# EFFICIENT INTERFERER CANCELATION BASED ON GEOMETRICAL INFORMATION OF THE REVERBERANT ENVIRONMENT

*D.Riva, P.Pagani, F.Antonacci, G.Prandi, A.Sarti, M.Tagliasacchi, S.Tubaro*

Politecnico di Milano, Dipartimento di Elettronica ed Informazione, Piazza Leonardo da Vinci, 32, 20133 Milano
driva/antonacc/prandi/sarti/tagliasa/tubaro@elet.polimi.it

## ABSTRACT

*This paper concerns the problem of separation of wideband acoustic sources in reverberant environments. The idea is to perform separation in two stages: first we assess the position of desired and interferer sources and then we use this knowledge to steer the sensitivity of a microphone array towards the desired source. Notice, however, that the steering of a null towards its DOA is not sufficient to perform cancelation of the interferer in reverberant environments, due to reflections. For this reason, nulls of the beampattern are steered not only towards the interferer location but also towards its most important echoes. The DOAs of the echoes are predicted using a fast beam tracing algorithms. Experimental results confirm that an improvement Signal to Interference Ratio of 5 dB over Blind Source Separation techniques can be reached in a moderately reverberant environment with moving sources.*

## 1. INTRODUCTION AND PREVIOUS WORK

This paper concerns the problem of wideband acoustic separation using microphone arrays in reverberant environments. Applications in which microphone arrays are involved (i.e. separation of acoustic sources, localization, etc.), aim at being *robust against reverberations*. In this paper we propose to use *reverberations as an additional source of information*.

In recent years separation of acoustic sources has been an increasing field of study for researchers. Several industrial applications can be devised for the field of acoustic separation ranging from teleconferencing to audio-surveillance. A categorization of separation techniques which is useful for our discussion distinguishes between signal based and geometry based approaches. The first class relies only on some prior knowledge about the signal to perform separation, while the second class exploits the knowledge about the spatial distribution of acoustic energy to steer the sensitivity of the microphone array.

The approach presented in this work is a hybrid between these two classes, as it obtains the estimation of sources locations through the TRINICON algorithm (signal-based) and uses this information to suitably design a beamformer through Generalized Sidelobe Canceling (geometry-based).

Blind Source Separation techniques belong to the signal-based class. A further categorization distinguishes between methods based on Second Order Statistics (BSS-SOS) [1] and Higher Order Statistics (BSS-HOS) [2]. BSS-SOS methods lack of effectiveness when dealing with non-gaussian signals, like speech. An example of BSS-HOS methods is TRINICON: it exploits non-gaussianity, non-stationarity and non-whiteness of speech signals to iteratively estimate the de-mixing filters in reverberant environments. A summary about statistical principles of Blind Source Separation can be found in [3].

In [4] the authors point out that the extremes of the de-mixing filters convey information about the Time Differences of Arrival, which in turn can be used to localize sources in space. During the adaptive estimation process, the first de-mixing filters taps in the TRINICON algorithm that reach the convergence are those related to the extremes, as detailed in [5]. As a result, we can state that the localization task can be accomplished before the separation task. When sources are free to move, the channel impulse responses between each source and each microphone are time varying; as a consequence the de-mixing filters are time varying as well. Therefore, when the channel variation is faster than the convergence rate of the iterative process, the algorithm is unable to separate signals. At the same time, however, the location of the extremes of the de-mixing filters vary accordingly to the source movement. As a consequence, when sources are moving, we are able to localize but not to separate sources.

Geometry based approaches use prior information about the distribution of energy of the signal to steer the sensitivity of the microphone array towards the desired source. It is possible to do a general categorization of beamforming techniques used in literature in the following terms:

- depending on the bandwidth of the signals, a beamformer may be *narrow-band* or *broadband*;
- a beamformer may be *data-independent* or *data-driven*, according to the fact that data are used or not in the assessment of the spatial response;
- a beamformer may be designed to work with *nearfield* or *farfield* acoustic waves.

In this work we are interested in farfield and broadband beamformers. Different criteria may be adopted to design the weights of the beamformer. Direct Synthesis of the Beampattern [6] aims at synthesizing the beampattern through the minimization of a cost function based on the difference between the desired beampattern and the actual beampattern in some relevant points inside the mainlobe and the sidelobe regions. Linearly Constrained Minimum Variance Beamformers (LCMV) obtain a minimum variance spatial response while preserving the gain of the beamformer in specific Directions of Arrival of interest. One of the problems related to the use of LCMV beamformers is the dependency of the beamformer on both the data and the constraints [7], which makes the optimization a difficult task. Generalized Sidelobe Cancelers provide an elegant solution to this problem by dividing the filter into two components, constraints-dependent and constraints-independent [8]. The constraint independent component is estimated in an iterative fashion.

In this paper we propose a hybrid approach that merges Blind Source Separation and beamforming: the separation process is performed in two steps. First, the location of the acoustic sources is obtained through the de-mixing filters, as envisioned in [9]. With this information we are able to determine the direction of arrival of the desired and interferer sources. When dealing with reverberant environments a significative fraction of the total acoustic energy derives from reflections. For this reason the nulls of the beampattern are directed not only towards the direction of arrival of the interferer but also towards its most important echoes. Given the prior information about the environment information and the estimated position of the sources, the DOAs of the most significative echoes are obtained through a fast beam tracing algorithm [10]. The beamformer is finally designed in an iterative fashion: only the constraints which mainly contribute to a significative reduction of the interferer energy are retained. In order to predict which constraints are the most significative, an iterative algorithm is proposed.

The rest of this paper is organized as follows: in Section 2 the proposed approach is described. Section 3 presents some experimental results which confirm the validity of the proposed approach. Finally Section 4 proposes future developments and makes some final conclusions.

## 2. THE PROPOSED SOLUTION

The scheme of the proposed solution is depicted in Figure 1. We can distinguish three main blocks: localization, computation of the Directions of Arrival and beamforming. The algorithms used to accomplish localization, DOA computation and beamforming tasks are dependent upon the experimental conditions. The estimation of the de-mixing filters enables us to localize sources in space. Given the environment geometry and the estimations of source positions, we can properly design the weights $\{w_i\}_{i=1}^{N}$ of the beamformer. The following paragraphs detail each of the blocks composing our system. The techniques used to localize sources and compute their DOAs have been the subject of previous publications, therefore just a brief summary is provided here. On the contrary, although beamforming is accomplished with Generalized Sidelobe Cancelling, the iterative construction of the beamformer presents some new aspects and details are given in subsection 2.3.

Two different microphone arrays perform localization and separation. In particular the localization array is composed by four pairs of microphones disposed in the proximity of the walls of the environment. The separation array is composed by a 35 microphones linear array.

### 2.1 Localization

Consider the case of two sources. In order to estimate the de-mixing filters we adopt the TRINICON algorithm on microphone pairs.

Time Differences of Arrival are related to the de-mixing filters by equations (1) and (2).

$$\hat{\tau}_1 = \frac{\arg\max_n |w_{12}(n)| - \arg\max_n |w_{22}(n)|}{f_s} , \qquad (1)$$

$$\hat{\tau}_2 = \frac{\arg\max_n |w_{11}(n)| - \arg\max_n |w_{21}(n)|}{f_s} , \qquad (2)$$

where $n$ is the tap index of the de-mixing filter. When sources are observed from multiple pairs of microphones, source positions are estimated through triangulation. However, the collection of TDOAs exhibit the presence of outliers due to noise, therefore a tracking algorithm is needed to remove outliers. Due to the strong nonlinearity that relates TDOAs and source positions, a linear dynamical model does not fit the problem. As a consequence, the tracking algorithm used in this work is based on particle filtering. The computational cost required by conventional particle filtering, however, is generally too high. For this reason in [11] and [12] we propose to adopt Regularized Particle Filter and Swarm Intelligence particle filter. The results obtained confirm that in a moderately reverberant environment, the root mean square localization error is smaller than $0.3m$. We must observe that the algorithm proposed in this paper is sensitive to localization errors: beamforming may impose nulls of the beampattern towards undesired directions. We will experimentally verify the sensitivity of the separation capabilities of the proposed algorithm to the localization error.

### 2.2 Computation of the Directions of Arrival

A fundamental assumption behind the computation of DOAs is that the acoustic sources are sufficiently distant from the microphones such that the acoustic waves impinging on the array are planar. The planar wavefront assumption simplifies the computation of DOAs: in fact we can expect that a source is observed from the array under a constant angle from all the microphones. Therefore the DOAs can be computed at a reference microphone, denoted by index $M$. In some situations, however, the farfield wavefront assumption could bring significative errors.

At this stage we assume that the environment is two dimensional. This assumption significatively reduces the computational cost of the beam tracing algorithm. As a final result, the path tracing returns a set $\mathscr{L}$ of paths from each source to the reference microphone. A path $P_i$ in the set $\mathscr{L}$ has the following form:

$$P_i = \{S_l, R_{1i}, R_{2i}, \ldots, R_{Ki}, M\} , \qquad (3)$$

where $S_l$ is the $l$-th source, $R_{ji}$ , $j = 1, \ldots, K$ , $K \geq 0$ denote the bounce point on the walls and $M$ is the reference microphone. In order to compute the DOAs of the source echoes, we are interested in determining the angle formed by the acoustic ray in the last reflection and the axis of the array, i.e. the bounce which conveys the acoustic energy to the reference microphone. Let $\angle\overline{R_{Ki}M}$ denote the angle formed by the segment $\overline{R_{Ki}M}$ and $\alpha$ the angle of the axis of the microphone array. Both angles $\alpha$ and $\angle\overline{R_{Ki}M}$ are computed with respect to a reference direction. The DOA $D(\overline{R_{Ki}M})$ of the path $P_i$ is computed as

$$D(\overline{R_{Ki}M}) = \angle\overline{R_{Ki}M} - \alpha .$$

In order to assess which are the most important reflections and make a sorted list of them, we must be able to compute the amplitude of each reflection. The amplitude $A_i$ of the acoustic path $P_i$, whose length is $l_i$ is computed as

$$A_i = \frac{r^K}{l_i} , \qquad (4)$$

where $r$ is the reflection coefficient (assumed to be equal for all the walls) and $K$ is the number of bounces involved in the path $i$. As a result, the sorted map $S_{l,i=1:N}$ which makes a correspondence between the angle $D(\overline{R_{Ki}M})$ and its amplitude is generated.

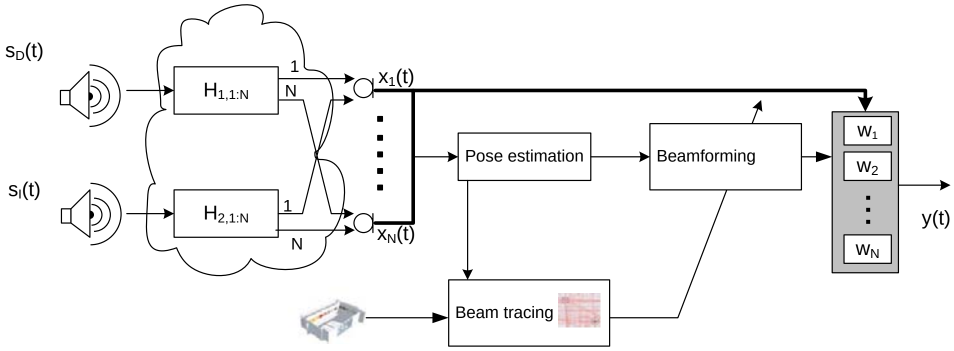


Figure 1: The proposed solution: we can observe three main components: localization, computation of DOAs and beamforming

## 2.3 Beamforming

Different algorithms have been conceived to steer the sensitivity of the array towards the desired source and reject the interferer. In this work we adopt the Generalized Sidelobe Canceler. In order to extend the GSC in a wideband setting, the computation of the weights is performed on 30 significative frequencies. Intermediate frequencies are then computed with a linear interpolation of the amplitude and cubic interpolation of the phase. A comprehensive summary on Generalized Sidelobe Canceler is in [7].

The linear separation array is composed by 35 elements. Their displacements are shown in Figure 2. Central microphones ($d = 5cm$) are used for high frequencies ($1400 - 3000Hz$), while microphones displaced by $10cm$ and $15cm$ are used, respectively, for mid-band ($1000 - 1400Hz$) and low-band frequencies ($300 - 1000Hz$).

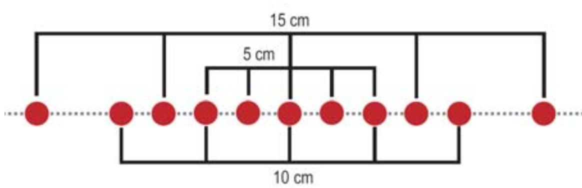


Figure 2: Array composition and microphones displacement

If we try to cancel all the interferer echoes at once, the separation capabilities may result worse than expected. Two reasons explain this behavior:

- when an echo is cancelled, an overlap between the DOAs of the echoes of interferer and desired sources may occur. This configuration causes the attenuation not only of the interferer but also of the desired source;
- the addition of a new constraint may cause a degradation of the shape of the beampattern.

A justification of the above issues is in [7].

Let $N$ denote the number of echoes in the sorted list for the desired and interferer sources. In order to prevent an undesired degradation of the separation capabilities, we propose a null selection algorithm. This algorithm builds the beamformer iteratively, by adding a single constraint at time and verifying each time that the separation capabilities of the beampattern are improved with respect to the previous step.

With the information provided by $S_{l,i=1:N}$ we are able to compute the function $C_{\mathrm{des}}(\theta)$ that inform us about the angular distribution of the echoes from the target source. At the same time the function $C_{i,\mathrm{int}}(\theta)$ inform us about the angular distribution of the energy from the interferer up to the $i$-th echo in the $S_{l,i=1:N}$ sorted list. The functions $C_{\mathrm{des}}(\theta)$ and $C_{i,\mathrm{int}}(\theta)$ are then used to estimate which echoes of the interferer are most likely disturbing in the separation process. The process is repeated until the beampattern exhibits a significative degradation when a new constraint is added.

Details about the null selection process are found in the following paragraphs.

1: compute the function $C_{\mathrm{des}}(\theta)$, which provides information about the spatial distribution of the energy of the desired source

$$C_{\mathrm{des}}(\theta) = \sum_{l=0}^{l=N} \delta(\theta - D(\overline{R_{Kl,\mathrm{des}}M})) \,, \tag{5}$$

where $D(\overline{R_{Kl,\mathrm{des}}M})$ is the $l$-th DOA referred to the desired source.

2: **for** $i = 1$ to $N$ **do**

3: compute the function $C_{i,\mathrm{int}}(\theta)$, which yields information about the echoes of the interferer in the sorted list $S_{i=1:N,\mathrm{int}}$ up to the $i$-th reflection:

$$C_{i,\mathrm{int}}(\theta) = \sum_{l=0}^{l=i} \delta(\theta - D(\overline{R_{Kl,\mathrm{int}}M})) \,, \tag{6}$$

where $D(\overline{R_{Kl,\mathrm{int}}M})$ is the $l$-th DOA of the interferer in the sorted list.

4: Steering of the array sensitivity towards a specific direction implies a leakage of energy from adjacent directions, as the extension of the array and the number of microphones are finite. Functions $C_{\mathrm{des}}(\theta)$ and $C_{\mathrm{int}}(\theta)$ are smoothed to obtain, respectively, $C^{lp}_{\mathrm{des}}(\theta)$ and $C^{lp}_{i,\mathrm{int}}(\theta)$. In particular the leakage depends on the position of the constraint (lobes close to $-\pi/2$ and $\pi/2$ are wider). Therefore the smoothing filter is a function of the angle. The dependency of the cutoff frequency of the filter on angle and frequency has been determined heuristically.

5: The decorrelation function $R_i(\theta)$ is computed:

$$R_i(\theta) = \frac{C^{lp}_{i,\mathrm{int}}(\theta)}{C^{lp}_{\mathrm{des}}(\theta)} .$$

6: For each DOA of the interferer to be canceled, the value of the function $R_i(\theta)$ is checked: if the value is below a threshold this DOA is discarded, since the cancelation of the interferer results also in the cancelation of the desired source.

7: Constraints are imposed: the beampattern must exhibit unitary and zero gain towards, respectively, the desired source and the echoes of the interferer survived to the null selection process.

8: The root mean square error between the actual beamformer and the desired one is computed. The error is computed on equally spaced angles. Only when the difference at step $i$ is lower than the difference at step $i-1$ the new beamformer is accepted.

9: **end for**

Preliminary experimental results confirm that the above algorithm prevents the cancelation of nulls that may result detrimental for the separation capabilities.

## 3. APPLICATION SCENARIO AND EXPERIMENTAL RESULTS

In order to validate the proposed algorithm, we have conducted two simulations. The metric used to evaluate the effectiveness of the separation algorithm is the Signal to Interference Ratio (SIR). The SIR compares the power of the desired and interferer sources after separation. Original unmixed signals must be available to estimate SIR, therefore a correct computation of SIR can be done only in a simulation context. In particular let $\Delta SIR$ denote the SIR difference between the output (after separation) and the input (before separation) stages.

A comparison of $\Delta SIR$ of the proposed approach and TRINICON is given. We must observe, however, that the algorithm proposed in [2] makes use of two sensors and, although possible, to our knowledge no implementation has been given that accounts for more sensors. On the other hand, the number of microphones used in our approach is scalable. We propose here to use 35 microphones. As a consequence, we can infer that part of the improvement of the separation capabilities of our algorithm with respect to TRINICON takes from the different number of microphones used.

The first test aims at finding the sensitivity of the proposed approach against localization error. Reverberations were simulated using the beam tracer in [10] and the environment is $30m \times 20m \times 3m$. The reference microphone is located in coordinates $(0,0)$. The farfield assumption in this experiment does not hold, since the source is not sufficiently apart from the array. The actual position of the interfering source is $(2m, 2.39m)$. In order to test the sensitivity, the interferer has been uncorrectly localized. The wrong locations are represented by the symbols $\varepsilon_1$ to $\varepsilon_8$. Table 1 shows the wrong locations together with the DOA offset (calculated from the reference microphone) with respect to the actual source position.

Five echoes of the interferer were canceled. The corresponding SIR is shown in Figure 3. We can see that the SIR decreases in an almost linear fashion as the localization error increases. As a reference performance index, the TRINICON algorithm in the same situation achieves $18dB$ as $\Delta SIR$.

Table 1: Wrong locations, localization and DOA errors

| Position | Coord. (x,y) | Loc. Error | Angle offset |
|---|---|---|---|
| $\varepsilon_1$ | 1.88m,2.49m | 0.15m | 2.95° |
| $\varepsilon_2$ | 1.77m,2.58m | 0.30m | 5.52° |
| $\varepsilon_3$ | 1.66m,2.68m | 0.45m | 8.22° |
| $\varepsilon_4$ | 1.54m,2.78m | 0.60m | 10.02° |
| $\varepsilon_5$ | 1.43m,2.87m | 0.75m | 13.51° |
| $\varepsilon_6$ | 1.31m,2.97m | 0.90m | 16.2° |
| $\varepsilon_7$ | 1.2m,3.06m | 1.05m | 18.59° |
| $\varepsilon_8$ | 1.08m,3.16m | 1.20m | 21.13° |

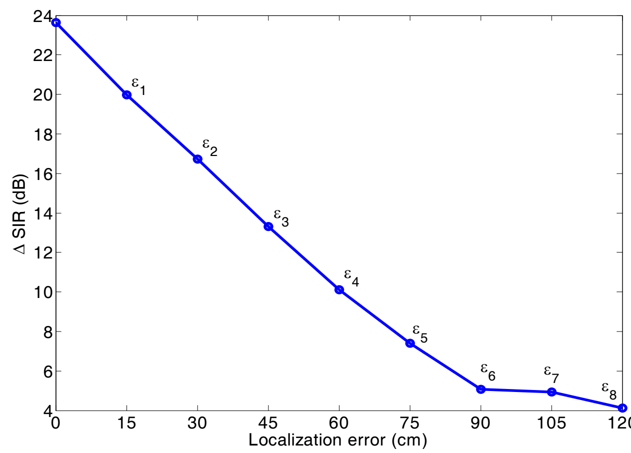


Figure 3: $\Delta SIR$ vs. the localization error of the interferer

The second test aims at comparing the SIR obtained by TRINICON and the proposed algorithm in a moving source setting. A room contains two sources free to move and active at the same time. The environment has dimensions $5m \times 5m \times 2.7m$ and the reverberation time is $0.25s$. The dimensions of the environment make the farfield hypothesis not valid. Trajectories of the sources have been generated using the Langevin dynamical model. In order to smooth random variations among different trajectories, the average $\Delta SIR$ of seven realizations vs. number of echoes canceled is plotted in Figure 4. The $\Delta SIR$ of TRINICON is plotted as a reference performance index. We can observe from Figure 4 that the proposed approach achieves a Signal to Interference Ratio of 15 dB which turns out to be effective for many applications. In the same situation, the $\Delta SIR$ obtained by TRINICON is about $4dB$: since sources were both in movement, the TRINICON algorithm was unable to identify the correct de-mixing filters. When sources were static, the TRINICON $\Delta SIR$ was only $6dB$ lower than the $\Delta SIR$ of the proposed approach. Due to the moving sources setting, TRINICON is unable to correctly identify the de-mixing filters, as a consequence the difference between $\Delta SIR$ increases to $11dB$. This fact proves the efficiency of the proposed approach over other BSS techniques. It is also worthwhile to observe that TRINICON works with only two microphones, while the proposed approach uses the array depicted in Figure 2.

Although the null selection process should prevent a

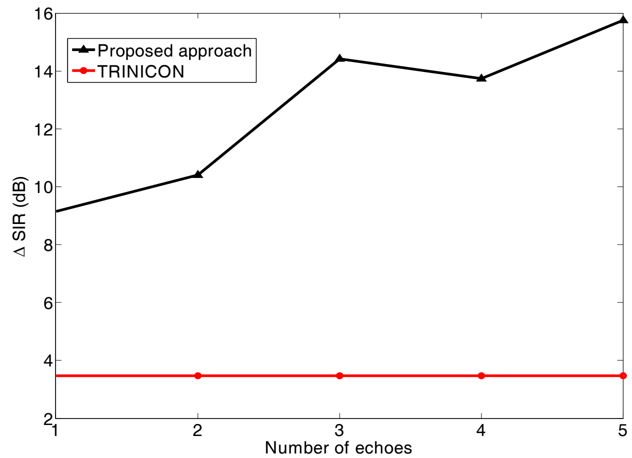


Figure 4: $\Delta SIR$ vs. the number of echoes cancelled for the proposed approach. As a reference the SIR obtained by TRINICON is also shown.

degradation of the beampattern when new constraints are added, we can notice from Figure 4 that $\Delta SIR$ slightly decreases going from 3 to 4 echoes.

## 4. CONCLUSIONS AND FUTURE WORK

In this paper we have described an algorithm for beamforming making use of a priori information acquired by the combination of localization and path tracing algorithms. The proposed algorithm is suitable in the context of moving sources, when the blind source separation techniques fail due to rapid channel variations. Experimental results confirm that the $\Delta SIR$ achieved in a simulation scenario reaches up to about $16dB$ when 5 echoes are canceled.

As motivated in Section 2, the different techniques used to perform localization, DOA computation and beamforming are interchangeable. In order to compare different beamforming algorithms, the authors are now working on the implementation of the beamforming by direct synthesis of the spatial response.

## REFERENCES


[1] H.C.Wu and J. C. Principe, "Simultaneous diagonalization in the frequency domain (sdif) for source separation," in *Proc. ICA*, 1999, pp. 245–250.

[2] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: A versatile framework for multichannel blind signal processing," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, May 2004.

[3] J.F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, pp. 2009–2025, Oct. 1998.

[4] H. Buchner, R. Aichner, J. Stenglein, H. Teutsch, and W. Kellermann, "Simultaneous localization of multiple sound sources using blind adaptive MIMO filtering," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2005)*.

[5] F. Antonacci, D. Saiu, P. Russo, A. Sarti, M. Tagliasacchi, and S. Tubaro, "Experimental evaluation of a localization algorithm for multiple acoustic sources in reverberating environments," in *Proc. of EUropean SIgnal Processing COnference EUSIPCO '06, Florence*, Sep. 2006.

[6] P.Y. Zhou and M.A. Ingram, "Pattern synthesis for arbitrary arrays using an adaptive array method," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 5, May 1999.

[7] H.L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation and Modulation Theory*, Wiley Interscience, New York, 2002.

[8] L.R. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propag.*, vol. AP-30, pp. 27–34, Jan. 1982.

[9] H.Buchner, R.Aichner, and W.Kellermann, "Simultaneous localization of multiple sound sources using blind adaptive mimo filtering," in *Proc. of 2005 IEEE International Conference on Acoustics, Sound and Signal Processing (ICASSP'05), Philadelphia*, Mar. 2005.

[10] M. Foco, P. Polotti, A. Sarti, and S. Tubaro, "Sound spatialization based on fast beam tracing in the dual space," in *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFx-03), London, Great Britain*, Sep. 2003.

[11] F. Antonacci, M. Matteucci, D. Migliore, D. Riva, A. Sarti, M. Tagliasacchi, and S. Tubaro, "Tracking multiple acoustic sources in reverberant environments using regularized particle filter," in *Proc. of 15th International Conference on Digital Signal Processing, DSP 2007*, Jul. 2007.

[12] F. Antonacci, D. Riva, A. Sarti, M. Tagliasacchi, and S. Tubaro, "Efficient tracking of multiple acoustic sources in reverberant environments using particle filtering with swarm intelligence," in *Proc. of 15th EUropean SIgnal Processing COnference, EUSIPCO'07, Poznan, Poland*, Sep. 2007.