Applied Acoustics 140 (2018) 188–197

Contents lists available at ScienceDirect

Applied Acoustics

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

A constrained total least squares calibration method for distributed microphone array

Rong Wang, Zhe Chen, Fuliang Yin*

School of Information and Communication Engineering, Dalian University of Technology, Dalian 116023, China

ARTICLE INFO

Keywords:
Calibration
Multidimensional scaling
Total least squares
Distributed microphone array

ABSTRACT

Microphone positions have to be calibrated in distributed microphone array applications. A constrained total least squares calibration method for distributed microphone arrays is proposed in this paper. All the source event positions are first estimated by the weighted multidimensional scaling algorithm. Then the suitable source events are picked up by the TDOA selection strategy at each node. Finally, the node microphones are calibrated by the total least squares, and further refined based on constrained total least squares when the estimated results suffer from large errors. The proposed method can obtain higher calibration accuracy and works well in noise and reverberation conditions. Simulation and real-world experiment results reveal the validity of the proposed method.

1. Introduction

Distributed microphone array (DMA) is a promising approach for acoustic capture and processing system [1]. Due to the unconstrained network structure and flexible deployment, DMA has been widely used in video conference systems, speech enhancement, sound source localization, speaker tracking, security monitor, and sniper detection, etc. [2–8]. Generally, the nodes in DMA are often randomly placed and may even vary over time, so the node microphone positions have to be calibrated in advance in DMA applications [9].

A number of node microphone calibration methods in the DMA have been proposed. According to the artificial calibration signal involving or not, the existing DMA calibration methods can be divided into two categories: the passive methods and the active ones [7]. The passive calibration method only uses the existing ambient signals, but is usually very time consuming and has high computational cost [10–14]. For the active calibration, as integrating assistant devices, e.g. one loudspeaker or radio unit in each node, it can effectively simplify the calibration work. In fact, it is also common in some consumer devices like laptops, smart phones, and experimental sensor nodes that all have at least one loudspeaker and one microphone in a known relative distance [15,16]. For instance, Raykar et al. assumed that each node has at least one microphone and one loudspeaker with the known node microphones number, and used maximum likelihood (ML) technique to estimate Time of Flight (TOF) [15]. Haddad et al. used least squares (LS) method to localize mobile device itself with additional loudspeakers [16]. Zhayida et al. used only Time of Arrival (TOA) measurements between receivers and sound sources with unknown positions to calibrate node positions [17], inferred the structure from distance measurements by audio and radio. Anwar et al. used a mobile beacon transmitting acoustic and RF signals simultaneously to obtain the Time Difference of Arrival (TDOA) and Angle of Arrival (AOA), and calibrated the node microphones by ML technique [18]. Usually, the node calibration methods with assistant devices can obtain measurement parameters conveniently, but have limited application scenarios due to the fact that two kinds of specific signals were exploited or additional loudspeaker was equipped. Besides, the signals transmitting from the nodes lead to a great energy consumption.

Due to simple hardware structure and energy efficiency, the nodes without assistant devices in DMA are preferred in many cases. The relevant calibration methods were studied by some researchers [19–22]. Valente et al. used a probing acoustic source and maximum likelihood estimation (MLE) to calibrate the unknown nodes configured some microphones regularly [19]. Khanal et al. employed a transducer source with a microphone attached above the center of source to calibrate the nodes based on least squares (LS) algorithm [20]. Plinge and Fink used a white noise as calibration signal to calibrate the nodes in the circular array equipped five microphones based on TDOA and Direction of Arrival (DOA) [21]. Gaubitch et al. utilized the Gaussian modulated sinusoidal pulse signal and TOA information to estimate the locations of microphones based on LS algorithm [22]. Although the existing calibration methods can calibrate the nodes without assistant devices, but they have complex calculation and low calibration accuracy in noise and reverberation conditions.

* Corresponding author.
E-mail address: flyin@dlut.edu.cn (F. Yin).

https://doi.org/10.1016/j.apacoust.2018.05.022
Received 3 June 2017; Received in revised form 23 October 2017; Accepted 23 May 2018

In order to calibrate the microphone nodes accurately in DMA, a constrained total least squares-based calibration method for DMA is proposed in this paper. All the source event positions are first estimated by the known reference node based on the weighted multidimensional scaling (MDS) algorithm [23], then the perfect source events are picked up by the TDOA selection strategy at each ordinary node. Finally, the undetermined node microphones are calibrated by the total least squares (TLS) algorithm, and further refined based on the constrained total least squares (CTLS) method with a geometry constraint when the estimated results suffer from large errors. The proposed method can calibrate the node positions even when the acoustic source and the nodes are asynchronous. Furthermore, it can obtain high calibration accuracy in noise and reverberation conditions.

The rest of the paper is structured as follows. Section 2 presents the weighted MDS and TLS algorithms. Section 3 describes the proposed node calibration method based on the constrained TLS. In Section 4, the results of the simulation and real-world experiments are given and discussed. Finally, some conclusions are drawn in Section 5.

2. Weighted multidimensional scaling localization and total least squares algorithm

2.1. Weighted MDS localization algorithm [23]

MDS is a statistical technique often used in information visualization for exploring similarities or dissimilarities in data. The MDS algorithm treats each object as a point in a multidimensional space, and the distances between pairs of points are connected to similarities. The localization technique based on MDS, first proposed by Shang et al. [24], has offered a new idea for determining the positions of the object in network. Given the measured distances between acoustic source and nodes, the source positions can be estimated based on MDS algorithm. In order to achieve better calibration performance, the weighted MDS is utilized in this paper.

Let $\mathbf{S} = [x, y, z]^T$ be the source event position to be determined. The reference node with known position has M microphones, and the coordinate of the m-th microphone in reference node is $\mathbf{P}_{0,m} = [x_{0,m}, y_{0,m}, z_{0,m}]^T$, $m = 1, 2, \cdots, M$, $(M \geqslant 3)$. Thus the true distance d_m between the source event and the m-th microphone in reference node is

$$d_m = \sqrt{(\mathbf{P}_{0,m}-\mathbf{S})^T(\mathbf{P}_{0,m}-\mathbf{S})},\ \ m = 1, 2, \cdots, M \tag{1}$$

where T denotes the transpose operator. The range measurement can be modeled as

$$r_m = d_m + q_m,\ \ m = 1, 2, \cdots, M \tag{2}$$

where q_m is the range error at the m-th microphone in reference node. The distance between the m-th microphone and n-th microphone in reference node $d_{m,n}$ is

$$d_{m,n} = \sqrt{(\mathbf{P}_{0,m}-\mathbf{P}_{0,n})^T(\mathbf{P}_{0,m}-\mathbf{P}_{0,n})},\ \ m, n = 1, 2, \cdots, M \tag{3}$$

The scalar product matrix is defined as $\mathbf{B}$, and its (m, n) entry is given by [23]

$$B_{m,n} = 0.5(d_m^2 + d_n^2 - d_{m,n}^2) \tag{4}$$

where $m, n = 1, 2, \cdots, M$

Let $\mathbf{P} = [\mathbf{P}_{0,1}, \mathbf{P}_{0,2}, \cdots, \mathbf{P}_{0,M}]^T$ be the coordinate matrix of microphone in reference node, and $\mathbf{S}$ be the unknown source event position. Then the coordinate matrix $\mathbf{X} = [\mathbf{P}_{0,1}-\mathbf{S}, \mathbf{P}_{0,2}-\mathbf{S}, \cdots, \mathbf{P}_{0,M}-\mathbf{S}]^T$ can be expressed as

$$\mathbf{X} = \mathbf{P} - \mathbf{1}_M\mathbf{S}^T \tag{5}$$

where $\mathbf{1}_M$denotes the $M \times 1$ vector with all elements equal to unity. So the scalar product matrix, $\mathbf{B} = \mathbf{X}\mathbf{X}^T$, is a rank-3 symmetric and positive semi-definite matrix [25].

According to the orthogonality between the signal subspace and the noise subspace of the scalar product matrix, the source event position satisfies the following equation [23]

$$\mathbf{B}\begin{bmatrix}\mathbf{1}_M^T\\ \mathbf{P}^T\end{bmatrix}^{\dagger}\begin{bmatrix}1\\ \mathbf{S}\end{bmatrix} = \mathbf{0}_M \tag{6}$$

where $\mathbf{0}_M$ denotes the $M \times 1$ column vector with zero entities, and $(\cdot)^{\dagger}$ represents the Moore-Penrose inverse. For convenience, the pseudo-inverse matrix is expressed as

$$\mathbf{A} = \begin{bmatrix}\mathbf{1}_M^T\\ \mathbf{P}^T\end{bmatrix}^{\dagger} \tag{7}$$

Replacing the true range d_m by its noisy measurement value r_m in the scalar product matrix, the noisy matrix $\widehat{\mathbf{B}}$ is constructed from Eq. (4). The source event position can be estimated by the weighted MDS algorithm [26]

$$\widehat{\mathbf{S}} = -(\mathbf{A}_2^T\widehat{\mathbf{B}}^T\mathbf{W}\widehat{\mathbf{B}}\mathbf{A}_2)^{-1}\mathbf{A}_2^T\widehat{\mathbf{B}}^T\mathbf{W}\widehat{\mathbf{B}}\mathbf{A}_1 \tag{8}$$

where $\mathbf{A}_1$ is the first column of $\mathbf{A}$, and others $\mathbf{A}_2$, *i. e.* $\mathbf{A} = [\mathbf{A}_1|\mathbf{A}_2]$; $\mathbf{W}$ is the weighting matrix to whiten the residuals vector and can be obtained via regularization technique [26].

2.2. Total least squares algorithm

The linear parameter estimation problem can be formulated in matrix form

$$\mathbf{A}_0\theta = \mathbf{b} \tag{9}$$

When the matrix $\mathbf{A}_0$ is over-determined and all noises are confined to vector $\mathbf{b}$, the LS solution $\widehat{\theta}_{LS}$ of Eq. (9) is

$$\widehat{\theta}_{LS} = \underset{\widehat{\theta}}{\operatorname{argmin}}(\mathbf{A}_0\widehat{\theta}-\mathbf{b})^T(\mathbf{A}_0\widehat{\theta}-\mathbf{b}) = (\mathbf{A}_0^T\mathbf{A}_0)^{-1}\mathbf{A}_0^T\mathbf{b} \tag{10}$$

When the matrix $\mathbf{A}_0$ also suffers from noise, *i. e.*, $\mathbf{A} = \mathbf{A}_0 + \triangle\mathbf{A}$, $\mathbf{b} = \mathbf{b}_0 + \triangle\mathbf{b}$, where $\triangle\mathbf{A}$ and $\triangle\mathbf{b}$ are the noise components of $\mathbf{A}$ and $\mathbf{b}$, respectively. The $\widehat{\theta}_{LS}$ is generally biased, so total least squares (TLS) algorithm is needed to solve θ. Really, the idea of TLS is to look for the minimal correction terms to make $\mathbf{A}\theta = \mathbf{b}$ consistent when errors occur in both $\mathbf{A}$ and $\mathbf{b}$. Let $\mathbf{C} = [\mathbf{A}, \mathbf{b}]$, $\triangle\mathbf{C} = [\triangle\mathbf{A}, \triangle\mathbf{b}]$. The TLS solution of $\mathbf{A}\theta = \mathbf{b}$ can be formulated as the following optimization problem [27]

$$\begin{cases}\min\limits_{\theta,\Delta\mathbf{C}} \|\Delta\mathbf{C}\|_F\\ s.\ t.\ (\mathbf{A} + \Delta\mathbf{A})\theta = \mathbf{b} + \Delta\mathbf{b}\end{cases} \tag{11}$$

where $\|\cdot\|_F$ denotes the Frobenius norm [27]. The singular value decomposition of $\mathbf{C}$ is expressed as $\mathbf{C} = \mathbf{U}\Sigma\mathbf{V}^T$, where right singular matrix $\mathbf{V} = [v_1, v_2, \cdots, v_{L+1}]$, and singular matrix $\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_{L+1})$, here L is the column number of the matrix $\mathbf{A}$. If the smallest singular value σ_{L+1} is unique, then the TLS solution θ_{TLS} can be obtained as [27]

$$\widehat{\theta}_{TLS} = \frac{1}{v(L+1, L+1)}\begin{bmatrix}v(1, L+1)\\ \vdots\\ v(L, L+1)\end{bmatrix} \tag{12}$$

where $v(i, L+1)$ is the i-th element in the $(L+1)$-th column of $\mathbf{V}$.

If the smallest singular value σ_{L+1} is repeated k times, any v belonging to the subspace spanned by $[v_{L-k+2}, v_{L-k+3}, \cdots, v_{L+1}]$ can be used to determine the TLS solution, which implies that the solution is not unique. In this case, the minimum norm method [28] and the optimal least squares approximation method [29] can be used to obtain unique solution.