing, automatic camera steering, hands-free audio devices, and human-computer interaction systems. Because of the diversity of wideband signals, it is more difficult than the localization problem of narrowband source. A microphone array is an essential tool, which estimates the source location by overlapping several measurements from the set of microphones so that the hidden spatial information can be revealed.

Localization problem is estimating the location of the source based on the signal data received by multiple microphones. For a given source, the signals received by different microphones will be delayed by the spatial distances between microphones. Indeed, localization methods are exploiting these spatial differences to pinpoint the source. They can be generally divided into two types of methods. One is based on a relative time delay estimation, which estimates the source location based on the time delay information about the arrival of acoustic signals from different microphones. Examples of this method include

*The paper is supported by RGC Grant PolyU. (5301/12E), PolyU Central Research Grant No. G-YJ36, the Grant of Chongqing Science and Technology Commission (No. cstc2013jcyjA1407) and the Program of Chongqing Innovation Team Project in University under Grant (No. KJTD201308)

[4, 5, 2] and more references can be found in [3]. This kind of methods is computationally efficient. However, the relative time delay estimation become very poor in a reverberant environment and the accuracy of localization is not ensured. The other is based on the beamforming technique. Broadband beamforming is a general technique for spatial filtering [17] and signal enhancement [19, 14], which can also be applied directly to the localization problem by exploring the steering nature of beamformers. It takes all the information in the computational domain into account and it can produce better accuracy than the method based on the time delay estimation. The steered response power (SRP) algorithm [20, 6, 10] belongs to this kind. The particle filtering method [18] has also been employed in tracking the source, which uses a state space approach. The location is modeled as a Markov process where the current state is estimated by filtering on previous states.

In the literature, these methods are formulated in a non-reverberant environment. Methods based on delay information become inaccurate due to the presence of reverberation which creates many mirror images of the source. On the other hand, the steering nature of beamformers could differentiate small power level differences between the true source and reflected sources. But this has not been explored in the literature. In view of this, we propose a new approach here to solve the indoor acoustic source localization problem using array response powers, based on the broadband beamformer design method in [7].

The proposed method includes the several steps. First, a room acoustic model is employed to estimate the transfer functions between any two points in the computational domain. This will characterize the room acoustics with reverberation. Second, a series of beamformer design problems with specified passband and stopband regions will be solved to yield a collection of beams for the computational domain. Moreover, instead of solving the filter coefficients, the corresponding optimal frequency response vectors are sought instead as described in [7, 9]. In addition, the designed beamformers will be standardized via certain mild requirements derived in this paper. Third, by using the steering vectors and calculating the beamformer output levels for all mesh points, a two-dimensional function can be constructed which contains the output levels at each point. Finally, we prove that the source location will correspond to the maximum value of this two-dimensional output level function, and hence it can be solved by maximizing the function. The structure and property of the underlying optimization problem are explored in order to develop a more specific technique for finding the global solution to the localization problem.

To the best of our knowledge, this is the first method which employs a room acoustic model for profiling the reverberation characteristics for indoor sound source localization. In addition, when the steered response power is applied, it has the difficulty in standardizing different beams in various locations and the effect of filter length is also of concerns. These problems have been addressed fully in this paper where the conditions on the beamformer standardization are derived properly; the frequency response functions are optimized directly instead, which serves as the performance limit of long filters. Finally, we show that the source location is indeed the global maximum value of the constructed two-dimensional output level function. As a result, the localization problem is reduced to the optimization of a two-dimensional function, which can be tackled by certain existing methods.

The rest of the paper is organized as follows. In Section 2, we review the acoustic model for transfer functions. Based on estimated transfer functions, we develop a new framework for the acoustic source localization problem in Section 3. Detailed analysis of the problem is also carried out. In Section 4, the implementation details of the proposed method are described. For illustration, several numerical examples are solved in Section 5 to demonstrate the efficiency and effectiveness of the proposed method.

2 Transfer Function

In this section, we describe the image model [1] for estimating transfer functions with reverberation and employed the implementation proposed by [11]. A general model of the transfer function should be a complex function with respect to the frequency and it is also related to the source location and receiving location. It can be defined as $T(\boldsymbol{r}, \boldsymbol{q}, f, \theta)$, where $\boldsymbol{r}$ is the source location, $\boldsymbol{q}$ is the receiving location, f denotes the frequency of the signal and θ denotes the uncertainties of this model.

For specific surroundings, the acoustic transfer function is relatively stable. In the non-reflection case, the following transfer function

$$T(\boldsymbol{r}, \boldsymbol{q}, f, \theta) = \frac{1}{||\boldsymbol{r} - \boldsymbol{q}||} e^{-j2\pi f \frac{||\boldsymbol{r}-\boldsymbol{q}||}{c}}. \tag{2.2}$$

is often used, while in the indoor situation, the image model can be deployed. The model was proposed in [1] and it is the limit of the sum of many transfer functions (2.2) computed from all possible reflections of six walls. A fast implementation of the method was proposed in [11]. Essentially, given an enclosure with dimensions $\boldsymbol{L} = [L_x, L_y, L_z]^T$ with a source point $\boldsymbol{r}$ and a microphone receiver $\boldsymbol{q}$, the basic idea of the image model is to mimic images from an infinite grid of mirror rooms expanding in all three dimensions. Each mirror image is a replica of the source signal delayed by a lag τ and attenuated by an amplitude factor A:

$$A(\boldsymbol{u}, \boldsymbol{v}) = \frac{\beta_{x,1}^{|v_x-u_x|}\beta_{x,2}^{|v_x|}\beta_{y,1}^{|v_y-u_y|}\beta_{y,2}^{|v_y|}\beta_{z,1}^{|v_z-u_z|}\beta_{z,2}^{|v_z|}}{4\pi \cdot d(\boldsymbol{u}, \boldsymbol{v})},$$

where $\boldsymbol{u} = (u_x, u_y, u_z)^T$ and $\boldsymbol{v} = (v_x, v_y, v_z)^T$, $\boldsymbol{\beta} = \{\beta_{x,i}, \beta_{y,i}, \beta_{z,i}, i = 1, 2\}$ are the reflection coefficients for the six enclosure surfaces, and $d(\cdot, \cdot)$ represents the distance as

$$d(\boldsymbol{u}, \boldsymbol{v}) = \|\mathrm{diag}(2u_x - 1, 2u_y - 1, 2u_z - 1) \cdot \boldsymbol{r} + \boldsymbol{q}\mathrm{diag}(2v_x, 2v_y, 2v_z) \cdot \boldsymbol{L}\|.$$

Then the indoor room impulse response between such a pair of source-receiver can be estimated by

$$h(\boldsymbol{r}, \boldsymbol{q}, t) = \sum_{\boldsymbol{u}=0}^{1} \sum_{\boldsymbol{v}=-\infty}^{+\infty} A(\boldsymbol{u}, \boldsymbol{v}) \cdot \delta(t - \tau(\boldsymbol{u}, \boldsymbol{v})),$$

where $\delta(\cdot)$ denotes the Dirac impulse function and $\tau(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{u}, \boldsymbol{v})/c$ is the time delay of the considered image source. By using a frequency domain representation, this room impulse response yields the required transfer function which can be employed in designing indoor beamformers as illustrated in [12, 13].

3 Problem Formulation

In this section, we propose a new framework of the localization problem formulation based on the estimated room transfer function described in the last section. In particular, it includes the steps of forming a collection of optimized beamformers for the computational domain, and a two dimensional function of the beamformer output levels is constructed for global optimization. Analysis on the variation of output levels for different beamformers are carried out, and conditions on the standardization are derived. Furthermore, the properties of the constructed output level function are studied.

3.1 Microphone Array Model

In this section, we describe the formulation of beamformer outputs based on the frequency response functions directly instead of filters as proposed in [7]. We assume that there is a microphone array of N elements which samples sound signals. Suppose that the input data is given by $X(f)$ at location $\boldsymbol{r}$ and that the locations of microphones are $\boldsymbol{q}_i, i = 1, \ldots, N$. The data received by the microphones can be denoted by

$$Y_i(\boldsymbol{r}, f) = X(f)T(\boldsymbol{r}, \boldsymbol{q}_i, f, \theta), \qquad i = 1, \ldots, N. \tag{3.1}$$

The received signals are processed by the filters behind. Suppose that the filter frequency responses are given by $H_i(f),\ i = 1, \ldots, N$. By summing up the processed data from all filters, the output response of the microphone array is given by

$$Y(\boldsymbol{r}, f) = \sum_{i=1}^{N} Y_i(\boldsymbol{r}, f)H_i(f) = \sum_{i=1}^{N} X(f)T(\boldsymbol{r}, \boldsymbol{q}_i, f, \theta)H_i(f). \tag{3.2}$$

Define a response function as

$$G(\boldsymbol{r}, f, \theta) = \sum_{i=1}^{N} H_i(f)T(\boldsymbol{r}, \boldsymbol{q}_i, f, \theta) = \boldsymbol{T}^{\intercal}(\boldsymbol{r}, f, \theta)\boldsymbol{H}(f), \tag{3.3}$$

where

$$\boldsymbol{T}(\boldsymbol{r}, f, \theta) = [T(\boldsymbol{r}, \boldsymbol{q}_1, f, \theta), \ldots, T(\boldsymbol{r}, \boldsymbol{q}_N, f, \theta)]^{\intercal}$$

is the transfer function vector and

$$\boldsymbol{H}(f) = [H_1(f), \ldots, H_N(f)]^{\intercal}$$

is the frequency response vector. Then, the output response (3.2) becomes

$$Y(\boldsymbol{r}, f) = X(f) \sum_{i=1}^{N} T(\boldsymbol{r}, \boldsymbol{q}_i, f, \theta)H_i(f) = X(f)G(\boldsymbol{r}, f, \theta). \tag{3.4}$$

3.2 An ε-accuracy Scheme Formulation

Since we need to compare different output levels of all optimized beamformers in the computational domain, there is a need to standardize the beamformers so that the output level difference due to different designs is within a small tolerant. In this and the following sections, the design requirements are derived in Theorem 1. Moreover, in Corollary 1 and 2, we show that, under the design requirements, location variations correspond to the difference in beamformer output levels. In particular, in Theorem 2, it is shown that the true location will have the maximum output level value of all. As a result, we can formulate the global maximization problem based on the output levels for the speaker location.

Equation (3.2) yields the output of the microphone array system. Our object is to derive the location information from this output. Note that in (3.2), the frequency response vector is yet to be designed. Note that Y is a function of $(\boldsymbol{r}, f)$, we can design the frequency response vector for a specific location. Since the speaker can be in all possible locations within the room, we will design a series of frequency response vectors such that all possible

locations can be localized. That is, we define a frequency response function $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ for each location $\bar{\boldsymbol{r}}$. Then, (3.2) becomes

$$\begin{aligned} Y(\boldsymbol{r}, \bar{\boldsymbol{r}}, f) =& \sum_{i=1}^{N} Y_i(\boldsymbol{r}, f) H_i(\bar{\boldsymbol{r}}, f) \\ =& X(f) \sum_{i=1}^{N} T(\boldsymbol{r}, \boldsymbol{q}_i, f, \theta) H_i(\bar{\boldsymbol{r}}, f) \\ =& X(f) G(\boldsymbol{r}, \bar{\boldsymbol{r}}, f, \theta), \end{aligned} \tag{3.5}$$

where the true source is located at $\boldsymbol{r}$, and $G(\boldsymbol{r}, \bar{\boldsymbol{r}}, f, \theta)$ is defined in (3.3) with the frequency response vector $\boldsymbol{H}(f)$ replaced by $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$.

In designing the frequency response function $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$, we need a criterion to measure whether $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ is suitable for localization. This criterion is set up for a given output function (3.5). In the ideal case, for any given source location $\boldsymbol{r}$, we can find a corresponding location $\bar{\boldsymbol{r}}$ by the output function (3.5) according to the criterion, which satisfies $\bar{\boldsymbol{r}} = \boldsymbol{r}$. It can be seen that this localization is exact, but it's difficult to pinpoint in practice. Instead, we can consider an ε-accuracy scheme; that is, for any given source location $\boldsymbol{r}$, we can find a location $\bar{\boldsymbol{r}}$ by the output function (3.5), which guarantees $||\bar{\boldsymbol{r}} - \boldsymbol{r}|| \leq \varepsilon$. The ε-accuracy scheme can be implemented in practice if ε is chosen adaptively. Basically, if ε is smaller, the localization accuracy should be better, while the implementation cost will become more expensive.

3.3 Beamformer Design

To implement an ε-accuracy scheme, the choice of frequency response vector which determines the properties of the output function (3.5) is very important. We choose the maximizer as the criterion, that is, $\bar{\boldsymbol{r}}$ is obtained by finding the maximizer of the output function (3.5). Therefore, we need to design the frequency response vector such that the output function (3.5) will achieve the maximum near the source location $\boldsymbol{r}$. The designed beamformers should include the location information. This can be done by the following specific beamformer design method.

From (3.3), the function $G(\boldsymbol{r}, f, \theta)$ is the actual response of the beamformer output, which is independent of the input signal $X(f)$. We can design the filter response vector such that it fits a given desired response function. We choose a given region called passband region, which is denoted by Ω_p and a given region called stopband region, which is denoted by Ω_s. Denote the specified space-frequency domain by $\Omega = \Omega_p \cup \Omega_s$. Then, we choose a cost function as

$$E(\boldsymbol{H}) = \max_{\Omega, \theta} \left| \boldsymbol{T}^{\intercal}(\boldsymbol{r}, f, \theta) \boldsymbol{H}(f) - G_d(\boldsymbol{r}, f) \right|, \tag{3.6}$$

where $G_d(\boldsymbol{r}, f)$ is a given desired response function. For the choice of $G_d(\boldsymbol{r}, f)$, its magnitude should be chosen as 1 in passband region and 0 in stopband region. Here, we set $G_d(\boldsymbol{r}, f)$ as

$$G_d(\boldsymbol{r}, f) = \begin{cases} e^{-j2\pi f \frac{||\boldsymbol{r} - \boldsymbol{r}_c||}{c}}, & (\boldsymbol{r}, f) \in \Omega_p \\ 0, & (\boldsymbol{r}, f) \in \Omega_s \end{cases}, \tag{3.7}$$

where $\boldsymbol{r}_c$ is the spatial point of a given reference location, which is often chosen as the centre of the microphone array.

The design problem is to find a frequency response vector $\boldsymbol{H}(f)$ such that the beamformer output $G(\boldsymbol{r}, f, \theta)$ fits a given desired response $G_d(\boldsymbol{r}, f)$; that is, the objective is to minimize the cost function (3.6)

$$\min_{\boldsymbol{H}} \quad E(\boldsymbol{H}). \tag{3.8}$$

It can be seen that this problem is a minimax optimization problem. A practical way of solving the minimax optimization problem is the discretization method, where the specified domain is replaced by a sufficiently dense grid. Note that the optimization problem (3.8) can be solved for each given frequency. Therefore, we decompose the domain Ω as

$$\Omega = \bigcup_{f} \Omega_f, \tag{3.9}$$

where, for each frequency f, Ω_f is the corresponding spatial domain in Ω for f. Denote the set of all the frequencies f by I_Ω. Then, for each frequency $f \in I_\Omega$, we consider the subproblem as

$$\min_{\boldsymbol{H}(f)} \quad E_f(\boldsymbol{H}(f)), \tag{3.10}$$

where

$$E_f(\boldsymbol{H}(f)) = \max_{\Omega_f} \left| \boldsymbol{T}^{\intercal}(\boldsymbol{r}, f, \theta)\boldsymbol{H}(f) - G_d(\boldsymbol{r}, f) \right|. \tag{3.11}$$

For this subproblem, note that the function inside the magnitude is linear with respect to the frequency response vector $\boldsymbol{H}(f)$. Hence, this problem can be transformed into a linear semi-definite programming problem, and an interior point method can be applied to solve it efficiently and effectively. Details of the transformation and the method can be seen in [7] and [8].

Denote the optimal solution by $\boldsymbol{H}^*(f)$, then the optimal value of this problem is given by $E^* = E(\boldsymbol{H}^*)$, which can also be computed by

$$E^* = E(\boldsymbol{H}^*) = \max_{I_\Omega} E_f(\boldsymbol{H}^*(f)). \tag{3.12}$$

We define E^* as the performance value of the beamformer. The smaller the E^*, the better the performance of the beamformer.

Note that the frequency response vector obtained by solving the problem (3.8) is for a specific location. We need to cover the whole domain by seeking the response function $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ for each location $\bar{\boldsymbol{r}}$. In this way, we have designed beamformer for each location $\bar{\boldsymbol{r}}$. For convenience, we denote the designed beamformer for the specific location $\bar{\boldsymbol{r}}$ as $\bar{\boldsymbol{r}}$-beamformer. In the following, the spatial domains Ω_f is aligned for all frequencies to be the same set, denoted by S. That is,

$$\Omega_f = S, \quad \forall f \in I_\Omega.$$

Then, the design domain Ω is $I_p \times S$, where I_p is the frequency domain including the human vocal spectrum.

For the spatial domain S, we define it to be the set of discrete locations in the room. To set up the beamformer, the passband and stopband are denoted by S_p and S_s, respectively. For the set of the $\bar{\boldsymbol{r}}$-beamformer, we choose the domain S_p containing $\bar{\boldsymbol{r}}$, that is, $\bar{\boldsymbol{r}} \in S_p$. Then, the response function $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ can be obtained by solving this beamformer problem.

3.4 Optimization Problem

After all the response functions $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ are obtained, we can formulate the localization optimization problem. First, we analyze the property of the $\bar{\boldsymbol{r}}$-beamformer.

Theorem 3.1. *Suppose that the performance value of the $\bar{\boldsymbol{r}}$-beamformer problem (3.8) is*

$$E^* < 0.5. \tag{3.13}$$

Then, $\forall(\boldsymbol{r}', f) \in \Omega_p$ and $\forall(\boldsymbol{r}'', f) \in \Omega_s$, we have

$$|Y(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f)| \le |X(f)|/2 \le |Y(\boldsymbol{r}', \bar{\boldsymbol{r}}, f)|. \tag{3.14}$$

The equality is true if and only if $|X(f)| = 0$.

Proof. Since $(\boldsymbol{r}', f) \in \Omega_p$, we have

$$\begin{aligned}
&|Y(\boldsymbol{r}', \bar{\boldsymbol{r}}, f)| = |X(f)| \cdot |G(\boldsymbol{r}', \bar{\boldsymbol{r}}, f, \theta)| \\
=&|X(f)| \cdot |G(\boldsymbol{r}', \bar{\boldsymbol{r}}, f, \theta) - G_d(\boldsymbol{r}', f) + G_d(\boldsymbol{r}', f)| \\
\ge&|X(f)| \cdot (|G_d(\boldsymbol{r}', f)| - |G(\boldsymbol{r}', \bar{\boldsymbol{r}}, f, \theta) - G_d(\boldsymbol{r}', f)|) \\
=&(1 - E^*) \cdot |X(f)| \\
\ge&|X(f)|/2.
\end{aligned}$$

Since $(\boldsymbol{r}'', f) \in \Omega_s$, we have

$$\begin{aligned}
&|Y(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f)| = |X(f)| \cdot |G(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f, \theta)| \\
=&|X(f)| \cdot |G(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f, \theta) - G_d(\boldsymbol{r}'', f) + G_d(\boldsymbol{r}'', f)| \\
\le&|X(f)| \cdot (|G_d(\boldsymbol{r}'', f)| + |G(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f, \theta) - G_d(\boldsymbol{r}'', f)|) \\
=&E^* \cdot |X(f)| \\
\le&|X(f)|/2.
\end{aligned}$$

From these two equations, it's easy to see that the symbol "$\ge$" becomes "$>$" if and only if $|X(f)| > 0$. This completes the proof. □

By Theorem 3.1, the beamformer is designed properly so that we can observe a clear difference between the passband region and the stopband region, which is essential for localization. It follows from Theorem 3.1 that we have the following.

Corollary 3.2. *Suppose that the performance value of the $\bar{\boldsymbol{r}}$-beamformer problem (3.8) satisfies the condition (3.13). Then, $\forall \boldsymbol{r}' \in S_p$ and $\forall \boldsymbol{r}'' \in S_s$ and $\forall f \in I_p$, we have*

$$|Y(\boldsymbol{r}'', \bar{\boldsymbol{r}}, f)|^2 \le |X(f)|^2/4 \le |Y(\boldsymbol{r}', \bar{\boldsymbol{r}}, f)|^2. \tag{3.15}$$

The equality is true if and only if $|X(f)| = 0$.

Corollary 3.2 gives the results for one beamformer. For two simultaneous beamformers, similar results can be obtained. Denote the spatial domain of the passband region and stopband region of $\bar{\boldsymbol{r}}$-beamformer by $S_p(\bar{\boldsymbol{r}})$ and $S_s(\bar{\boldsymbol{r}})$, respectively. The result is summarized below.

Corollary 3.3. *Suppose that for any two points $\bar{\boldsymbol{r}}', \bar{\boldsymbol{r}}'' \in S$, the performance values of $\bar{\boldsymbol{r}}'$-beamformer and $\bar{\boldsymbol{r}}''$-beamformer satisfy the conditions (3.13). If the true source location $\boldsymbol{r}$ belongs to $S_p(\bar{\boldsymbol{r}}')$ and also belongs to $S_s(\bar{\boldsymbol{r}}'')$, we have*

$$|Y(\boldsymbol{r}, \bar{\boldsymbol{r}}'', f)|^2 \leq |Y(\boldsymbol{r}, \bar{\boldsymbol{r}}', f)|^2, \quad \forall f \in I_p. \tag{3.16}$$

The equality is true if and only if $|X(f)| = 0$.

Next, we consider the localization optimization problem. Denote the complementary set of $S_s(\bar{\boldsymbol{r}})$ by

$$(S_s(\bar{\boldsymbol{r}}))^c = S \setminus S_s(\bar{\boldsymbol{r}}),$$

and the performance value of the $\bar{\boldsymbol{r}}$-beamformer by $E^*(\bar{\boldsymbol{r}})$. We have the following result.

Theorem 3.4. *Suppose that $|X(f)| > 0$ and for any $\bar{\boldsymbol{r}} \in S$, the performance value $E^*(\bar{\boldsymbol{r}})$ is less than 0.5. If*

$$\bar{\boldsymbol{r}} \in S_p(\bar{\boldsymbol{r}}), \qquad \forall \bar{\boldsymbol{r}} \in S, \tag{3.17}$$

then the true source location $\boldsymbol{r}$ satisfies

$$\boldsymbol{r} \in S_s(\bar{\boldsymbol{r}}^*)^c, \tag{3.18}$$

where $\bar{\boldsymbol{r}}^*$ is the maximizer of the function $|Y(\boldsymbol{r}, \bar{\boldsymbol{r}}, f)|^2$ with respect to $\bar{\boldsymbol{r}}$.

Proof. By (3.17), we have $\boldsymbol{r} \in S_p(\boldsymbol{r})$. Then, it follows by (3.15) that

$$|X(f)|^2/4 \leq |Y(\boldsymbol{r}, \boldsymbol{r}, f)|^2.$$

Hence, the maximizer point $\bar{\boldsymbol{r}}^*$ must satisfy

$$|X(f)|^2/4 < |Y(\boldsymbol{r}, \bar{\boldsymbol{r}}^*, f)|^2.$$

Then, it follows by (3.15) that $\boldsymbol{r}$ must not belong to $S_s(\bar{\boldsymbol{r}}^*)$, or the equation above can not be satisfied. Thus, we have the result (3.18). This completes the proof. □

By Theorem 3.4, the function $Y(\boldsymbol{r}, \bar{\boldsymbol{r}}, f)|^2$ can be used as the performance indicator to be maximized for localization if and only if $|X(f)| > 0$. To prevent the case of $|X(f)| = 0$, where the maximizer is not unique, we define the indicator function as

$$F(\bar{\boldsymbol{r}}) = \int_{I_p} |Y(\boldsymbol{r}, \bar{\boldsymbol{r}}, f)|^2 df, \tag{3.19}$$

where, for an acoustic signal, there exists some frequency f such that $|X(f)| > 0$. Then, the localization problem can be formulated into the following optimization problem:

$$\max_{\bar{\boldsymbol{r}} \in S} \quad F(\bar{\boldsymbol{r}}). \tag{3.20}$$

4 Implementation

As described in the previous section, the localization problem is reduced to forming a collection of optimized beamformers for the computational domain and maximizing on the beamformer output levels, in this section, we discuss the implementation details and summarize the final algorithm.

4.1 Choice of Beamformers

From the previous section, we can see that the conditions in Theorem 3.4 must be satisfied. Thus, the choice of beamformers can not be arbitrary. To formulate the $\bar{\boldsymbol{r}}$-beamformer design problem for each $\bar{\boldsymbol{r}}$, we should define a suitable passband region and the corresponding stopband region. First, for the frequency domain, sound produced by the human vocal tract has the range from about $100Hz$ to $6kHz$ [16]. The frequency domain I_p should cover at least part of this range $[100Hz, 6kHz]$ for practical applications. In fact, the frequency range of male voice is relatively lower while the frequency range of female voice is relatively higher.

For the the spatial domain, it follows by Theorem 3.4 that condition (3.13) must be satisfied for any $\bar{\boldsymbol{r}}$. To satisfy this condition, we need to have a reasonable transition region for all the beamformer design problems. If the transition region is too narrow, the actual response G will become oscillatory at the edges of the passband region as well as the stopband region. The reason is that the actual response G is continuous with respect to $\boldsymbol{r}$ while the desired response G_d is not continuous. Then, $|G|$ should be allowed to transit from 1 to 0 gradually to avoid the performance value (3.12) of the beamformer being greater than or equal to 0.5, which contradicts condition (3.13).

Following Theorem 3.4, the stopband region should be as large as possible in order to achieve good localization accuracy. Also, as a rule of thumb, it is usually not easy to design a good beamformer with a very large passband region. Hence, for any $\bar{\boldsymbol{r}}$, the $\bar{\boldsymbol{r}}$-beamformer should be defined with a relatively small passband region.

4.2 Computation of Cost Function

Since the localization problem is a precursor to track speaker movements, the implementation efficiency is very important. However, the cost function (3.19) is expensive to compute. This is because the computation of the function (3.5) requires the value of frequency response vector $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$. Therefore, for any $\bar{\boldsymbol{r}}$ in S, $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ is obtained by solving a minimax optimization problem (3.8). Hence, the computation of the cost function (3.19) is rather expensive. To simplify the computation of $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$, we choose a sufficiently dense discrete set of points in the spatial domain S, which replaces the original continuous set S. We then only consider the point $\bar{\boldsymbol{r}}$ in a finite grid of S, which can be denoted by $S^d = \{\bar{\boldsymbol{r}}_k : k = 1, \dots n_S\}$, where n_S is the number of points in the grid. For a typical room, an example of S^d can be seen in Figure 1.

Then, we can calculate the optimal solution $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ over the finite set S^d. This will be implemented offline and the optimized solutions can be stored in advance to form different beams in the domain. For a point $\bar{\boldsymbol{r}}$ which is not in S^d, we can use interpolation to calculate the function value $F(\bar{\boldsymbol{r}})$. When the finite set S^d is sufficiently dense, the approximation error should be small.

Since the frequency response function $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ are computed in advance, the main computational effort of the function (3.5) is the computation of $Y_i(\boldsymbol{r}, f)$, which is the discrete Fourier transform of the received data. This can be implemented very efficiently in practice so that the evaluation of the cost function can be achieved in real time.

4.3 Decomposition Method

The problem (3.20) is a continuous optimization problem. We need to develop an efficient and effective method to solve this problem and to implement it in real time. However, this problem is very complicated and there are many local maximizers. This can be seen in the figures in the examples below.

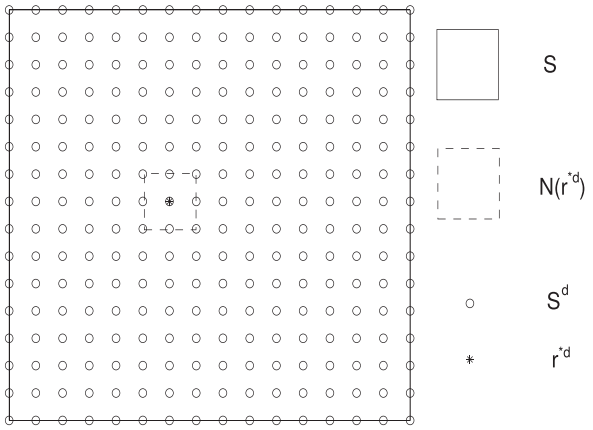

Figure 1: An example of S, S^d, $\boldsymbol{r}^{*d}$ and $N(\boldsymbol{r}^{*d})$.

Note that the size of the spatial domain S is limited and the feasible set is bounded.We can find the solution over the mesh first to obtain an initial approximate solution, and refine the global solution in a small area by a local optimization method. Hence, we can solve this problem by the following algorithm in two steps as follows:

1. Solve the problem (3.20) with $\bar{\boldsymbol{r}}$ constrained in its discrete set S^d. That is,

$$\max_{\bar{\boldsymbol{r}} \in S^d} \quad F(\bar{\boldsymbol{r}}). \tag{4.1}$$

2. Solve the problem (3.20) with the search area constrained to only a small neighbourhood of $\bar{\boldsymbol{r}}^{*d}$, that is

$$\max_{\bar{\boldsymbol{r}} \in N(\bar{\boldsymbol{r}}^{*d})} \quad F(\bar{\boldsymbol{r}}), \tag{4.2}$$

   where $N(\bar{\boldsymbol{r}}^{*d})$ denotes a neighborhood of $\bar{\boldsymbol{r}}^{*d}$.

In the first stage, It can be seen that the optimization problem is discrete. That is, the variables are defined in a discrete set S^d. Since $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ over the finite set S^d is stored in advance, the computation of $F(\bar{\boldsymbol{r}})$ over the finite set is computed very efficiently. The problem can be solved by comparing all the function values in S^d and therefore the maximizer can be found quickly.

In the second stage, after solving the subproblem (4.1), the maximizer is obtained, which can be denoted by $\bar{\boldsymbol{r}}^{*d}$. Then the solution of problem (3.20) is close to the point $\bar{\boldsymbol{r}}^{*d}$. Hence, a local search in a small neighbourhood will be sufficient to refine and pinpoint the final source location. Since the optimization problem is continuous, while the feasible set is very small. General continuous optimization method can be applied to solve this problem.

An example of $N(\bar{\boldsymbol{r}}^{*d})$ can be seen in Figure 1. From a practical point of view, it might not be necessary to obtain a solution with high precision. For example, it take 10 and 100 iterations to obtain the solution with precision less than $10^{-3}m$ and 10^{-4} respectively. Then, it's not necessary to obtain the solution less than 10^{-4}. We only take a few iterations to obtain an acceptable accuracy like precision less than 10^{-3}. Hence, it will be very simple to solve this problem. The total implementation complexity of this method is not high. Therefore, it can be implemented in real time for practical applications.

5 Numerical Examples

In this section, we show how the framework of localization works. We demonstrate that if (3.13) is satisfied, the localization method will work and the accuracy can be controlled as well. Three numerical examples will be given to illustrate the proposed method. The computations were performed in Matlab running on a 2GHz PC with 2G RAM. The software to solve the semi-definite programming problem (3.10) is SDPA 6.0 [15] and the function to find the maximum of the problem (3.20) is *fminicon* in Matlab. The interpolation method for the function $F(\bar{\boldsymbol{r}})$ is the cubic spline interpolation.

To show how the proposed method works, we define the sets $S_p(\bar{\boldsymbol{r}})$ and $S_s(\bar{\boldsymbol{r}})$ for any $\bar{\boldsymbol{r}}$ by

$$\begin{aligned} S_p(\bar{\boldsymbol{r}}) = \{\boldsymbol{r} =& (r_x, r_y, r_z) : r_x \in [\bar{r}_x - h_x, \bar{r}_x + h_x], \\ & r_y \in [\bar{r}_y - h_y, \bar{r}_y + h_y], \\ & r_z \in [\bar{r}_z - h_z, \bar{r}_z + h_z]\}, \end{aligned} \tag{5.1}$$

$$\begin{aligned} S_s(\bar{\boldsymbol{r}}) = \{\boldsymbol{r} =& (r_x, r_y, r_z) : r_x \notin [\bar{r}_x - h_x - d_x, \bar{r}_x + h_x + d_x], \\ & r_y \notin [\bar{r}_y - h_y - d_y, \bar{r}_y + h_y + d_y], \\ & r_z \notin [\bar{r}_z - h_z - d_z, \bar{r}_z + h_z + d_z]\}, \end{aligned} \tag{5.2}$$

where h_x, h_y and h_z are small positive numbers, and d_x, d_y and d_z are three positive numbers. For comparison, we first compare our proposed method with the method in [3], where the frequency response vector $\boldsymbol{H}(f)$ is given by

$$H_i(f) = e^{-j2\pi f \frac{||\boldsymbol{r}-\boldsymbol{r_i}||-||\boldsymbol{r}-\boldsymbol{r_c}||}{c}}. \tag{5.3}$$

This definition of $\boldsymbol{H}(f)$ is very easy for implementation. It's not necessary to solve an optimization problem to obtain the frequency response value. However, the localization accuracy is not ensured, which can be seen in the following examples.

In the first example, we consider a non-reflective case, where an acoustic signal of length 0.2s is used and the received data of the microphones are simulated by the transfer function (2.2). There are 25 microphones, which are placed at the z-coordinate of $1.5m$ and the (x, y)-coordinates $\{(x, y) : x = -2.5 + 1.25k, y = -2.5 + 1.25l; k, l = 0, \ldots, 4\}$. The center point $(0m, 0m)$ is set as the reference point. The placement of the microphone array and the reference point can be seen in Figure 2. The feasible source location field is given by a $5m \times 5m$ region with the z-coordinate 0 and (x, y)-domain $[-2.5m, 2.5m] \times [-2.5m, 2.5m]$.

We choose the discrete grid set of $[-2.5m, 2.5m] \times [-2.5m, 2.5m]$ by taking the points every $0.1m$, that is, the total number of the discrete set is $51 * 51 = 2601$. For the frequency domain, we set $I_p = [0.2kHz, 1kHz]$. For the parameters of the specified domain Ω, we set

$$h_x = h_y = 0.025m, \quad d_x = d_y = 0.4m.$$

For each discrete point $\bar{\boldsymbol{r}}$, we set up the problem (3.8) and solve it by semi-definite programming software. The performance values $E^*(\bar{\boldsymbol{r}})$ can be obtained and are depicted as Figure 3. It can be seen that all these values are less than 0.5 and the condition (3.13) is satisfied. Then, we can use the obtained solution $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ to do the localization.

With these settings, we can formulate problem (4.1) and estimate the source location. For example, set the source location as $(0.8m, 2.25m)$. Then, by apply the proposed method, we can locate the source point as $(0.8034m, 2.2970m)$, where the localization error is about

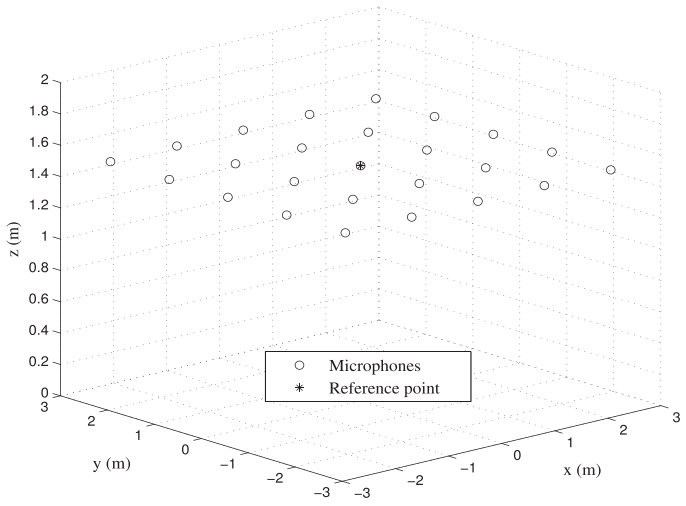

Figure 2: The placement of microphone array and the reference point in the first example.

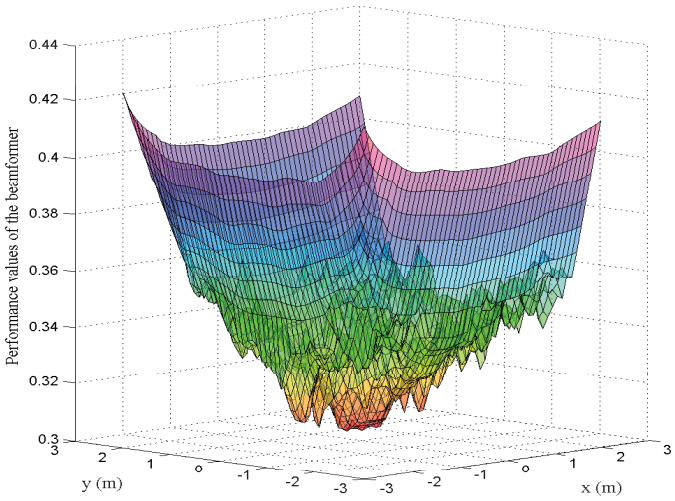

Figure 3: Performance values $E^*(\bar{\boldsymbol{r}})$ (3.12) of the beamformer design problem with the specified domain $\Omega(\bar{\boldsymbol{r}})$ in the first example.

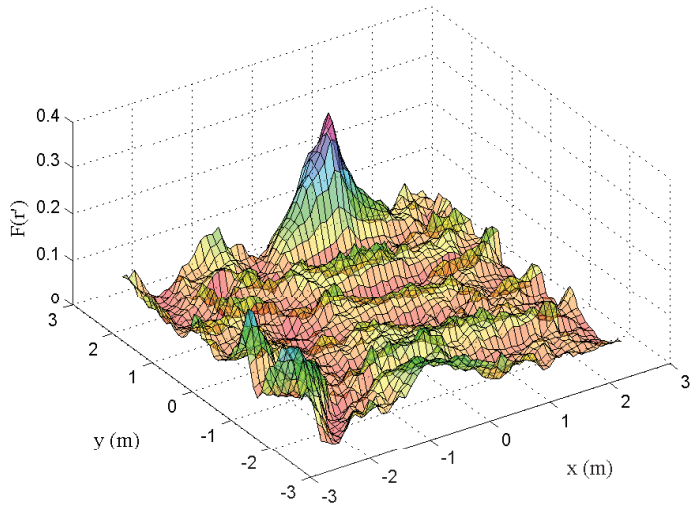

Figure 4: The values $F(\bar{\boldsymbol{r}})$ (3.19) using the proposed method in the first example, where the source location is $(0.8m, 2.25m)$.

$0.0472m$. The function value (3.19) of the discrete grid set can be seen in Figure 4, where the points around $(0.8m, 2.25m)$ have the largest values compared to the other points.

For comparison, we use the the frequency response function defined in (5.3), the F function values (3.19) of the discrete grid set are depicted in Figure 5. It follows from Figure 5 that there exist many peaks and it's difficult to judge whether the maximum point is the real source point. We can find the best point manually, where it's around $(0.6m, 0.7m)$, which is far away from the original source location. Hence, not all beamformers can be employed here.

Next, we randomly choose 100 sample points in $[-2.5m, 2.5m] \times [-2.5m, 2.5m]$ as the source locations by using the function $5 * rand(2, 100) - 2.5$. Then, we apply the proposed method to locate these points. The average localization error is $0.058m$ and the average running time to locate the source point is 1.06 seconds. The efficiency and the effectiveness of the algorithm are acceptable.

In the second example, we consider a reflection case, where an acoustic signal of $1s$ is used and the received data of the microphones are simulated by the transfer function given by [11]. The room structure is $[-2m, 2m] \times [-2m, 2m] \times [0m, 3m]$. We use the reverberation time $T_{60} = 0.2s$ which is larger than the value estimated by the Sabine equation to compute the transfer function. The acoustic signal is received by the microphone array with 25 elements. The microphones are placed with the coordinates which can be seen from Figure 6. The point $(0m, 0m, 2.9m)$ is set as the reference point. The feasible source location field is in the region with the z-coordinate $1.5m$ and (x, y)-domain $[-2m, 2m] \times [-2m, 2m]$.

We choose the discrete grid set of $[-2m, 2m] \times [-2m, 2m]$ by taking the points every $0.1m$, and the setting of h_x ,h_y, d_x, d_y are the same as those given in the first example. Then, for each discrete point $\bar{\boldsymbol{r}}$, we set up the problem (3.8) and solve it by semi-definite programming software. The performance values $E^*(\bar{\boldsymbol{r}})$ of the problem can be obtained and are depicted as Figure 7. It can be seen that all these values are less than 0.5 and the condition (3.13) is satisfied. Then, we use the obtained solution $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ to do the localization.

With these settings, we can formulate the problem (4.1) and estimate the source loca-

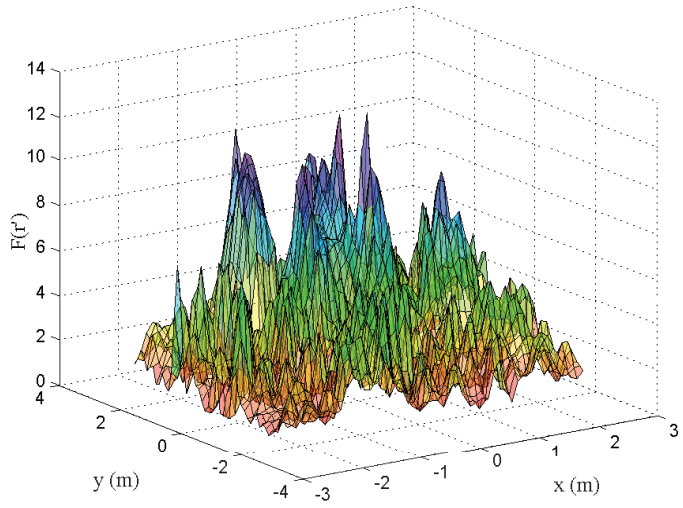

Figure 5: The values $F(\bar{\boldsymbol{r}})$ (3.19) using (5.3) in the first example, where the source location is $(0.8m, 2.25m)$.

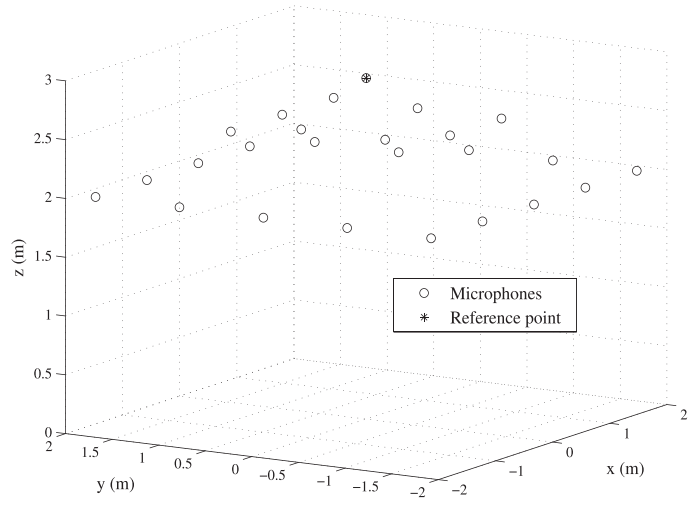

Figure 6: The placement of microphone array and the reference point in the second example.

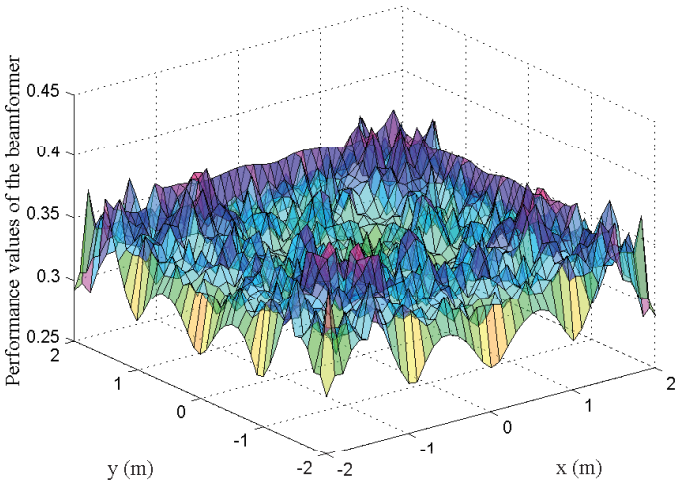

Figure 7: Performance values $E^*(\bar{\boldsymbol{r}})$ (3.12) of the beamformer design problem with the specified domain $\Omega(\bar{\boldsymbol{r}})$ in the second example.

tion. For example, we randomly set the source location as $(-0.9661m, -0.474m)$. Then, by applying the proposed method, for about one second, we can localize the source point as $(-0.9879m, -0.4707m)$, where the localization error is about $0.022m$. The function (3.19) of the discrete grid set can be seen in Figure 8, where the points around $(-0.5m, 1.2m)$ have the largest values compared to the other points.

Next, we randomly choose 100 sample points in $[-2m, 2m]\times[-2m, 2m]$ as source locations by using the function $4 * rand(2, 100) - 2$. Then, we apply the proposed method to locate these points. The average localization error is $0.072m$ and the average running time to locate the source point is 1.21 seconds. The efficiency and the effectiveness of the algorithm are acceptable.

In the third example, we consider another reflective case with stronger reflection parameters of six walls. An acoustic signal of $3s$ is used and the received data of the microphones are simulated by the transfer function given in [11]. The room structure is $[-2.5m, 2.5m] \times [-1.5m, 1.5m] \times [0m, 2.5m]$. We use the reverberation time $T_{60} = 0.45s$ which is larger than the value estimated by the Sabine equation to compute the transfer function. The acoustic signals are received by the microphone array with 21 elements, which are placed at the z-coordinate $2.4m$ and the (x, y)-coordinates $\{(x, y) : x = -2.4+0.8k, y = -1.2+1.2l; k = 0, \ldots, 6, l = 0, \ldots, 2\}$, shown in Figure 9. The point $(0m, 0m, 2.4m)$ is set as the reference point. The feasible source location field is in the region with the z-coordinate $1m$ and (x, y)-domain $[-2.5m, 2.5m] \times [-1.5m, 1.5m]$. We choose the discrete grid set of $[-2.5m, 2.5m] \times [-1.5m, 1.5m]$ by taking the points every $0.1m$, and the setting of h_x ,h_y, d_x, d_y are the same as those given in the first example.

Similarly, we can set up the problem (3.8) for each discrete point $\boldsymbol{r}'$, and solve it by semi-definite programming software. The performance values $E^*(\bar{\boldsymbol{r}})$ of the problem can be obtained and are depicted in Figure 10. It can be seen that there are a few points with the performance values larger than 0.5. Then, the condition (3.13) is not satisfied completely. To demonstrate the effect of this condition, we still use the obtained solution $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ to do the localization.

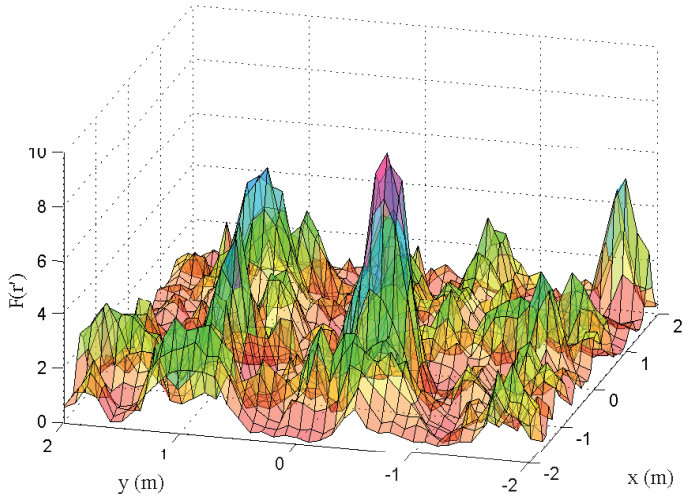

Figure 8: The $F(\bar{\boldsymbol{r}})$ (3.19) using proposed method in the second example, where the source location is $(-0.5m, 1.2m)$.

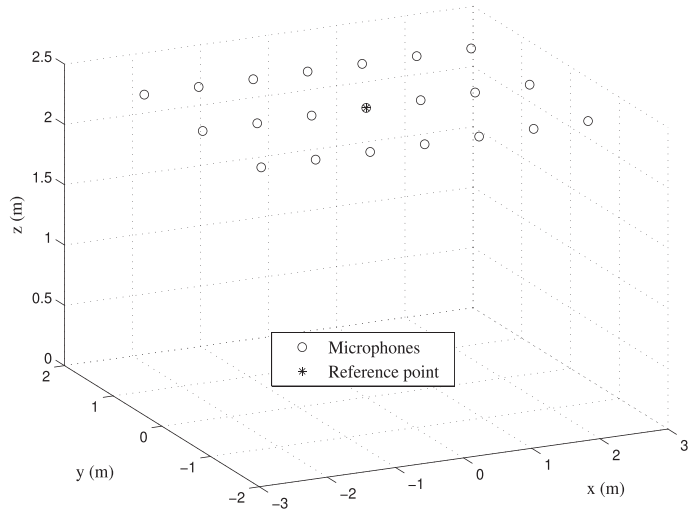

Figure 9: The placement of microphone array and the reference point in the third example.

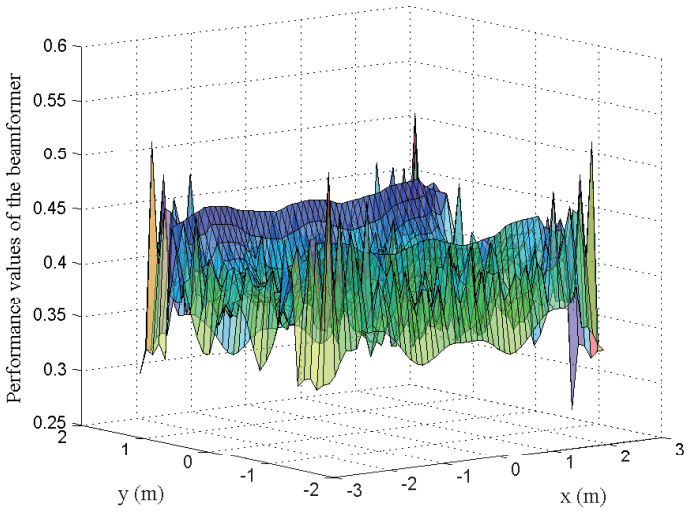

Figure 10: Performance values $E^*(\bar{\boldsymbol{r}})$ (3.12) of the beamformer design problem with the specified domain $\Omega(\bar{\boldsymbol{r}})$ in the third example, where the number of microphones is 21.

By randomly choosing 100 sample points in $[-2.5m, 2.5m] \times [-1.5m, 1.5m]$ as the source locations uniformly, we locate these points with the proposed method. We find that there are four cases with larger localization errors. For example, when the source location is $(-0.9027m, 0.7952m)$, we localize the point as $(-2.2998m, -1.3119m)$, which is $2.5283m$ away from the true source point. For the other 96 source points, the average localization error is $0.051m$. The reason for the larger errors is because the condition (3.13) is not fully satisfied. To overcome it, we increment the microphone number to 25 by adding 4 more microphones. The 4 additional microphones are placed at the z-coordinate $2.4m$ and the (x, y)-coordinates $\{(x, y) : x = \pm 2.4m, y = \pm 0.6m\}$. Then, by formulating the problem (3.8) for each discrete point $\bar{\boldsymbol{r}}$, we solve $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$. The performance values $E^*(\bar{\boldsymbol{r}})$ of the problem are depicted as Figure 11. It can be seen that all the performance values are now less than 0.5. Then, the condition (3.13) is satisfied.

Next, the obtained solution $\boldsymbol{H}(\bar{\boldsymbol{r}}, f)$ can be used to do the localization. By randomly choosing 100 sample points in $[-2.5m, 2.5m] \times [-1.5m, 1.5m]$ as the source locations uniformly, we can now locate all these points with the proposed method. The average localization error is $0.047m$ and the average running time is 1.57 seconds.

6 Conclusion

In this paper, we have developed a new framework for the acoustic source localization. By formulating a series of beamformer design problems with specified passband regions and stopband regions within a room, the localization problem can be formulated as a global optimization problem of a two-dimensional function. In order to implement the method in real-time, the frequency response vectors corresponding to a set of grid points are computed in advance and stored. We have also provided several numerical examples to demonstrate the efficiency and effectiveness of the proposed method.

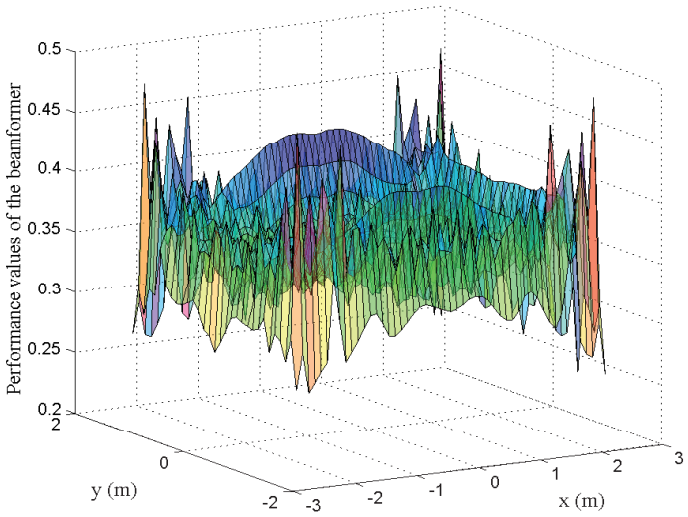

Figure 11: Performance values $E^*(\bar{\boldsymbol{r}})$ (3.12) of the beamformer design problem with the specified domain $\Omega(\bar{\boldsymbol{r}})$ in the third example, where the number of microphones is 25.

References

[1] J.B. Allen and D.A. Berkeley, Image method for efficiently simulating small-room acoustics, *Journal of the Acoustical Society of America* 65 (1979) 943–950.

[2] M.S. Brandstein, J.E. Adcock and H.F. Silverman, A closed form location estimator for use with room environment microphone arrays, *IEEE Trans. Speech and Audio Processing* 5 (1997) 45–50.

[3] M. Brandstein and D.B. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*, Springer-Verlag, Berlin, 2001.

[4] Y.T. Chan and K.C. Ho, A simple and efficient estimator for hyperbolic location, *IEEE Trans. Signal Processing* 42 (1994) 1905–1915.

[5] J.W. Choi and Y.H. Kim, Spherical beam forming and MUSIC methods for the estimation of location and strength of spherical sound sources, *Mechanical Systems and Signal Processing* 9 (1995) 569–588.

[6] J.P. Dmochowski, J. Benesty and S. Affes, A generalized steered response power method for computationally viable source localization, *IEEE Trans. Audio, Speech and Lang. Proc.* 15 (2007) 2510–2526.

[7] Z.G. Feng, K.F.C. Yiu and S. Nordholm, A two-stage method for the design of near-field multi-dimensional broadband beamformer, *IEEE Trans. Signal Processing* 59 (2011) 3647–3656.

[8] Z.G. Feng, K.F.C. Yiu and S. Nordholm, Placement design of microphone arrays in near-field broadband beamformers, *IEEE Trans. Signal Processing* 60 (2012) 1195–1204.

[9] Z.G. Feng and K.F.C. Yiu, The design of multi-dimensional acoustic beamformers via window function, *Digital Signal Processing* 29 (2013) 107–116.

[10] K.-W. Kim, D.-H. Seo, J. Chang and Y.-H. Kim, Sound source localization using scattered acoustic pressure on the surface of rigid sphere and its performance, *Proceedings of 20th International Congress on Acoustics 2010*, Sydney, Australia 2010.

[11] E.A. Lehmann and A.M. Johansson, Diffuse reverberation model for efficient image-source simulation of room impulse responses, *IEEE Trans. Audio, Speech and Lang. Proc.* 18 (2010) 1429–1439.

[12] Z.B. Li and K.F.C. Yiu, A least-squares indoor beamformer design, *Pacific Journal of Optimization* 9 (2013) 697–707.

[13] Z.B. Li, K.F.C. Yiu and S. Nordholm, On the indoor beamformer design with reverberation, *IEEE Transactions on Audio, Speech and Language Processing* 22 (2014) 1225–1235.

[14] W. Liu and S. Weiss, *Wideband Beamforming*, John Wiley and Sons, Chichester, 2010.

[15] M. Yamashita K. Fujisawa and M. Masakazu, Implementation and evaluation of SDPA 6.0 (SemiDefinite Programming Algorithm 6.0), *Optimization Methods and Software* 18 (2003) 491–505.

[16] H. Silverman, Some analysis of microphone arrays for speech data acquisition, *IEEE Trans. Acoustics, Speech and Signal Processing* ASSP-35 (1987) 1699–1712.

[17] B.D. Van Veen and K.M. Buckley, Beamforming: A versatile approach to spatial filtering, *IEEE ASSP Mag.* 5 (1988) 4–24.

[18] D.B. Ward, E.A. Lehmann and R.C. Williamson, Particle filtering algorithms for tracking an acoustic source in a reverberant environment, *IEEE Trans. Speech and Audio Processing* 11 (2003) 826–836.

[19] K.F.C Yiu, N. Grbic, K.L. Teo and S. Nordholm, A new design method for broadband microphone arrays for speech input in automobiles, *IEEE Signal Processing Letters* 9 (2002) 222–224.

[20] D.N. Zotkin and R. Duraiswami, Accelerated speech source localization via a hierarchical search of steered response power, *IEEE Trans. Speech and Audio Processing* 12 (2004) 499–508.

Manuscript received 5 March 2014
revised 7 July 2014, 7 October 2014
accepted for publication 25 November 2014

Zhiguo Feng
College of Mathematics, Chongqing Normal University
Chongqing, China and
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hunghom, Kowloon, Hong Kong China
E-mail address: 18281102@qq.com

Ka-fai Cedric Yiu
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hunghom, Kowloon, Hong Kong, China
E-mail address: macyiu@polyu.edu.hk

Randolph Chi-Kin Leung
Department of Mechanical Engineering
The Hong Kong Polytechnic University
Hunghom, Kowloon, Hong Kong, China
E-mail address: mmrleung@polyu.edu.hk

Sven Nordholm
Department of Electrical and Computer Engineering
Curtin University, Perth, Australia
E-mail address: S.Nordholm@curtin.edu.au