Applied Acoustics 122 (2017) 146–151

Contents lists available at ScienceDirect

Applied Acoustics

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

On the design of time-domain implementation structure for steerable spherical modal beamformers with arbitrary beampatterns

Weiyi Ren, Huawei Chen *, Weixia Gao

College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

ARTICLE INFO

Article history:
Received 19 December 2016
Received in revised form 21 February 2017
Accepted 22 February 2017
Available online 7 March 2017

Keywords:
Spherical modal beamforming
Microphone array
Broadband beamforming
Beam steering
Time-domain structure

ABSTRACT

It is noted that the existing time-domain implementation structure for steerable spherical modal beamformers is only applicable to the specific beamformers with rotationally symmetric beampatterns about the look direction, which may limit its applications. To overcome the restriction, this paper presents an alternative time-domain implementation structure for spherical modal beamformers using the Wigner-D function, which enables three-dimensional beam steering with arbitrary patterns. In particular, a necessary condition for guaranteeing only real-valued operations to make the time-domain implementation viable is derived. Design examples are presented to demonstrate the effectiveness of the presented time-domain modal beamformer structure.

1. Introduction

Spherical microphone arrays are able to analyze three-dimensional (3-D) sound fields effectively and facilitate array signal processing in the spherical harmonics domain, thus have found a variety of applications such as sound field reproduction, beamforming, sound localization, and noise reduction, acoustic absorption measurement, among others [1–5]. Spherical array beamforming in the spherical harmonics domain, also known as spherical modal beamforming, offers several advantages when compared with conventional beamforming in the element-space domain [6]. One of the advantages of spherical modal beamforming is that beampattern design can be decoupled from beampattern steering in the spherical harmonics domain, which results in efficient implementation of steerable modal beamformers in 3-D space.

Spherical modal beamformers can be implemented either in the frequency domain or in the time domain. In comparison to the time-domain implementation, the frequency-domain implementation is usually block-based processing with the discrete Fourier transform and it may not be suitable for time-critical speech and audio applications due to its associated time delay [7]. It is noted that the frequency-domain implementation approaches for beam steering of spherical modal beamformers with arbitrary beampatterns have been available in the literature [8,9]. In contrast, however, the existing time-domain implementation approach for steerable spherical modal beamformers is only applicable to the specific beamformers, i.e., the beampatterns should be rotationally symmetric [7], which may limit its applications. By rotationally symmetric, it means that the beampattern is rotationally invariant with respect to the look direction, i.e., the beampattern will not change when rotated at an arbitrary angle along the look direction. In contrast, non-rotationally symmetric implies that the beampattern will change when rotated along the look direction. In some practical applications, however, a non-rotationally symmetric beampattern may be desired. For instance, as noted in [8], the recording of sound sources of interest in an auditorium with a spherical array placed at the seating area requires a mainlobe that is wide along the azimuth dimension but narrow along the elevation dimension to cover the entire stage. Inspired by the Wigner-D function [6], a time-domain implementation structure for spherical modal beamformers which enables 3-D beam steering with arbitrary patterns is developed in this paper. In particular, a necessary condition for guaranteeing only real-valued operations to make the time-domain implementation viable is also derived. The advantage of the proposed structure is that it is applicable to not only rotationally symmetric beampatterns but also to non-rotationally symmetric beampatterns in the time-domain implementation.

2. Spherical modal beamforming

The standard spherical (r, θ, ϕ) coordinate system is used hereafter, where θ and ϕ denote elevation and azimuth angles, respectively [10]. Consider a unit magnitude plane wave with wave

* Corresponding author.
E-mail address: hwchen@nuaa.edu.cn (H. Chen).

http://dx.doi.org/10.1016/j.apacoust.2017.02.013
0003-682X/

number k impinging on a sphere of radius a from direction $\Omega = (\theta, \phi)$, where $k = 2\pi f/c$ with f denoting the frequency and c the speed of sound. Then, the frequency-domain expression of the sound pressure at a point $\Omega_s = (\theta_s, \phi_s)$ on the sphere surface can be expressed as [11]:

$$p(ka, \Omega, \Omega_s) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} b_n(ka) Y_n^{m*}(\Omega) Y_n^m(\Omega_s) \tag{1}$$

$$p_{nm}(ka, \Omega) = b_n(ka) Y_n^{m*}(\Omega) \tag{2}$$

where $Y_n^m(\Omega)$ is the spherical harmonic of order n and degree m, which is defined as follows [13]:

$$Y_n^m(\Omega) = Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{jm\phi} \tag{3}$$

where $(\cdot)!$ denotes the factorial function, $P_n^m(\cdot)$ are the associated Legendre functions, and $j = \sqrt{-1}$. The superscript * denotes complex conjugation. The mode strength $b_n(ka)$ for order n is related to both frequencies and sphere configurations [1]. $p_{nm}(ka, \Omega)$ are the spherical harmonic coefficients of $p(ka, \Omega, \Omega_s)$ which are obtained by performing the spherical Fourier transform [12]:

$$p_{nm}(ka, \Omega) = \int_{\Omega_s \in S^2} p(ka, \Omega, \Omega_s) Y_n^{m*}(\Omega_s) d\Omega_s \tag{4}$$

where the integral $\int_{\Omega_s \in S^2} d\Omega_s = \int_0^{2\pi} \int_0^{\pi} \sin\theta_s d\theta_s d\phi_s$covers the entire surface of the unit sphere S^2. In practice, since the number of microphones mounted on a sphere is usually limited, the microphone positions are required to satisfy the following discrete orthonormality condition

$$\sum_{s=1}^{M} \alpha_s Y_n^{m*}(\Omega_s) Y_{n'}^{m'}(\Omega_s) = \delta_{n-n'} \delta_{m-m'} \tag{5}$$

where $\delta_{n-n'}$ and $\delta_{m-m'}$ are the Kronecker delta functions, α_s are real-valued parameters depending on the spatial sampling scheme, and M denotes the number of microphones. Then, (4) can be approximated by its discrete version

$$p_{nm}(ka, \Omega) = \sum_{s=1}^{M} \alpha_s p(ka, \Omega, \Omega_s) Y_n^{m*}(\Omega_s). \tag{6}$$

Accordingly, the array order is limited to N which satisfies $(N+1)^2 \leqslant M$ [14].

By the discrete spherical Fourier transform for the sound pressure samples, the beampattern which is the array's response to a unit input signal from Ω can be expressed in the spherical harmonics domain as [14]:

$$B(f, \Omega) = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm}(ka, \Omega) w_{nm}^*(f) \tag{7}$$

where $w_{nm}(f)$ are the spherical Fourier coefficients of the array weights $w(f, \Omega_s)$.

3. Main results

The 3-D rotation of the beampattern $B(f, \Omega)$ can be achieved by using the Wigner-D function and the rotated beampattern can be expressed as [8,6]

$$\begin{aligned} B^r(f, \Omega) &= \Lambda(\alpha, \beta, \gamma) B(f, \Omega) \\ &= \sum_{n=0}^{N} \sum_{m=-n}^{n} b_n(ka) \sum_{m'=-n}^{n} D_{m'm}^{n*}(\alpha, \beta, \gamma) Y_n^{m'*}(\Omega) w_{nm}^*(f) \\ &= \sum_{n=0}^{N} \sum_{m'=-n}^{n} b_n(ka) Y_n^{m'*}(\Omega) w_{nm'}^{r*}(f) \end{aligned} \tag{8}$$

where $w_{nm'}^r(f) = \sum_{m=-n}^{n} D_{m'm}^n(\alpha, \beta, \gamma) w_{nm}(f)$ are the rotated array weights, $\Lambda(\alpha, \beta, \gamma)$ denotes the rotation operation, (α, β, γ) represents the rotation angle, and $D_{m'm}^n(\alpha, \beta, \gamma)$ is the Wigner-D function defined as

$$D_{m'm}^n(\alpha, \beta, \gamma) = e^{-jm'\alpha} d_{m'm}^n(\beta) e^{-jm\gamma} \tag{9}$$

with $d_{m'm}^n(\beta)$ denoting the real-valued Wigner-d function which can be written in terms of the Jacobi polynomial [15]:

$$d_{m'm}^n(\beta) = \varsigma_{m'm} \sqrt{\frac{s!(s+\mu+\upsilon)!}{(s+\mu)!(s+\upsilon)!}} [\sin(\beta/2)]^{\mu} [\cos(\beta/2)]^{\upsilon} P_s^{(\mu,\upsilon)}(\cos\beta) \tag{10}$$

where $\mu = |m'-m|, \upsilon = |m'+m|, s = n-(\mu+\upsilon)/2$, and $\varsigma_{m'm} = 1$ when $m \geqslant m', \varsigma_{m'm} = (-1)^{m-m'}$ when $m < m'$.

3.1. Proposed time-domain beamformer structure

According to (8) and [8], the structure for a steerable spherical modal beamformer usually consists of three stages: modal transformation, beam steering, and pattern forming. Denote $x_s(l) = x_s(t)|_{t=lT_s}$ as the discrete-time series received at the sth microphone, where T_s denotes the sampling interval. Note that the spherical harmonics are independent of frequency. Performing the spherical Fourier transform to $x_s(l)$ yields the modal transformation of $x_s(l)$

$$x_{nm'}(l) = \sum_{s=1}^{M} \alpha_s x_s(l) Y_n^{m'*}(\Omega_s) \tag{11}$$

Denote the real and imaginary parts of $x_{nm'}(l)$ as

$$\tilde{x}_{nm'}(l) = \sum_{s=1}^{M} \alpha_s x_s(l) \mathrm{Re}[Y_n^{m'}(\Omega_s)] \tag{12}$$

$$\overset{smile}{x}_{nm'}(l) = \sum_{s=1}^{M} \alpha_s x_s(l) \mathrm{Im}[Y_n^{m'}(\Omega_s)] \tag{13}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ stand for the real and imaginary parts, respectively. Then, (11) can be expressed as

$$x_{nm'}(l) = \tilde{x}_{nm'}(l) - j\overset{smile}{x}_{nm'}(l) \tag{14}$$

With beam steering in the spherical harmonics domain using the Wigner-D function, $x_{nm'}(l)$ now becomes

$$x_{nm}^r(l) = \sum_{m'=-n}^{n} D_{m'm}^{n*}(\alpha, \beta, \gamma) x_{nm'}(l) \tag{15}$$

In frequency-domain implementation for spherical modal beamformers, a set of complex-valued array weights $w_{nm}(f)$ are employed in the pattern forming stage. In contrast, for time-domain implementation, the complex-valued array weights $w_{nm}(f)$ are replaced by a bank of finite impulse response (FIR) filters with real-valued coefficients such that

$$w_{nm}^*(f) = \mathbf{h}_{nm}^T \mathbf{e}(f) \tag{16}$$

where $\mathbf{h}_{nm} = \left[h_{nm}^1, h_{nm}^2, \ldots, h_{nm}^L\right]^T$ is the impulse response of the FIR filter corresponding to the spherical harmonics of order n and degree m, the superscript T denotes the transpose, and L is the tap length of each FIR filter. And $\mathbf{e}(f) = \left[1, e^{-j2\pi f T_s}, \ldots, e^{-j(L-1)2\pi f T_s}\right]^T$. Accordingly, the output time series of the beam-steered beamformers can be expressed as