Applied Soft Computing 13 (2013) 1486–1490

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A hybrid descent method with genetic algorithm for microphone array placement design☆

Zhibao Li[a], Ka Fai Cedric Yiu[b,*], Zhiguo Feng[c]

[a] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[b] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[c] College of Mathematics, Chongqing Normal University, Chongqing, PR China and the Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

ARTICLE INFO

Article history:
Received 31 December 2011
Accepted 13 February 2012
Available online 8 April 2012

Keywords:
Beamformer design
Microphone array
Microphone placement
Genetic algorithm
Hybrid descent method

ABSTRACT

In beamformer design, the microphone locations are often fixed and only the filter coefficients are varied in order to improve on the noise reduction performance. However, the positions of the microphone elements play an important role in the overall performance and should be optimized at the same time. However, this nonlinear optimization problem is non-convex and local search techniques might not yield the best result. This problem is addressed in this paper. A hybrid descent method is proposed which consists of a genetic algorithm together with a gradient-based method. The gradient-based method can help to locate the optimal solution rapidly around the start point, while the genetic algorithm is used to jump out from local minima. This hybrid method has the descent property and can help us to find the optimal placement for better beamformer design. Numerical examples are provided to demonstrate the effectiveness of the method.

1. Introduction

Broadband beamformers have been studied extensively due to their wide applications in wireless communications, sonar, radar, speech and acoustics [1–4]. There are many model-based approaches developed today to design the beamformer filter coefficients in the literature [5–11]. Besides the optimal filter coefficients, the placement of microphone array elements also plays an important role in the overall performance of the beamformers and different microphone array configurations perform significantly differently. Therefore, finding a good placement for the microphone array configuration is important in enhancing the performance of the broadband beamformers.

The microphone array placement problem has been addressed partially in the array thinning technique [12–19], which attempts to reduce the number of elements and adjust the positions of the remaining elements to retain performance. Since this problem is nonlinear, various global optimization methods have been developed, including evolutionary programming [20,21], genetic algorithm [22,23], simulated annealing algorithm [24–26] and pattern search algorithm [27]. However, the array thinning technique is essentially one-dimensional and is more suited for antenna design. In formulating the multi-dimensional design problem, a nonlinear optimization problem under the l_2-norm was proposed in [28], which allow the microphones to move around in a multi-dimensional solution space in search of better performances. However, the objective function is highly nonlinear and is nonconvex with respect to the placement variables, it is therefore hard to tackle with traditional gradient-based methods. The problem is complicated further by the influence of different filter lengths, in which optimal designs might vary for different filter lengths. This problem can be avoided by considering the performance limit when the filter length is sufficiently long. By taking the limit of the filter length to infinity, it was shown in [28] that the filter coefficients are defined by a set of reduced one-dimensional convex optimization problems and the coefficients can be sought efficiently for a particular array configuration. By treating this as a subproblem, we can reduce the original mixed optimization problem to the placement problem with microphone locations being the only decision variables.

In this paper, we will study the optimal placement problem in two dimensions and propose a hybrid descent method [29] to handle the nonconvexness. The proposed hybrid descent method consists of a genetic algorithm combining with a gradient-based

☆ This work was supported by RGC Grant PolyU. (5365/09E) and the Research Committee of the Hong Kong Polytechnic University. The second and third authors would also like to thank the support of NSFC Grant (10901170).
* Coresponding author.
E-mail addresses: zbli0307@163.com (Z. Li), macyiu@polyu.edu.hk (K.F. Cedric Yiu), z.feng.scholar@gmail.com (Z. Feng).

1568-4946/$ – see front matter
doi:10.1016/j.asoc.2012.02.027

method. The literature in genetic algorithm is vast and it has been employed for beamformer design recently [30] to enhance speech recognition accuracy. Since the formulated nonlinear optimization problem will have many local minima, we employ the genetic algorithm here to generate the new feasible solutions in order to jump out from the local minima while retaining previous useful information. Once a descent point is identified, the gradient-based method is executed to descend to the nearest local minimum rapidly. This method has a nice descent property and the objective function is improved monotonically. Numerical experiments will be given to demonstrate that much better placements can be found.

The rest of this paper is organized as follows. In Section 2, we formulate the placement problem and discuss the performance limit of the filter length. In Section 3, we propose the hybrid descent method with a genetic algorithm for solving the placement problem. A two-dimensional numerical example will be used to demonstrate the proposed method and the performance of the designed beamformer will be given and compared in Section 4.

2. Problem formulation

For an N elements microphone array, assume each element is an L-tap finite impulse response (FIR) filter. Define the microphone locations to be $\boldsymbol{r}_i$, $i = 1, 2, \ldots, N$, the transfer function of the ith microphone is given by

$$A_i(\boldsymbol{r}, f) = \frac{1}{\|\boldsymbol{r} - \boldsymbol{r}_i\|} e^{(-j2\pi f \|\boldsymbol{r}-\boldsymbol{r}_i\|)/c}, \tag{2.1}$$

where $\boldsymbol{r}$ is the location of the sound source and c is the speed of sound in the air. If the signals received by this microphone array are sampled synchronously at the rate of f_s per second, the frequency responses of these FIR filters are

$$H_i(\boldsymbol{h}, f, L) = \boldsymbol{h}_i^T \boldsymbol{d}_0(f), \quad i = 1, 2, \ldots, N, \tag{2.2}$$

where $\boldsymbol{h}_i$ is the coefficients of the ith FIR filter, and $\boldsymbol{d}_0(f)$ is the Kronecker delta frequency response, defined as

$$\boldsymbol{h}_i = [h_i(0), h_i(1), \ldots, h_i(L-1)]^T,$$

$$\boldsymbol{d}_0(f) = \left[1, e^{-j2\pi f/f_s}, \ldots, e^{-j2\pi f(L-1)/f_s}\right]^T.$$

Suppose the desired response is $G_d(\boldsymbol{r}, f, L)$, then the FIR filter broadband beamformer design problem is to find a set of coefficients $\boldsymbol{h} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_N]^T$ for the FIR filters, such that the beamformer output

$$G(\boldsymbol{r}, f, L) = \sum_{i=1}^{N} H_i(\boldsymbol{h}, f, L) A_i(\boldsymbol{r}, f) = \boldsymbol{A}^T(\boldsymbol{r}, f) \boldsymbol{H}(\boldsymbol{h}, f, L) \tag{2.3}$$

is sufficiently close to $G_d(\boldsymbol{r}, f, L)$, where $\boldsymbol{A}^T(\boldsymbol{r}, f) = [A_1(\boldsymbol{r}, f), \ldots, A_N(\boldsymbol{r}, f)]^T$ is the transfer function vector defined in (2.1), and $\boldsymbol{H}(\boldsymbol{h}, f, L) = [H_1(\boldsymbol{h}, f, L), \ldots, H_N(\boldsymbol{h}, f, L)]^T$ is the frequency filter response vector defined in (2.2).

There are many objective criteria exist to evaluate the error between $G(\boldsymbol{r}, f, L)$ and $G_d(\boldsymbol{r}, f, L)$, such as the l_∞-norm (for example [31,32]) and the l_1-norm minimax (for example [33,11]). However, these design techniques are often more expensive than the use of an l_2-norm. In the following, we adopt the l_2-norm criterion for the design of broadband beamformer. For a given placement of the microphone array, the beamformer design problem can be formulated to find the filter coefficients $\boldsymbol{h}$ such that the objective function

$$E(\boldsymbol{h}) = \frac{1}{\|\Omega\|} \int_{\Omega} \rho(\boldsymbol{r}, f) \|\boldsymbol{A}^T(\boldsymbol{r}, f) \boldsymbol{H}(\boldsymbol{h}, f, L) - G_d(\boldsymbol{r}, f, L)\|^2 \, d\boldsymbol{r} \, df, \tag{2.4}$$

is minimized, where Ω is a specified spatial-frequency domain as the definition field of $G_d(\boldsymbol{r}, f, L)$, and $\rho(\boldsymbol{r}, f)$ is a positive weighting function. Usually, the domain $\Omega = \Omega_p \bigcup \Omega_s$ is consisting of the passband region Ω_p and stopband region Ω_s. Then, for a given placement of the microphone array, the beamformer design problem is formulated as

$$\min_{\boldsymbol{h} \in \mathbb{R}^{N \times L}} E(\boldsymbol{h}). \tag{2.5}$$

In fact, the optimal solution of (2.5) depends on the filter length, which can be denoted by $\boldsymbol{h}^*(L)$. We define the performance limit as $\inf_L E(\boldsymbol{h}^*(L))$, which is the infimum of the objective function (2.4) for all filter lengths and all filter coefficients.

The direct numerical computation of $\inf_L E(\boldsymbol{h}^*(L))$ is impossible, since the filter length can approach to infinity. However, it follows from [34] that in the case of

$$G_d(\boldsymbol{r}, f, L) = e^{-j2\pi f \tau_L / f_s} \hat{G}_d(\boldsymbol{r}, f), \tag{2.6}$$

where $\tau_L \in [0, L-1]$ is the group delay satisfying $\lim_{L\to\infty} \tau_L = \infty$ and $\lim_{L\to\infty} L - \tau_L = \infty$, and $\hat{G}_d(\boldsymbol{r}, f)$ is a response function independent of filter length L, the performance limit $\inf_L E(\boldsymbol{h}^*(L))$ can be computed by solving

$$\min_{\tilde{\boldsymbol{H}} \in \Gamma^N} E(\tilde{\boldsymbol{H}}), \tag{2.7}$$

where

$$E(\tilde{\boldsymbol{H}}) = \frac{1}{\|\Omega\|} \int_{\Omega} \rho(\boldsymbol{r}, f) \|\boldsymbol{A}^T(\boldsymbol{r}, f) \tilde{\boldsymbol{H}}(f) - \hat{G}_d(\boldsymbol{r}, f)\|^2 \, d\boldsymbol{r} \, df, \tag{2.8}$$

and Γ is given by

$$\begin{aligned} \Gamma = \{ & u(f) + jv(f) : u(f) \text{ and } v(f) \text{ are continuous, absolute integrable,} \\ & \text{and the right-hand and left-hand derivatives exist, } v(0) = 0, \\ & v(f_s/2) = 0\}. \end{aligned} \tag{2.9}$$

Note that the performance limit is computed for a given placement of the microphone array. Different placements will yield different performance limits. Thus the placement design is to find the placement of the microphone array such that the corresponding performance limit is minimized. To formulate this properly, let $\boldsymbol{\lambda} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \in \Lambda \subset \mathbb{R}^{3\times N}$ denote the microphone locations, where Λ is the set of all possible $\boldsymbol{\lambda}$, the placement problem is given by

$$\begin{aligned} &\min_{\boldsymbol{\lambda} \in \Lambda, \tilde{\boldsymbol{H}}^N \in \Gamma^N} && E(\boldsymbol{\lambda}, \tilde{\boldsymbol{H}}) \\ &s.t && \|\boldsymbol{r}_i - \boldsymbol{r}_j\|^2 \geq \varepsilon_d, \ldots i, j = 1, 2, \ldots, N, i \neq j, \end{aligned} \tag{2.10}$$

where

$$E(\boldsymbol{\lambda}, \tilde{\boldsymbol{H}}) = \frac{1}{\|\Omega\|} \int_{\Omega} \rho(\boldsymbol{r}, f) \|\boldsymbol{A}^T(\boldsymbol{\lambda}, \boldsymbol{r}, f) \tilde{\boldsymbol{H}}(f) - \hat{G}_d(\boldsymbol{\lambda}, \boldsymbol{r}, f)\|^2 \, d\boldsymbol{r} \, df \tag{2.11}$$

is the same as (2.8), but considering $\boldsymbol{\lambda}$ as the new decision vector. The constraints $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|^2 \geq \varepsilon_d$, $i, j = 1, 2, \ldots, N$, $i \neq j$, are for practicality so that the set of microphone elements should be kept at least a certain minimum distance mutually for proper functioning, and ε_d is the square of the minimum distance between two different microphone elements.

3. A hybrid descent method

In general, the combined nonconvex optimization problem (2.10) has two kinds of variables, it is difficult to solve as a whole. Note that the objective function (2.11) is nonconvex and highly nonlinear with respect to the placement variables $\boldsymbol{\lambda}$, but it is convex with respect to the filter coefficient variables $\tilde{\boldsymbol{H}}$. For a given