Signal Processing 107 (2015) 54–67

Contents lists available at ScienceDirect

Signal Processing

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

Localizing multiple audio sources in a wireless acoustic sensor network

Anthony Griffin [a,*], Anastasios Alexandridis [a,b], Despoina Pavlidi [a,b], Yiannis Mastorakis [a,b], Athanasios Mouchtaris [a,b]

[a] *Foundation for Research & Technology – Hellas, Institute of Computer Science (FORTH-ICS), Heraklion, Crete GR-70013, Greece*
[b] *University of Crete, Department of Computer Science, Heraklion, Crete GR-70013, Greece*

A R T I C L E I N F O

Article history:
Received 18 February 2014
Received in revised form
1 August 2014
Accepted 5 August 2014
Available online 23 August 2014

Keywords:
Acoustic sensors
Acoustic source localization
Location estimation
Microphone arrays
Wireless acoustic sensor networks

A B S T R A C T

In this work, we propose a grid-based method to estimate the location of multiple sources in a wireless acoustic sensor network, where each sensor node contains a microphone array and only transmits direction-of-arrival (DOA) estimates in each time interval, reducing the transmissions to the central processing node. We present new work on modeling the DOA estimation error in such a scenario. Through extensive, realistic simulations, we show that our method outperforms other state-of-the-art methods, in both accuracy and complexity. We also present localization results of real recordings in an outdoor cell of a sensor network.

1. Introduction

Microphone arrays have become increasingly popular due to their ability to perform direction-of-arrival (DOA) estimation. Identifying the direction of incoming sound is the basis for performing many operations, such as beamforming, speech enhancement, and distant sound acquisition. However, in many situations not only the DOA, but also the actual location of a sound source in space is required. Wireless acoustic sensor networks (WASNs), where a number of microphones or microphone arrays are distributed over an area, have emerged from the need to provide better spatial coverage and perform localization. WASNs have attracted a lot of interest due to their variety of application in hearing aids, ambient intelligence, hands-free telephony and acoustic monitoring [1].

Source localization in a WASN is a challenging task as the sensor network poses many constraints related to time-synchronization, power and bandwidth limitations, etc. For these reasons, approaches that require the transmission of the full audio signals to the central processing node are often unsuitable as they are bandwidth consuming, and the required transmission power can reduce the battery-life of the sensors. Moreover, such approaches require the signals to be synchronized. The work in [2] circumvented the problem of synchronization by using special nodes that used their internal Global Positioning System (GPS) chips to resample the audio samples with a network-common timestamp. However, the full audio signals still need to be transmitted to the central processing node.

By allowing increased computational ability in the nodes, the absolute minimum transmission bandwidth can be attained when each sensor node only transmits DOA estimates to the central processing node [3–5].

* Corresponding author.
E-mail addresses: doctorgriffin@yahoo.co.uk (A. Griffin), analexan@ics.forth.gr (A. Alexandridis), pavlidi@ics.forth.gr (D. Pavlidi), jmastor@csd.uoc.gr (Y. Mastorakis), mouchtar@ics.forth.gr (A. Mouchtaris).

http://dx.doi.org/10.1016/j.sigpro.2014.08.013
0165-1684/

Localization using bearing-only (i.e., DOA) estimates can also tolerate unsynchronized output given that the sources are static or that they move at a rather slow rate relative to the analysis frame.

The bearing-only localization problem for a single-source has been thoroughly investigated and a variety of estimators are available in the literature. Closed-form solutions include the Stansfield estimator [6], which is a weighted linear least squares estimator. The weights are determined from range information between the source and the sensors. When range information is not available, the Stansfield estimator reduces to the Orthogonal Vector (OV) estimator [7]—the unweighted version of the Stansfield estimator.

While simple in their implementation, these linear least squares algorithms suffer from increased estimation bias. For this reason, maximum-likelihood (ML) and non-linear least squares (NLS) algorithms have been investigated [8–12]. A comparison between the Stansfield estimator and the ML estimator in [8] reveals that the Stansfield estimator provides biased estimates even for a large number of measurements and that the bias may not vanish as the number of measurements increases. The work in [9] forms geometric relationships between the measured data and formulates the localization problem as a constrained optimization task, while [10] proposes a variant of the ML estimator that theoretically performs better than the traditional ML approach. Estimators that take into account the velocity of a moving source—especially for vehicle tracking—are discussed in [12,11].

The aforementioned methods consider the problem of localizing a single source. However, in many realistic scenarios multiple sources may co-exist in an area and the location of all sources may need to be known. The bearing-only multiple source localization problem of acoustic sources poses many challenging issues. First of all, the so-called *data association problem* occurs, where the central processing node receiving DOA estimates for multiple sources from the different sensors cannot know to which source they belong. Erroneous DOA combinations across the sensors will result in “ghost sources” that do not correspond to real sources. A solution to this problem was given in [13] and later generalized in [14] but has been found to be Non-deterministic Polynomial-time hard (NP-hard) when the number of sensors is ≥ 3. Another solution is discussed in [15,16] but is suitable only for noiseless scenarios. The work in [17] proposes a solution based on statistical clustering of the intersection of bearing lines. However, they again consider idealized scenarios of no missed detections and no spurious measurements. Localization of multiple sources by angle and frequency measurements is considered in [18], but this method will fail if the sources contain the same frequencies, and thus it cannot be applied to the case of acoustic sources. A method for multiple source localization using non-linear least squares that tries to surpass the data association problem is discussed in [19]. However, ghost sources are not eliminated, leading to severe performance degradation.

Our previous experience with DOA estimation [20,21] has revealed that when the sources are close together some arrays might only detect one source. This is a valid observation made from experiments using real recorded signals [20,21]. As a result, the DOAs of some sources from some sensors might be missing. This problem of missing DOA estimates as a function of the sources' locations is an important aspect which—to the best of our knowledge—has not been widely examined so far.

Our work in [22] considers a method for localizing two sources using far-field DOA measurements in an outdoor WASN. This paper extends [22] to more than two sources. Moreover, this paper proposes a novel iterative grid-based approach that can be thought of as an alternative solution to the NLS estimator. Other iterative solutions for source localization have also been proposed, the most popular of which are Steered Response Power (SRP) based approaches [23]. However, when applied to a WASN, such approaches require a significantly higher amount of information to be transmitted to the central processing node. In our approach only DOA estimates are transmitted to the central node, keeping bandwidth requirements to the minimum. When localizing a single source, our grid-based approach maintains the accuracy of the standard NLS, while performing much better in terms of computation time.

The computational efficiency allows our approach to be extended to localize multiple sources. To do so, we apply the single-source grid-based method to each possible combination of DOA measurements from the sensors and then solve the data association problem using a sub-optimal—yet efficient—method which relies on the estimated locations and the corresponding DOA combinations to decide on the actual source locations. Our approach is real-time and as our simulations and real experiments show, it remains accurate.

Our simulations use new results that we present here to model the DOA estimation error of the algorithm of [21] and consider the problem of missing DOAs as a function of source location, which makes them more realistic than simulations considered so far. The problem of missing DOAs when the sources are close together occurs very often in practice as our real experiments in this paper suggest.

The remainder of the paper is organized as follows. Section 2 sets up the basic definitions and assumptions for the problem. Section 3 reviews single source localization methods using DOA estimates and proposes the intersection point and the grid-based method. Then, Section 4 discusses the multiple source localization problem extending the intersection point and the grid-based methods for multiple sources. Simulation results and real experiments that compare the proposed methods with other state-of-the-art methods in realistic scenarios are presented in Section 5. Finally, Section 6 concludes the paper.

2. The framework

Our framework is a wireless acoustic sensor network whose M nodes are each equipped with a microphone array—which we will also refer to as a sensor. This enables each node to generate a direction-of-arrival estimate for any sources that it can “hear” (any sources whose signal-to-noise ratio (SNR) at the node is high enough to be detected). It is important to note that each node's estimates consist of direction only, and no range information,