SOURCE LOCALIZATION IN REVERBERANT ENVIRONMENTS USING SPARSE OPTIMIZATION

Jonathan Le Roux, Petros T. Boufounos, Kang Kang, John R. Hershey*

Mitsubishi Electric Research Laboratories (MERL)
201 Broadway, Cambridge, MA 02139, USA

ABSTRACT

In this paper, we demonstrate that recently-developed sparse recovery algorithms can be used to improve source localization in reverberant environments. By formulating the localization problem in the frequency domain, we are able to efficiently incorporate information that exploits the reverberation instead of considering it a nuisance to be eliminated. In this formulation, localization becomes a joint-sparsity support recovery problem which can be solved using model-based methods. We also develop a location model which further improves performance. Using our approach, we are able to recover more sources that the number of sensors. In contrast to conventional wisdom, we demonstrate that reverberation is beneficial in source localization, as long as it known and properly accounted for.

Index Terms— compressive sensing, joint sparsity, microphone array, reverberation, source localization

1. INTRODUCTION

Recent work on compressive sensing and sparse representations has provided new powerful theory and tools for signal acquisition and processing applications. Efficient sparse and model-based sparse reconstruction is now possible using a number of new algorithms and signal models, which are able to capture a large variety of the signals we are interested in processing.

In this paper we examine how these new algorithms can be applied to improve source localization in broadband reverberant environments. Specifically, we propose a model-based sparsity-enforcing method which enables us to capture the reverberant characteristics of the environment and to localize more sources than the number of sensors. Our approach operates in the frequency domain, taking into account the frequency response of the environment and, accordingly, formulating a separate narrowband localization problem at each frequency. To ensure that all problems share information and provide a consistent localization output, we enforce a joint-sparsity constraint across the problems.

This approach extends and improves earlier work [1], which used simple joint-sparsity models in anechoic environments. In addition, we further enhance the model using a source location model which takes into account the fundamental localization ambiguity due to the array configuration. We also demonstrate that, in contrast to conventional wisdom, reverberation can be beneficial in source localization, as long as it is known and properly accounted for. Note that the assumption that the reverberation is known is also investigated in previous work on the reconstruction of more sources than sensors [2].

*Part of this work was performed while K. Kang was an intern at MERL in Summer 2012. He is now with Texas Instruments, Dallas, TX.

Our work builds on recent theoretical and algorithmic results on compressed sensing (CS) and sparse representations [3,4], including model-based CS [5]. We also capitalize on earlier sparsity-based work on sensors arrays [6,7], focused either on narrowband arrays or on broadband arrays observing signals with a time-sparse structure.

There is a large body of work on localization of audio sources, e.g., [8–10]. Previous works generally do not assume knowledge on reverberation conditions but attempt to be robust to reverberation effects, which are considered a distortion to be accounted for. Thus, their performance decreases as reverberation time and intensity increases. Instead, we demonstrate that, when the reverberation is known, it can be beneficial if properly exploited, improving localization performance.

2. BACKGROUND IN JOINT-SPARSITY MODELS

Sparsity has recently emerged as a powerful tool in signal processing applications. A sparse model assumes that a signal, when transformed to an appropriate basis, has very few significant coefficients which can explain most of its energy. Promoting sparsity in some appropriate domain has proved to be a very efficient computational method to capture the structure of most natural and man-made signals processed by modern signal processing systems.

Sparsity is typically useful when inverting a linear system—usually underdetermined—of the form

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1}$$

where $\mathbf{y}$ is the M-dimensional measurement vector and $\mathbf{x}$ is the N dimensional sparse signal vector. For an underdetermined system, the following optimization determines the sparsest solution:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{y} \approx \mathbf{A}\mathbf{x}, \tag{2}$$

where the ℓ_0 norm counts the number of non-zero coefficients of $\mathbf{x}$. This is a combinatorially complex problem, but under certain conditions on $\mathbf{A}$, solution is possible in polynomial time using a convex relaxation of the ℓ_0 norm [3, 4] or one of the many greedy algorithm developed over the recent years. These include the orthogonal matching pursuit (OMP), the Compressive Sampling Matching Pursuit (CoSaMP) [11], the Subspace Pursuit (SP) [12], and the iterative hard thresholding (IHT) [13].

Joint- and group-sparsity models, and their variations, provide further structure to the signal of interest [5, 14–25]. Joint sparsity can be considered as a special case of group sparsity, thus we only describe the latter. Under this model, the signal coefficients are partitioned into groups $\mathcal{G}_i$, which partition the coefficient index set $\{1, \ldots, N\}$. The group-sparsity model assumes that only a few of these groups contain non-zero coefficients, and most groups contain

all-zero coefficients. Group sparsity can also be enforced using a convex optimization problem or one of many greedy algorithms.

In joint-sparsity models, several sparse signals are measured together. The model assumption is that all the signals share the same sparsity pattern. In other words, the significant signal coefficients are located at the same positions for all signals. By considering the whole acquisition as a linear system, it is straightforward to show that these models are a special case of group-sparsity models and a similar approach can determine the sparse output.

Model-based compressed sensing [5] further enables models with more complex constraints and structure than typical sparsity or group-sparsity problems. As shown in [5], it is straightforward to modify standard algorithm such as CoSaMP [11] to enforce model-based sparsity. All that is necessary is a model-based thresholding function, which replaces the standard thresholding function in standard algorithms, and truncates the signal according to the model. Indeed, in this paper we modify CoSaMP with a model-based truncation which enforces both a joint-sparsity model and a signal location model we develop.

3. MODEL

We consider M sensors, each receiving a signal $y_m(t)$, $m = 1, \ldots, M$, where t denotes a time index. We discretize the area of interest to N gridpoints, indexed using $n = 1, \ldots, N$. Each of those points is a potential source location, emitting a signal $x_n(t)$. This signal is exactly equal to zero if there is no source in that location and non-zero if a source emits from that location.

In a reverberant environment, the signal $x_n(t)$ from each source n arrives to each sensor m through multiple paths. Assuming the sources and sensors are omnidirectional, the contribution $\tilde{x}_{mn}(t)$ of source n to the signal received at sensor m, which we shall refer to as the source image of n at m, can be represented as the convolution of the original signal at source n with a filter $a_{mn}(t)$ referred to as the room impulse response (RIR) for that pair of source and sensor:

$$\tilde{x}_{mn}(t) = (a_{mn} * x_n)(t). \tag{3}$$

Directionality of the sensors can be incorporated in $a_{mn}(t)$, while that of the sources can be handled by extending the grid.

If the window of analysis is sufficiently long, typically longer than the length of the RIR, the relationship between the source signal, the source image and the RIR can be reasonably well approximated in the frequency domain using the narrowband approximation:

$$\tilde{X}_{mn}(\omega) \approx A_{mn}(\omega) X_n(\omega), \tag{4}$$

where ω denotes the discretized angular frequency, and $\tilde{X}_{mn}$, A_{mn} and X_n denote the discrete Fourier transform (DFT) of the windowed $\tilde{x}_{mn}$, a_{mn} and x_n, respectively.

Summing over the contributions from all potential sources, we model the signal received at sensor m as:

$$Y_m(\omega) \approx \sum_n A_{mn}(\omega) X_n(\omega), \tag{5}$$

which can be rewritten, for each ω, as a matrix multiplication

$$\mathbf{Y}(\omega) \approx \mathbf{A}(\omega)\mathbf{X}(\omega), \tag{6}$$

where the data vector is $\mathbf{Y}(\omega) = [Y_1(\omega), \ldots, Y_M(\omega)]^T$ and the source vector is $\mathbf{X}(\omega) = [X_1(\omega), \ldots, X_N(\omega)]^T$. This model generalizes that in [1] to reverberant environments. We further assume that sources are broadband and that at most K of them exist in the scene at any time, where K is much smaller than N. The number of sources K might or might not be known in advance. Broadband sources emit signals in a wide bandwidth, i.e., $X_n(\omega)$ is non-zero for a wide range of possible ω's.

The localization problem can be formulated as that of estimating the support of $\mathbf{X}(\omega)$ given the mixing matrix $\mathbf{A}(\omega)$ and the observed signals $\mathbf{Y}(\omega)$. As we assume that there are only a small number of active sources compared to the total number of grid locations, the vector $\mathbf{X}(\omega)$ is sparse, and recovery of the support of $\mathbf{X}(\omega)$ can be considered as a classical compressed sensing problem. Furthermore, this support, or sparsity pattern, is common to all frequencies ω, or, in other words, the vectors $\mathbf{X}(\omega)$ are jointly sparse.

Under the broadband assumption, it is possible to identify the location of more than M sources using M sensors, although not necessarily reconstruct all M of them. The ambiguities that arise in the single frequency case manifest themselves for different locations at different frequencies. The joint-sparsity model exploits these differences to resolve these ambiguities and localize the sources.

4. ALGORITHMS

4.1. Location model

Even when considering the jointly sparse model, the coherence of columns of the matrix $\mathbf{A}(\omega)$ corresponding to neighboring locations can be high, impairing the performance of the system. If two sources are close to each other, localization and recovery become an ill-conditioned problem. Furthermore, if the discretization of the localization grid is very fine when compared to the coherence pattern of the array, nearby locations will look very similar to the location of an actual source, often confusing the recovery algorithm and hindering its ability to localize weaker sources.

Fortunately, it is unlikely in practice for two sources to be sufficiently close to each other as measured by the coherence pattern of the array. By actively enforcing a location model preventing sources from being too close to each other, the localization performance increases dramatically. This can be performed at each iteration of any greedy algorithm whenever a truncation is performed, simply by performing a model-based truncation [5].

The model-based truncation we enforce ensures that no two identified support locations are close to each other. Given a signal $\mathbf{X}$, our truncation, denoted $T(\mathbf{X}, K)$, simply returns K support locations, greedily starting from the grid point with maximum total energy, each time discarding all points in the neighborhood of the selected point before making another selection.

4.2. Joint-sparsity and model-based CoSaMP

As described above, the signal $\mathbf{X}(\omega)$ only contains at most K non-zero components (i.e., it is K-sparse) which have the same support in all frequencies. Under this formulation, localization becomes a joint-support recovery problem [6] and signal recovery becomes a joint-sparsity recovery problem.

Recent work in group- and joint-sparsity problems has demonstrated that the principles from simple-sparsity problems can be transferred to joint-sparsity ones if the norms in convex formulations and the truncations in greedy algorithms are not computed on each signal vector independently but on vectors comprising the energy of each of the groups in the support. Specifically, in our case,

Algorithm 1 The modified joint-sparsity CoSaMP algorithm used in the experiments in this paper.

initialize
Iteration count $l = 0$,
Initial signal estimate $\widehat{\mathbf{X}}^l(\omega) = (0)$, for all ω
Initial support estimate set $\mathcal{S}^l = \emptyset$.
while not converged **do**
  Increase iteration count: $l \leftarrow l + 1$
  Compute residual: $\mathbf{R}(\omega) = \mathbf{Y}(\omega) - \mathbf{A}(\omega)\widehat{\mathbf{X}}^{l-1}(\omega)$, for all ω
  Compute gradient: $\mathbf{G}^l(\omega) = \mathbf{A}^H(\omega)\mathbf{R}(\omega)$
  Compute proxy: $\mathbf{p} = E(\mathbf{G}^l)$ {as in (7)}
  Identify support candidate: $\mathcal{T}^l = T(\mathbf{p}, 2K) \cup \mathcal{S}^{l-1}$
  Invert over support: $\mathbf{B}^l(\omega) = \mathbf{A}^\dagger_{\mathcal{T}^l}(\omega)\mathbf{Y}(\omega)$, for all ω
  Compute final support: $\mathcal{S}^l = T(E(\mathbf{B}^l), K)$
  Truncate and update estimate: $\widehat{\mathbf{X}}^l(\omega) = \mathbf{B}^l(\omega)\big|_{\mathcal{S}^l}$
end while
return Signal estimate $\widehat{\mathbf{X}}^l(\omega)$, support estimate $\mathcal{S}^l$.

we consider a vector energy function

$$(E(\mathbf{X}))_n = \sqrt{\sum_\omega |X_n(\omega)|^2}. \quad (7)$$

Minimizing or constraining the ℓ_0 or the ℓ_1 norm of $E(\mathbf{X})$—also known as the mixed ℓ_0/ℓ_2 or ℓ_1/ℓ_2 norm of $\mathbf{X}$ with respect to grouping defined by ω—subject to $\mathbf{Y}(\omega) \approx \mathbf{A}(\omega)\mathbf{X}(\omega)$ for all ω generates solutions with common support structure, i.e., jointly sparse with respect to the support of each $\mathbf{X}(\omega)$.

To compute the joint-sparsity pattern in the experiments in the remainder of this paper, we use a variation of model-based CoSaMP [5, 11] described in Algorithm 1. This algorithm uses a model-based truncation which operates on the energy signal as described above. This enables to separate the impact of two different models: the joint-sparsity model and the location model.

We found experimentally that recovery performance is improved if appropriate normalizations of the columns of the matrix $\mathbf{A}(\omega)$ are used in the residual, inverse and gradient steps. Specifically, if $\mathbf{A}_n(\omega)$ denotes the n^{th} column of $\mathbf{A}(\omega)$, we obtained best performance when using, for the residual and inverse steps, a normalization over frequencies and sensors for each grid position: $\tilde{\mathbf{A}}_n(\omega) = \mathbf{A}_n(\omega)/\sqrt{\sum_{\omega',m} |\mathbf{A}_{mn}(\omega')|^2}$, and, for the gradient step, a normalization over sensors only: $\mathbf{A}'_n(\omega) = \mathbf{A}_n(\omega)/\sqrt{\sum_m |\mathbf{A}_{mn}(\omega)|^2}$.

The algorithm returns both a location estimate and a signal estimate. As mentioned earlier, under the broadband assumption, it is possible to identify the location of more than M sources using M sensors, but not necessarily to reconstruct all of them. For localization, the joint-sparsity model is able to resolve the ambiguities that arise at each frequency by exploiting their differences. However, full system inversion to recover the source signal and not just its support is not always possible unless all individual $\mathbf{A}(\omega)$ are invertible over the recovered support set or there is prior information or a model on the sources.

5. EXPERIMENTS

5.1. Setup

We conducted a series of simulations to evaluate the localization performance of our method. The task was to recover the location of multiple sources in a room given the mixed signals observed at three microphones. The number of sources was assumed known, and ranging from 1 to 14. We considered a rectangular room of dimensions 3.0 m × 2.0 m, as shown in Fig. 1, where the roof and floor were assumed anechoic. Two microphones were placed 15 cm to the left wall and 5 cm on each side of the median line, and a third one 1 m to the right wall on the median line. We considered as potential positions for the sources a 5 cm grid from 0.8 m to 1.5 m of the left wall, and from the bottom to the top wall. Source positions were randomly selected while ensuring that they are no less than 20 cm apart from each other. Localization of a source was assumed correct when the estimated location was the exact same point on the grid as the true source location, and incorrect otherwise. Figure 1 shows an example of configuration for 5 sources together with the neighorhoods corresponding to the location model. Source signals were assumed to be white Gaussian noise. The simulation results we report here are averaged on 100 experiments for each number of sources.

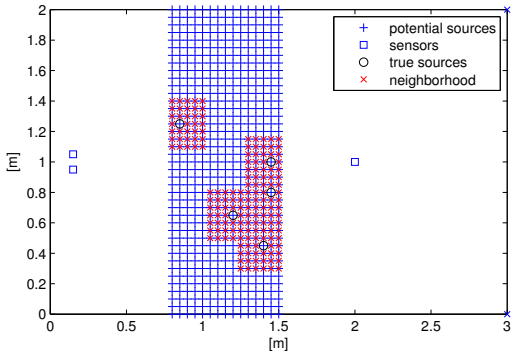

Fig. 1. Room setup.

Table 1. *RT60 (ms) for each absorption coefficient c.*

c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
RT60	412	208	139	104	79	56	51	41	29	9

Room impulse responses (RIR) for all pairs of source positions and microphones were simulated using the Roomsimove toolbox[1], with a 16 kHz sampling rate and assuming omnidirectional sources and sensors. Note however that directionality of both sources and sensors could be included in our model, for the sources for example by augmenting the grid to include discrete directions, and for the sensors simply by including directionality in the RIRs. Absorption coefficients for all walls were assumed equal, and varied from $c = 0.1$ to $c = 1.0$. The corresponding RT60 for a left-wall microphone and a source on the median line at 0.8 m of the left wall are shown in Table 1. Source images (contribution of a source) at each microphone were computed by convolving source signals with the corresponding RIR in the time domain.

We considered two types of signals: in the one denoted "raw data", invalid parts of the convolution are discarded, which amounts to assuming that we observe part of the convolution of an infinite source signal with the RIR; in the one denoted "known tail data" (KT), we assume that the source signal is of finite length, and that the source image at the microphone includes both the onset and the tail of the signal due to reverberation. This is arguably an ideal situation that could approximately occur in a forward-backward framework where the contributions of the signal from both the previous and next frame can be estimated. The raw data is used by default unless otherwise specified. We assumed that various lengths T of

[1] http://www.irisa.fr/metiss/members/evincent/Roomsimove.zip

signals could be observed at the microphone: 512, 1024, 2048 and 4096 samples. The KT data was assumed to be obtained by convolving a signal of length T with the RIRs and truncated at $2T$. For consistence, the raw data was assumed to be of length T and the DFT was computed after zero-padding to $2T$. We did not however notice a significant impact of zero-padding on performance.

5.2. Results

5.2.1. *Influence of the room reverberation time*

We first investigate the influence of the room RT60 on localization performance. We expect a decrease in the absorption coefficient c of the walls, i.e., an increase in reverberation time, to decrease the coherence of the mixing matrices $\mathbf{A}(\omega)$, and thus to result in an increase in localization performance. Indeed, as shown in Fig. 2 for $T = 4096$, lower c, or higher RT60, leads to dramatically better performance for all numbers of sources than lower RT60, in particular the anechoic case $c = 1$. Similar results were obtained for other values of T.

5.2.2. *Introduction of a location model*

Next, we compare the performance of our method with and without using a location model (LM) preventing estimated sources to be strictly closer than 20 cm to each other. As can be seen in Fig. 3 ($c = 0.3$), introducing a location model drastically improves performance, by up to almost 30 % absolute in the case of 14 sources for $T = 4096$. We can also see on this figure how performance improves with the window length.

5.2.3. *Enforcing length limitation*

We finally compare the localization performance in the realistic case of the raw data and in the ideal case of the KT data. In the latter, the original source data being of finite length, we investigated the possibility of enforcing that the recovered time-domain sources were also of finite length, in another flavor of model-based truncation. The algorithm only needs to be modified by introducing an extra step in the computation of the source estimate, which consists in taking the inverse DFT, truncating to finite length T by setting $\hat{x}(t) = 0$ for $t \in [T+1, 2T]$ and taking the DFT. Performance on raw data and KT data using the original algorithm and on KT data using the modified algorithm with finite length truncation (FLT) is shown in Fig. 4 for $c = 0.3$. The improvements obtained with KT data and using FLT are mainly seen for short signal lengths. For $T = 4096$, trying to exploit some knowledge gained on the source signals in neighboring frames may lead to some improvement, but FLT is likely not to help.

6. CONCLUSION

As we demonstrate, sparse recovery algorithms can be exploited to localize sources in reverberant environments. A joint-sparsity formulation in the frequency domain can properly account for the reverberation. Thus, many more sources than sensors can be accurately located. In contrast to conventional wisdom, we demonstrate that reverberation, if it is accurately known and properly exploited, helps instead of hurting. Our work only scratches the surface and raises a number of interesting questions. For example, we often do not know the RIR; future work will investigate its estimation and adaptation to room conditions. Furthermore, we only localize the sources. Reconstruction is not straightforward; it is also deferred to future publications.

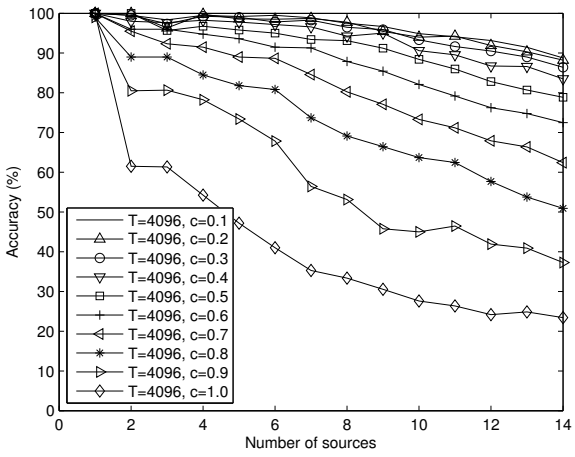

Fig. 2. Influence of room reverberation time on performance.

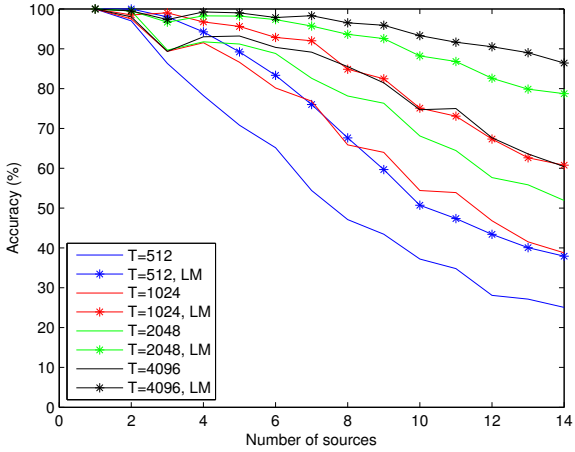

Fig. 3. Performance comparison with/without location model (LM).

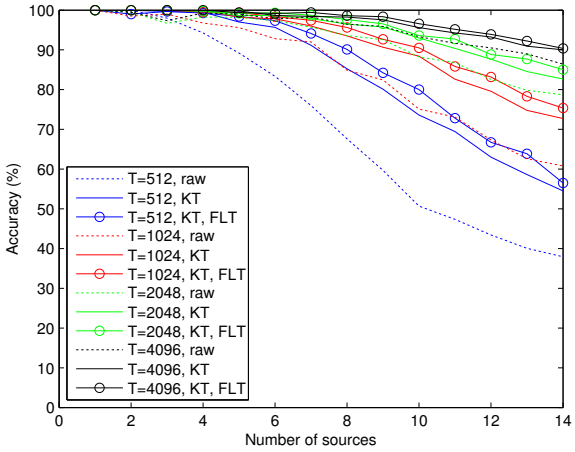

Fig. 4. Performance comparison of the original algorithm on raw data and known tail data (KT) and of the modified algorithm enforcing finite length (FLT) of the time-domain signal for the KT data.

7. REFERENCES

[1] P. T. Boufounos, P. Smaragdis, and B. Raj, "Joint sparsity models for wideband array processing," in *Proc. SPIE*, Aug. 2011.

[2] M. Kowalski, E. Vincent, and R. Gribonval, "Beyond the narrowband approximation: Wideband convex methods for underdetermined reverberant audio source separation," *IEEE Trans. ASLP*, 2009.

[3] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Info. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[6] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[7] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624–638, 2006.

[8] J. DiBiase, H. Silverman, and M. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays – Signal Processing Techniques and Applications*, M. Brandstein and D. Ward, Eds., chapter 8, pp. 157–180. Springer, 2001.

[9] B. Loesch, P. Ebrahim, and B. Yang, "Comparison of different algorithms for acoustic source localization," *ITG-Fachbericht-Sprachkommunikation 2010*, 2010.

[10] C. Blandin, A. Ozerov, and E. Vincent, "Multi-source TDOA estimation in reverberant audio using angular spectra and clustering," *Signal Processing*, 2011.

[11] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[12] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Info. Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[13] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[14] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[15] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[16] L. Peotta and P. Vandergheynst, "Matching pursuit with block incoherent dictionaries," *IEEE Trans. Signal Processing*, vol. 55, no. 9, pp. 4549 –4557, Sep. 2007.

[17] M. Fornasier and H. Rauhut, "Recovery algorithms for vector valued data with joint sparsity constraints," *SIAM J. Numer. Anal.*, vol. 46, no. 2, pp. 577–613, 2008.

[18] M. Fornasier and H. Rauhut, "Iterative thresholding algorithms," *Applied and Computational Harmonic Analysis*, vol. 25, no. 2, pp. 187–208, 2008.

[19] Y. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. Info. Theory*, vol. 56, no. 1, pp. 505–519, 2010.

[20] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 3075 –3085, Aug. 2009.

[21] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Info. Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[22] M. Kowalski and B. Torrésani, "Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients," *Signal, Image and Video Processing*, vol. 3, no. 3, pp. 251–264, Sep. 2009.

[23] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3042 –3054, 2010.

[24] P. Boufounos, G. Kutyniok, and H. Rauhut, "Sparse recovery from combined fusion frame measurements," *IEEE Trans. Info. Theory*, vol. 57, no. 6, pp. 3864–3876, 2011.

[25] S. Foucart, "Recovering jointly sparse vectors via hard thresholding pursuit," in *Proc. Sampling Theory and Applications (SampTA)*, Singapore, May 2011.