ARTICLE IN PRESS

Journal of Sound and Vibration ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Blind separation of incoherent and spatially disjoint sound sources

Bin Dong [a,*], Jérôme Antoni [a], Antonio Pereira [a,b], Walter Kellermann [c]

[a] *Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France*
[b] *Univ Lyon, École Centrale de Lyon, Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, F-69134, Écully, France*
[c] *Institute of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstaβe 7, D-91058 Erlangen, Germany*

ARTICLE INFO

Article history:
Received 20 October 2015
Received in revised form
16 June 2016
Accepted 14 July 2016
Handling Editor: K. Shin

Keywords:
Blind source separation
Source identification
Inverse problem
Joint approximate diagonalization

ABSTRACT

Blind separation of sound sources aims at reconstructing the individual sources which contribute to the overall radiation of an acoustical field. The challenge is to reach this goal using distant measurements when all sources are operating concurrently. The working assumption is usually that the sources of interest are incoherent – i.e. statistically orthogonal – so that their separation can be approached by decorrelating a set of simultaneous measurements, which amounts to diagonalizing the cross-spectral matrix. Principal Component Analysis (PCA) is traditionally used to this end. This paper reports two new findings in this context. First, a sufficient condition is established under which "virtual" sources returned by PCA coincide with true sources; it stipulates that the sources of interest should be not only incoherent but also spatially orthogonal. A particular case of this instance is met by spatially disjoint sources – i.e. with non-overlapping support sets. Second, based on this finding, a criterion that enforces both statistical and spatial orthogonality is proposed to blindly separate incoherent sound sources which radiate from disjoint domains. This criterion can be easily incorporated into acoustic imaging algorithms such as beamforming or acoustical holography to identify sound sources of different origins. The proposed methodology is validated on laboratory experiments. In particular, the separation of aeroacoustic sources is demonstrated in a wind tunnel.

1. Introduction

A fundamental issue in noise and vibration engineering is to identify sound sources of different origins. With the development of more and more stringent standards in terms of acoustical quality, especially in the transportation industry, the need of dedicated techniques for localizing, quantifying, and ranking sound sources has become crucial [1]. In this

Abbreviations: 2SO (Joint), Statistical and Spatial Orthogonality; 3S, Supervised Source Separation; BSS, Blind Source Separation; CLT, Central Limit Theorem; CSM, Cross-Spectral Matrix; ESM, Equivalent Source Method; EVD, Eigen-Value Decomposition; HELS, Helmholtz's Equation Least-Squares; ICA, Independent Component Analysis; JAD, Joint Approximate Diagonalization; MCMC, Markov Chain Monte Carlo; NAH, Near-field Acoustical Holography; PCA, Principal Component Analysis; PSD, Power Spectral Density; SNR, Signal-to-Noise Ratio; SO, Statistical Orthogonality (Only); SOI, Source Of Interest; SONAH, Statistically Optimized NAH; STFT, Short Time Fourier Transform; SVD, Singular Value Decomposition

* Corresponding author. Fax: +33 472 43 87 12.
E-mail addresses: bin.dong@insa-lyon.fr, sagittarius1013@gmail.com (B. Dong), jerome.antoni@insa-lyon.fr (J. Antoni), antonio.pereira@ec-lyon.fr (A. Pereira), wk@LNT.de (W. Kellermann).

http://dx.doi.org/10.1016/j.jsv.2016.07.018
0022-460X/

respect, a recurrent challenge is to separate the partial contributions of the different sound sources that contribute to the overall radiated noise, in particular when sources are all active at the same time and have overlapping frequency spectra.

Many methods have been proposed in the past to meet the above requirements and to replace the traditional subsystem masking techniques which are time consuming and prone to influence the operation conditions of the object under study. Among the most popular approaches, acoustic imaging techniques such as beamforming and Near-field Acoustical Holography (NAH) are particularly interesting because they allow contactless measurements recorded by an array of microphones and are rather universal in their principle. Introduction by Maynard, Williams and Lee in the 80's [2,3], NAH has the remarkable capability to indirectly reconstruct sound sources (typically parietal pressure and normal component of particle velocity) with a good spatial resolution and reasonable quantification. Several other sound imaging techniques have been proposed to meet different industrial needs, such as Statistically Optimized NAH (SONAH) [4,5], the Helmholtz's Equation Least-Squares (HELS) [6], the Equivalent Source Method (ESM) [7,8], Bayesian Focalization [9], to name just a few. Reviews of some of these methods can be found for instance in Refs. [10] and [11]. They will be referred to herein as “backpropagation” methods as they all aim at reconstructing the sound source distributions by backpropagating the measured acoustical pressure to the source domain. This implies solving an inverse problem.

Backpropagation methods are good enough to localize and identify the Sources Of Interest (SOIs) when their spacing is large enough as compared to the attainable spatial resolution (e.g. Rayleigh's limit) and when their relative levels are within the available dynamic range. In other situations, when the SOIs are physically very close to each other, slightly overlap in space, and/or exhibit significant differences in level, their visual separation by traditional acoustic imaging techniques may be difficult or even impossible. As a consequence, the reconstructed source distribution still contains a superposition of mixed components that remain to be unraveled.

One way to solve the problem is to exploit the property that sound sources of distinct physical origins can reasonably be assumed mutually independent and then to resort to statistical criteria to achieve their separation. This is the realm of “source separation”, whose objectives are of prime interest in practice.

Technically speaking, there are essentially two types of source separation methods found in the literature: Supervised Source Separation (3S) and Blind Source Separation (BSS). 3S methods can separate out any SOI for which an external reference is available. A “reference” is a signal measured simultaneously with the radiated acoustical field and which is perfectly coherent with the SOI (e.g. a vibration signal captured close to the SOI). This implies that it is uncorrelated – i.e. statistically orthogonal – with the other sources in the mixture. Thus, a mean-square-error prediction filter (also called Wiener filter) can be constructed which maps the reference signal to the sound measurements. By definition, the output of the prediction filter is an estimate of the SOI. Other – but theoretically equivalent – implementations are based on the use of partial coherences [12]. Due to its simplicity, the method began to attract attention in the late 70's right after the dual channel analyzers came out. References [13–16] report early applications to acoustic imaging (mainly NAH). The method has been extended later to account for various scenarii such as weakly nonstationary sources [17] and cyclostationary sources [18]. However, 3S methods have fundamental limitations: 1) references must be available, 2) they must be of excellent quality (in a sense to be described shortly), and 3) they must be of sufficient number (at least as many references as SOIs). Requirement (1) is not always fulfilled, in particular due to accessibility constraints or to limited numbers of tracks of the data acquisition system. Requirement (2) is probably the most difficult to attain: it implies the measurement of external signals with theoretically infinite Signal-to-Noise Ratios (SNR), which are fully coherent with the SOI and totally uncorrelated with the other sources. Positions where these conditions are met may not exist at all and, even if they do, their localization would ideally require solving the source identification problem first. Requirement (3) is also a strong one in particular when several sources are to be separated.

In order to alleviate some of these limits, Tomlinson made use of the Principle Component Analysis (PCA) in an attempt to correct a set of non-ideal references [19]. Another elegant solution has been proposed in Ref. [20] which is to replace external references by “numerical” ones returned by a first resolution of the inverse problem. Yet this is likely to succeed only in the case of sources which are initially well separated in space [21–23].

Indeed, early efforts have been spent to avoid the need of any reference at all. This brings us to the second group of source separation methods. Historically, first contributions to the subject are probably due to Price et al. [24] and Otte et al. [25] in the late 80's who proposed to decorrelate a set of measurements in order to force them to comply with the property of references. The resulting uncorrelated signals are then interpreted as “virtual sources”. Technically speaking, this amounts to the diagonalization of the Cross-Spectral Matrix (CSM), which may be achieved either by PCA, partial coherences (an implementation of Gram-Schmidt orthogonalization), Cholesky factorization, etc. As pointed out by Price et al., there is no reason that the so separated virtual sources are totally consistent with a physical origin. Reference [26] demonstrates that PCA separation actually holds provided that the SOI is dominating; similarly, virtual sources obtained by partial coherences are meaningful only if the iterative orthogonalization is performed in a pyramidal order where the n-th measurement (used at iteration n) contains no more contributions than n SOIs including the $n-1$ previously extracted ones (at iterations $1,\ldots,n-1$). Given an arbitrary numbering of the SOIs and of the measurements, this means that the first selected measurement must contain only SOI#1, the second selected measurement – to be orthogonalized with the first one – must contain only SOIs #1 and #2, the third selected measurement – to be orthogonalized with the subspace spanned by the first and the second ones – must contain only SOIs #1, #2 and #3, etc. Needless to say that such a pyramidal order is hardly encountered in practice. A similar limitation was pointed in Ref. [27] in 1976. Although it was early recognized that virtual sources do not match the SOIs in general, the technical literature contains numerous instances of the application of PCA in sound source