Sparse Acoustic Source Localization with Blind Calibration for Unknown Medium Characteristics

Cagdas Bilen, Srđan Kitić, Nancy Bertin, Rémi Gribonval

Sparse Acoustic Source Localization with Blind Calibration for Unknown Medium Characteristics

Çağdaş Bilen[1], Srđan Kitić[1], Nancy Bertin[2] and Rémi Gribonval[1]

[1]INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes Cedex, France.*

[2]IRISA - CNRS UMR 6074, Campus de Beaulieu, F-35042 Rennes Cedex, Rennes, France.

Abstract— **We consider the problem of audio source localization in a medium with unknown characteristics, particularly the speed of sound within the medium. We propose an algorithm to retrieve both the speed of sound and the localized signals under reasonable assumptions such as smoothness of medium structure. We present initial simulation results for stationary sinusoidal sources both for the case of uniform and non-uniform speed of sound through the medium to demonstrate the performance of the proposed algorithm.**

1 Introduction

Sparse recovery has become a popular research field especially after the emergence of compressed sensing theory [1]. Consequently many natural signals have also been shown to be compressible or essentially low dimensional which enabled sparsity inducing methods to be used for the estimation of these signals in linear inverse problems.

Acoustic source localization can also be formulated as a sparse recovery problem [2, 3]. In the continuous space and time, the acoustic pressure field in a 2-dimensional space, $p(\overrightarrow{s}, t)$, induced by the pressure emitted by the sources, $f(\overrightarrow{s}, t)$, is known to obey the differential wave equation

$$\Delta p(\overrightarrow{s}, t) - \frac{1}{c^2}\frac{\partial^2 p(\overrightarrow{s}, t)}{\partial t^2} = \begin{cases} 0, & \text{if no source at } \overrightarrow{s} \\ f(\overrightarrow{s}, t), & \text{if source at } \overrightarrow{s} \end{cases} \tag{1}$$

at all non-boundary locations. The parameter c in (1) represents the speed of sound in the medium. For inhomogeneous mediums, this constant can be represented more generally as $c(\overrightarrow{s}, t)$. In addition to the wave equation, the acoustic pressure field is also restricted by the boundary conditions. Both the wave equation and the boundary equations can be discretized straightforwardly into a set of linear differential equations represented as

$$(\mathbf{\Omega}_1 + \alpha\mathbf{\Omega}_2)\mathbf{x} = \mathbf{z} \tag{2}$$

in which $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ represent the discretized differential operators in space and time respectively, $\alpha \triangleq 1/c^2$ is a function of speed of sound, $\mathbf{x}$ and $\mathbf{z}$ represent the discretized pressure field and the source signals in vector form respectively.

The source localization problem can be described as the estimation of $\mathbf{z}$ (or equivalently $\mathbf{x}$) from limited number of measurements recorded by a number of microphones, i.e. from

$$\mathbf{y} = \mathbf{I}_s\mathbf{x} \tag{3}$$

where $\mathbf{I}_s$ is a matrix formed by keeping some of the rows of the identity matrix corresponding to microphone locations. Considering the sparse nature of the sources, source localization can be posed as a sparse recovery problem which can be solved with convex optimization such as in

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|(\mathbf{\Omega}_1 + \alpha\mathbf{\Omega}_2)\mathbf{x}\|_* \text{ s.t. } \mathbf{I}_s\mathbf{x} = \mathbf{y} \tag{4}$$

where the $*$ norm is often chosen as the ℓ_1 norm or the $\ell_{1,2}$ norm [3]. In many such approaches, the medium is assumed to be homogeneous and the medium properties (c) are assumed to be known. However, in practice this assumption may not be always valid since the behavior of the medium may vary depending on the physical conditions such as temperature, ambient pressure, or unknown inhomogeneities in the medium. These incorrect assumptions on α (and hence the sparsifying operator $\mathbf{\Omega}_1 + \alpha\mathbf{\Omega}_2$) may lead to significantly decreased sparsity or multiplicative perturbations that would highly reduce the recovery performance [4, 5].

In this abstract, we investigate the source localization problem formulated in (4) but with an unknown speed of sound c (or equivalently α). In Section 2, we formulate the source localization problem for a spatially varying c, i.e. a time invariant but inhomogeneous medium and present an algorithm to estimate the medium properties along with the sources under spatial smoothness assumptions. Preliminary experimental results with the proposed algorithm are then presented in Section 3, which is followed by the concluding remarks in Section 4.

2 Source Localization with Unknown Medium Parameters

Let us first assume that the medium is inhomogeneous and unknown, i.e. $\alpha = \alpha(\overrightarrow{s}, t)$ for the continuous domain. Consequently the discretized equation in (2) can be modified as

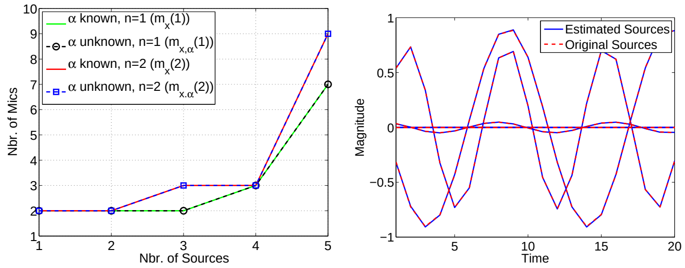
$$(\mathbf{\Omega}_1 + \mathrm{Diag}(\boldsymbol{\alpha})\mathbf{\Omega}_2)\mathbf{x} = \mathbf{z} \tag{5}$$

where $\mathrm{Diag}(\boldsymbol{\alpha})$ indicates a diagonal matrix with the entries of the vector $\boldsymbol{\alpha}$ along the diagonal. In this setting one can attempt to recover $\boldsymbol{\alpha}$ and $\mathbf{x}$ by minimizing

$$\hat{\mathbf{x}}, \hat{\boldsymbol{\alpha}} = \arg\min_{\mathbf{x},\boldsymbol{\alpha}} \|(\mathbf{\Omega_1} + \mathrm{Diag}(\boldsymbol{\alpha})\mathbf{\Omega_2})\mathbf{x}\|_{1,2} \text{ s.t. } \mathbf{I}_s\mathbf{x} = \mathbf{y} \tag{6}$$

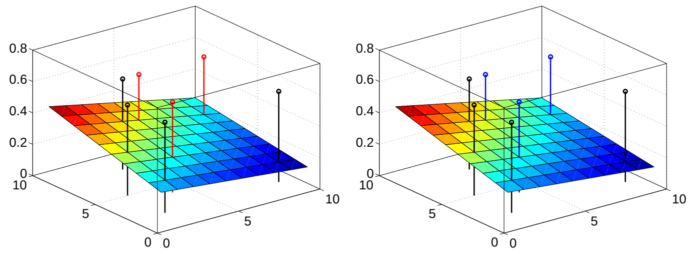
where the $\ell_{1,2}$ norm is defined as ℓ_1 norm along the spatial and ℓ_2 norm along the temporal dimension, assuming all the source locations are fixed along time. However the global minimum of the objective function in (5) often does not correspond to the actual $\mathbf{x}$ and $\boldsymbol{\alpha}$ due to the high number of degrees of freedom introduced by the unknown $\boldsymbol{\alpha}$.

*This work was partly funded by the European Research Council, PLEASE project (ERC-StG-2011-277906).



(a) (b)

Figure 1: (a) The minimum sufficient number of the microphones, $m_{\mathbf{x}.\boldsymbol{\alpha}}(n)$, with $n = 1, 2$ for successfully localizing different number of sources. The required number of microphones for the ideal case when the field $\boldsymbol{\alpha}$ is known, $m_{\mathbf{x}}(n)$, is also plotted for comparison. (b) The original and estimated signals emitted from the sources for a sample recovery.



(a) The original field $\boldsymbol{\alpha}$ (b) The estimated field $\hat{\boldsymbol{\alpha}}$

Figure 2: The original field $\boldsymbol{\alpha}$ and the estimated field $\hat{\boldsymbol{\alpha}}$ on a 10 by 10 2-dimensional grid for a sample experiment. The red, blue and black markers indicate the spatial positions of the original sources, estimated sources and the microphones respectively.

The intrinsic degrees of freedom in the structure of a real world medium is often much smaller. In this work we will assume the medium does not change in time and exploit the fact that the spatial variation in a natural medium is often limited and assume nth order spatial difference of the field $\boldsymbol{\alpha}$ along each spatial dimension is zero, i.e.

$$\mathbf{D}_n \boldsymbol{\alpha} = 0 \tag{7}$$

$$\Rightarrow \boldsymbol{\alpha} = \mathbf{A}_n \mathbf{p} \tag{8}$$

where $\mathbf{A}_n$ is an orthogonal basis for the null space of $\mathbf{D}_n$. It can be observed that the degrees of freedom for $\boldsymbol{\alpha}$ are reduced to n^2 (for 2-dimensional space) from the total size of the discretized points in space, which can be a significant reduction in practice. With the assumption of (8), the optimization in (6) can be modified as

$$\hat{\mathbf{x}}, \hat{\mathbf{p}} = \arg\min_{\mathbf{x},\mathbf{p}} \|(\boldsymbol{\Omega}_1 + \mathrm{Diag}(\mathbf{A}_n \mathbf{p})\boldsymbol{\Omega}_2)\mathbf{x}\|_{1,2} \text{ s.t. } \mathbf{I}_s \mathbf{x} = \mathbf{y} \tag{9}$$

Unlike the case with known α, (9) is not a linear optimization when jointly considering the variables $\mathbf{x}$ and α, hence the well developed convex optimization methods are not applicable. It is, however, linear in terms of each of the unknowns when the other is considered a constant. Therefore we propose a form of the alternating minimization approach which is often employed in such scenarios.

3 Experimental Results

In order to demonstrate the joint estimation performance of the proposed approach, a rectangular grid of 10 by 10 spatial dimensions is simulated with k randomly located stationary sources, k varying from 1 to 5. The sources emit random sinusoidals for a duration of 20 time instants. The pressure field is measured by m randomly located stationary microphones for 100 time instants, m varying from 2 to 10. The boundary conditions are modeled to be Neumann boundaries as in [2, 3]. The field $\boldsymbol{\alpha}$ is generated randomly to obey (8) with parameter n set as 1 and 2 and the field values to be between 0.1 and 0.6.

The field parameters $\mathbf{p}$ and the acoustic pressure $\mathbf{x}$ are estimated from the microphone measurements $\mathbf{y} = \mathbf{I}_s \mathbf{x}$ with an algorithm that minimizes (9) by alternating between unknowns applying the method of Alternating Direction Method of Multipliers (ADMM [6]). The details of the algorithm are not provided here due to space constraints.

Each of the randomly generated experiments is repeated 20 times with the same parameters to empirically estimate the minimum sufficient number of microphones for successful source localization for each set of n, m, k. After each simulation, k points with highest energy are chosen as estimates of the source locations. The minimum number of microphones that succeeds in localization of the sources in 80% of the 20 experiments is set as the minimum sufficient number of microphones, $m_{\mathbf{x},\boldsymbol{\alpha}}(n)$. The minimum sufficient number of the microphones for the recovery of the acoustic field when the medium parameter, $\boldsymbol{\alpha}$, is known, $m_{\mathbf{x}}$, is also computed the same way for comparison. Figure 1a shows the change of $m_{\mathbf{x}.\boldsymbol{\alpha}}(n)$ and $m_{\mathbf{x}}(n)$ with respect to the number of sources. A sample reconstruction of source signals is shown in Figure 1b. The original and the estimated field $\boldsymbol{\alpha}$ as well as the source and the microphone locations for the same sample reconstruction can also be seen in Figure 2.

The first thing to notice in the presented results in Figure 1a is that estimating the unknown α introduces almost no performance degradation with respect to perfectly knowing the medium parameters. This can be attributed to the very small degrees of freedom in $\boldsymbol{\alpha}$ thanks to the smoothness constraints. Hence estimating $\boldsymbol{\alpha}$ while localizing sources is possible whenever the sources can be localized with known $\boldsymbol{\alpha}$. In fact it is observed in the experiments that the estimation of $\boldsymbol{\alpha}$ is still accurate for many cases when source localization is not successful. A second observation is that the sufficient number of microphones is not significantly affected when $\boldsymbol{\alpha}$ is spatially varying ($n = 2$) compared to when it is spatially constant ($n = 1$). However the degrees of freedom within $\boldsymbol{\alpha}$ increases with n^2 and therefore the sufficient number of microphones is expected to increase more rapidly with increasing n.

4 Conclusion

In this work, the acoustic source localization problem with additional variables due to unknown medium properties is described and a sparse recovery approach is presented for estimating both the sources and the medium structure. The presented approach utilizes the assumption that the spatial variation in the medium is limited and the acoustic sources are sparse. Preliminary results that demonstrate the performance of the proposed approach are presented. The talk will include further results, the details on the recovery algorithm and discussions on the extensions of the algorithm for more realistic scenarios.

References

[1] D. Donoho, “Compressed sensing”, Information Theory, IEEE Transactions on, vol. 52, no. 4, 2006.

[2] S. Nam and R. Gribonval, “Physics-driven structured cosparse modeling for source localization”, in Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on.

[3] S. Kitić, N. Bertin and R. Gribonval, “A review of cosparse signal recovery methods applied to sound source localization”, Le XXIVe colloque Gretsi, Brest, France, Sep 2013.

[4] E. J. Candès, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements”, Communications on Pure and Applied Mathematics, vol. 59, no. 8, pp. 1207-1223, 2006.

[5] M. Herman and T. Strohmer, “General deviants: An analysis of perturbations in compressed sensing”, Selected Topics in Signal Processing, IEEE Journal of, vol. 4, pp. 342-349, April 2010.

[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers”, Foundations and Trends in Machine Learning, vol. 3, no. 1, pp. 1-122, 2011.