IEEE TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

www.ieee.org/sp/index.html

MAY 2007 VOLUME 15 NUMBER 4 ITASD8 (ISSN 1558-7916)

PAPERS

(Contents Continued on Back Cover)

(Contents Continued from Front Cover)

CORRESPONDENCE

ANNOUNCEMENTS

IEEE SIGNAL PROCESSING SOCIETY

The Signal Processing Society is an organization, within the framework of the IEEE, of members with principal professional interest in the technology of transmission, recording, reproduction, processing, and measurement of speech and other signals by digital electronic, electrical, acoustic, mechanical, and optical means, the components and systems to accomplish these and related aims, and the environmental, psychological, and physiological factors concerned therewith. All members of the IEEE are eligible for membership in the Society and will receive this TRANSACTIONS upon payment of the annual Society membership fee of $27.00 plus an annual subscription fee of $34.00. For information on joining, write to the IEEE at the address below. *Member copies of Transactions/Journals are for personal use only.*

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published eight times a year in January, February, March, May, July, August, September, and November by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $90.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm.

Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. GST Registration No. 125634188. Printed in U.S.A.

Direction of Arrival Estimation Using the Parameterized Spatial Correlation Matrix

Jacek Dmochowski, Jacob Benesty, *Senior Member, IEEE*, and Sofiène Affes, *Senior Member, IEEE*

***Abstract*—The estimation of the direction-of-arrival (DOA) of one or more acoustic sources is an area that has generated much interest in recent years, with applications like automatic video camera steering and multiparty stereophonic teleconferencing entering the market. DOA estimation algorithms are hindered by the effects of background noise and reverberation. Methods based on the time-differences-of-arrival (TDOA) are commonly used to determine the azimuth angle of arrival of an acoustic source. TDOA-based methods compute each relative delay using only two microphones, even though additional microphones are usually available. This paper deals with DOA estimation based on spatial spectral estimation, and establishes the parameterized spatial correlation matrix as the framework for this class of DOA estimators. This matrix jointly takes into account all pairs of microphones, and is at the heart of several broadband spatial spectral estimators, including steered-response power (SRP) algorithms. This paper reviews and evaluates these broadband spatial spectral estimators, comparing their performance to TDOA-based locators. In addition, an eigenanalysis of the parameterized spatial correlation matrix is performed and reveals that such analysis allows one to estimate the channel attenuation from factors such as uncalibrated microphones. This estimate generalizes the broadband minimum variance spatial spectral estimator to more general signal models. A DOA estimator based on the multichannel cross correlation coefficient (MCCC) is also proposed. The performance of all proposed algorithms is included in the evaluation. It is shown that adding extra microphones helps combat the effects of background noise and reverberation. Furthermore, the link between accurate spatial spectral estimation and corresponding DOA estimation is investigated. The application of the minimum variance and MCCC methods to the spatial spectral estimation problem leads to better resolution than that of the commonly used fixed-weighted SRP spectrum. However, this increased spatial spectral resolution does not always translate to more accurate DOA estimation.**

***Index Terms*—Circular arrays, delay-and-sum beamforming (DSB), direction-of-arrival (DOA) estimation, linear spatial prediction, microphone arrays, multichannel cross correlation coefficient (MCCC), spatial correlation matrix, time delay estimation.**

I. Introduction

PROPAGATING signals contain much information about the sources that emit them. Indeed, the location of a signal source is of much interest in many applications, and there exists a large and increasing need to locate and track sound sources. For example, a signal-enhancing beamformer [1], [2] must continuously monitor the position of the desired signal source in order to provide the desired directivity and interference suppression. This paper is concerned with estimating the direction-of-arrival (DOA) of acoustic sources in the presence of significant levels of both noise and reverberation.

The two major classes of broadband DOA estimation techniques are those based on the time-differences-of-arrival (TDOA) and spatial spectral estimators. The latter terminology arises from the fact that spatial frequency corresponds to the wavenumber vector, whose direction is that of the propagating signal. Therefore, by looking for peaks in the spatial spectrum, one is determining the DOAs of the dominant signal sources.

The TDOA approach is based on the relationship between DOA and relative delays across the array. The problem of estimating these relative delays is termed "time delay estimation" [3]. The generalized cross-correlation (GCC) approach of [4], [5] is the most popular time delay estimation technique. Alternative methods of estimating the TDOA include phase regression [6] and linear prediction preprocessing [7]. The resulting relative delays are then mapped to the DOA by an appropriate inverse function that takes into account array geometry.

Even though multiple-microphone arrays are commonplace in time delay estimation algorithms, there has not emerged a clearly preferred way of combining the various measurements from multiple microphones. Notice that in the TDOA approach, the time delays are estimated using only *two* microphones at a time, even though one usually has several more sensor outputs at one's disposal. The averaging of measurements from independent pairs of microphones is not an optimal way of combining the measurements, as each computed time delay is derived from only two microphones, and thus often contains significant levels of corrupting noise and interference. It is thus well known that current TDOA-based DOA estimation algorithms are plagued by the effects of both noise and especially reverberation.

To that end, Griebel and Brandstein [8] map all "realizable" combinations of microphone-pair delays to the corresponding source locations, and maximize simultaneously the *sum* (across various microphone pairs) of cross-correlations across all possible locations. This approach is notable, as it jointly maximizes the results of the cross-correlations between the various microphone pairs.

The spatial spectral estimation problem is well defined in the narrowband signal community. There are three major methods: the steered conventional beamformer approach (also termed the "Bartlett" estimate), the minimum variance estimator (also termed the "Capon" or maximum-likelihood estimator), and

Manuscript received September 6, 2006; revised November 8, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hiroshi Sawada.

The authors are with the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, Montréal, QC H5A 1K6, Canada (e-mail: dmochow@emt.inrs.ca).

provides an excellent overview of these approaches. These three approaches are unified in their use of the narrowband spatial correlation matrix, as outlined in the next section.

The situation is more scattered in the broadband signal case. Various spectral estimators have been proposed, but there does not exist any common framework for organizing these approaches. The steered conventional beamformer approach applies to broadband signals. The delay-and-sum beamformer (DSB) is steered to all possible DOAs to determine the DOA which emits the most energy. An alternative formulation of this approach is termed the "steered-response power" (SRP) method, which exploits the fact that the DSB output power may be written as a sum of cross-correlations. The computational requirements of the SRP method are a hindrance to practical implementation [8]. A detailed treatment of steered-beamformer approaches to source localization is given in [10], and the statistical optimality of the approach is shown in [11]–[13]. Krolik and Swingler develop a broadband minimum variance estimator based on the steered conventional beamformer [14], which may be viewed as an adaptive weighted SRP algorithm. There have also been approaches that generalize narrowband localization algorithms (i.e., MUSIC [15]) to broadband signals through subband processing and subsequent combining (see [16], for example). A broadband linear spatial predictive approach to *time delay estimation* is outlined in [17] and [18]. This approach, which is limited to linear array geometries, makes use of all the channels in a joint fashion via the time delay parameterized spatial correlation matrix.

This paper attempts to unify broadband spatial spectral estimators into a single framework and compares their performance from a DOA estimation standpoint to TDOA-based algorithms. This unified framework is the azimuth parameterized spatial correlation matrix, which is at the heart of all broadband spatial spectral estimators.

In addition, several new ideas are presented. First, due to the parametrization, well-known narrowband array processing notions [19] are applied to the DOA estimation problem, generalizing these ideas to the broadband case. A DOA estimator based on the eigenanalysis of the parameterized spatial correlation matrix ensues. More importantly, it is shown that this eigenanalysis allows one to estimate the channel attenuation from factors such as uncalibrated microphones. The existing minimum variance approach to broadband spatial spectral estimation is reformulated in the context of a more general signal model which accounts for such attenuation factors. Furthermore, the ideas of [17] and [18] are extended to more general array geometries (i.e., circular) via the azimuth parameterized spatial correlation matrix, resulting in a minimum entropy DOA estimator.

Circular arrays (see [20]–[22], for example) offer some advantages over their linear counterparts. A circular array provides spatial discrimination over the entire 360° azimuth range, which is particularly important for applications that require front-to-back signal enhancement, such as teleconferencing. Furthermore, a circular array geometry allows for more compact designs. While the contents of this paper apply generally to planar

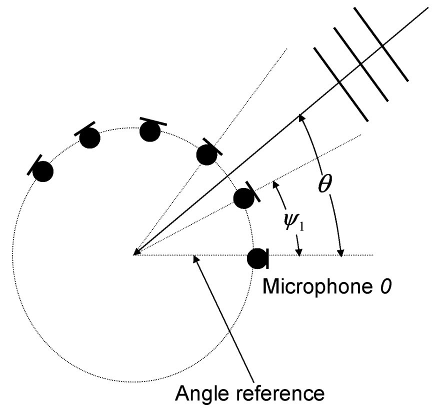

Fig. 1. Circular array geometry.

Section II presents the signal propagation model in planar (i.e., circular) arrays and serves as the foundation for the remainder of the paper. Section III reviews the role of the traditional, nonparameterized spatial correlation matrix in narrowband DOA estimation, and shows how the parameterized version of the spatial correlation matrix allows for generalization to broadband signals. Section IV describes the existing and proposed broadband spatial spectral estimators in terms of the parameterized spatial correlation matrix. Section V outlines the simulation model employed throughout this paper and evaluates the performance of all spatial spectral estimators and TDOA-based methods in both reverberation- and noise-limited environments. Concluding statements are given in Section VI.

The spatial spectral estimation approach to DOA estimation has limitations in certain reverberant environments. If an interfering signal or reflection arrives at the array with a higher energy than the direct-path signal, the DOA estimate will be false, even though the spatial spectral estimate is accurate. Such situations arise when the source is oriented towards a reflective barrier and away from the array. This problem is beyond the scope of this paper and is not addressed herein. Rather, the focus of this paper is on the evaluation of spatial spectral estimators in noisy and reverberant environments and on their application to DOA estimation.

II. Signal Model

Assume a planar array of $L+1$ elements in a 2-D geometry, shown in Fig. 1 (i.e., circular geometry), whose outputs are denoted by $x_l[n], l=0,1,\ldots,L$, where n is the time index. Denoting the azimuth angle of arrival by θ, propagation of the signal from a far-field source to microphone l is modeled as:

$$x_l[n] = \alpha_l s\left[n - t - f_l(\theta)\right] + v_l[n] \tag{1}$$

where α_l, $l=0,1,2,\ldots,L$, are the attenuation factors due to channel effects, t is the propagation time, in samples, from the

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | sales@docketalarm.com | 1-866-77-FASTCASE

fastcase® Smarter legal research.