



HAL
open science

Blind separation of acoustic sources from the principle of least spatial complexity

Bin Dong, Jérôme Antoni

► **To cite this version:**

Bin Dong, Jérôme Antoni. Blind separation of acoustic sources from the principle of least spatial complexity. International Conference Surveillance 6, Oct 2011, Compiègne, France. hal-01019695

HAL Id: hal-01019695

<https://hal.science/hal-01019695>

Submitted on 24 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Blind Separation of Acoustic Sources from the Principle of Least Spatial Complexity

DONG Bin, Jérôme ANTONI

Originally at: Laboratory of Roberval UMR 6253,
University of Technology of Compiègne, Research Center of Royallieu, BP 20529,
60205 COMPIEGNE CEDEX, FRANCE.

Currently at: Vibrations and Acoustic Laboratory,
INSA of Lyon (National Institute for Applied Sciences),
Building of St. Exupéry, 25 bis avenue of Jean Capelle,
69621 VILLEURBANNE CEDEX, FRANCE.

bin.dong@utc.fr

Abstract

A novel algorithm based on the principle of least spatial complexity is proposed to decompose an acoustic field distribution into uncorrelated acoustic sources with maximum spatial compactness. First of all, the acoustic field source distribution is reconstructed by backpropagating the acoustic pressures measured by an array of microphones. In a second step, the reconstructed source field is classically decomposed into uncorrelated sources by means of a principal component analysis. However, it is well-known that having uncorrelated sources is a necessary but by no means a sufficient condition: there exists an infinity of uncorrelated virtual sources that can produce the same acoustic field. In order to find a unique and optimal solution to source separation, it is proposed to invoke a “principle of least spatial complexity” which forces solutions with high spatial compactness. This is achieved thanks to recent results about optimization within the Stiefel manifold. An experiment is conducted to demonstrate the validity of the proposed algorithm. The final result proves that the principle of least spatial complexity and its related algorithm can successfully decompose an acoustic field into its actual constituents.

Keywords: Blind source separation; principle of least spatial complexity; spatial entropy; inverse acoustic problem; near-field acoustic holography

1. Introduction

Blind source separation (BSS) has been applied in many research areas. One of remarkable applications of BSS is to identify positions of faults in a diesel engine ^[1]. As we known, different sources originate from different phenomena. That means the total sound radiated by an engine can be simplified as a linear combination of distinct sound sources. However, due to acoustical signal complexity and the instability inherent to the backpropagation of the waves from the microphone array to the sourcesurface, classic approaches of BSS such as independent component analysis

(ICA) can hardly be applied directly. Another reason is that most of these approaches require the sources to be non-Gaussian (more exactly, no more than one Gaussian source), which is hardly verified when working in the frequency domain as is commonplace in acoustics. Indeed, when the acoustic signals are transformed from time domain to frequency domain, the Fourier Transform (FT) forces the acoustic signals to tend to a Gaussian distribution according to the central limit theorem (CLT). That means ICA could not find a relevant solution ^[2-4]. Therefore, there is still a need for alternative approaches to efficiently decompose an acoustic field into distinct sources (i.e. stemming from different origins) when these do not necessarily meet the assumptions of ICA. Acoustical array processing naturally suggests the exploitation of spatial information rather than probabilistic properties (i.e. independence). For these reasons, a method based on minimizing the spatial complexity is proposed.

The paper is organised as follows: the reconstruction of the acoustic field distribution on the surface of sources is introduced first. Then principal component analysis (PCA) is used to decompose the acoustic field into uncorrelated “virtual” sources which, however, do not enjoy the property of uniqueness. The principle of least spatial complexity is then invoked to find an optimal solution to the BSS problem. Finally, the validity of the proposed algorithm is demonstrated by a laboratory experiment. Properties of the proposed algorithm are then discussed in terms of measurements of the experiment results.

2. Methodology

2.1 Reconstruction of the acoustic field distribution

The pressure field is recorded by a microphone array delivering the M acoustic pressure signals $p(\mathbf{r}_m, t)$, $m=1, 2, \dots, M$, where \mathbf{r}_m is the position of the m -th microphone and t is time. The signal $p(\mathbf{r}_m, t)$ is divided into a series of snapshots $p(\mathbf{r}_m, t; \bar{\omega})$, where $\bar{\omega}$ stands for the snapshot label. The Fourier transform $p(\mathbf{r}_m, \omega; \bar{\omega})$ of $p(\mathbf{r}_m, t; \bar{\omega})$ will be considered from now on, where ω stands for radial frequency.

The acoustic pressure signal $p(\mathbf{r}_m, \omega; \bar{\omega})$ is assumed to be produced by n_s uncorrelated acoustic sources $s_i(\mathbf{r}, \omega; \bar{\omega})$, such that

$$p(\mathbf{r}_m, \omega; \bar{\omega}) = \int_{\Gamma} \mathbf{G}(\mathbf{r}_m \parallel \mathbf{r}, \omega) \sum_{i=1}^{n_s} s_i(\mathbf{r}, \omega; \bar{\omega}) d\Gamma(\mathbf{r}), \quad (1)$$

where \mathbf{r} is the position vector, $\mathbf{G}(\mathbf{r}_m \parallel \mathbf{r}, \omega)$ is the Green function in free space, and Γ is the source surface.

The continuous acoustic field distribution is next discretised on the source surface Γ . The position vector \mathbf{r} is replaced by N samples \mathbf{r}_l , $l = 1, 2, \dots, N$, associated with a small surface element $\Delta\Gamma_l$. The discrete form of the signal $p(\mathbf{r}_m, \omega; \bar{\omega})$ is then

$$p(\mathbf{r}_m, \omega; \bar{\omega}) = \sum_{l=1}^N \sum_{i=1}^{n_s} G(\mathbf{r}_m \parallel \mathbf{r}_l, \omega) s_i(\mathbf{r}_l, \omega; \bar{\omega}) \Delta\Gamma_l, \quad (2)$$

which may be rewritten in a matrix form as:

$$\begin{aligned} \mathbf{p}(\bar{\omega}) &= \begin{bmatrix} p(\mathbf{r}_1, \omega; \bar{\omega}) \\ \vdots \\ p(\mathbf{r}_M, \omega; \bar{\omega}) \end{bmatrix} \\ &= \begin{bmatrix} G(\mathbf{r}_1 \parallel \mathbf{r}_1, \omega) \Delta\Gamma_1 & \cdots & G(\mathbf{r}_1 \parallel \mathbf{r}_N, \omega) \Delta\Gamma_N \\ \vdots & \ddots & \vdots \\ G(\mathbf{r}_M \parallel \mathbf{r}_1, \omega) \Delta\Gamma_1 & \cdots & G(\mathbf{r}_M \parallel \mathbf{r}_N, \omega) \Delta\Gamma_N \end{bmatrix} \begin{bmatrix} s(\mathbf{r}_1, \omega; \bar{\omega}) \\ \vdots \\ s(\mathbf{r}_N, \omega; \bar{\omega}) \end{bmatrix} = \mathbf{G}\mathbf{s}(\bar{\omega}). \end{aligned} \quad (3)$$

The next step is to expand the source distribution $s(\mathbf{r}_l, \omega; \bar{\omega})$ on an optimal basis Φ with corresponding coefficient \mathbf{c} in order to allow its recovery [5]. That is:

$$\mathbf{s}(\bar{\omega}) = \begin{bmatrix} \Phi_1(\mathbf{r}_1, \omega) & \cdots & \Phi_K(\mathbf{r}_1, \omega) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathbf{r}_N, \omega) & \cdots & \Phi_K(\mathbf{r}_N, \omega) \end{bmatrix} \begin{bmatrix} c_1(\omega; \bar{\omega}) \\ \vdots \\ c_K(\omega; \bar{\omega}) \end{bmatrix} = \Phi \mathbf{c}(\bar{\omega}). \quad (4)$$

The index K of the optimal basis is equal to the number of microphones on the microphone array according to Ref. [5]. The recovery of the unknown coefficient \mathbf{c} from the measured pressure signals $\mathbf{p}(\bar{\omega})$ is obtained as

$$\mathbf{c}(\bar{\omega}) = [\mathbf{G}\Phi]^+ \mathbf{p}(\bar{\omega}), \quad (5)$$

where the exact expression of the pseudo-inverse $[\mathbf{G}\Phi]^+$ is detailed in Ref. [5] and not reported here due to lack of space. However, this is source *reconstruction* but not source *separation*: in other words, the coefficients $c_{ki}(\omega; \bar{\omega})$ assigned to the i -th acoustic source $s_i(\mathbf{r}_l, \omega; \bar{\omega})$ still remain unknown. It is the object of this paper to recover them.

In order to do so, let us first note that for uncorrelated sources the n_s coefficients $c_{ki}(\omega; \bar{\omega})$ attached to the same optimal basis Φ_k are necessarily mutually uncorrelated. Therefore, $c_k(\omega; \bar{\omega})$ may be further expanded as

$$c_k(\omega; \bar{\omega}) = \sum_{i=1}^{n_s} c_{ki}(\omega; \bar{\omega}) = \sum_{i=1}^{n_s} a_{ki}(\omega) \mathcal{S}_i(\omega; \bar{\omega}), \quad (6)$$

where $\mathcal{S}_i(\omega; \bar{\omega})$ is a ‘‘latent source’’, which is a statistical variable that does not depend on the index ‘‘ k ’’ of the basis function, and $a_{ki}(\omega)$ is the corresponding coefficient. Equation (5) may be rewritten more concisely in a matrix form:

$$\mathbf{c}(\bar{\omega}) = \begin{bmatrix} a_{11} \cdots a_{1n_s} \\ \vdots \ddots \vdots \\ a_{K1} \cdots a_{Kn_s} \end{bmatrix} \begin{bmatrix} \mathcal{S}_1(\omega; \bar{\omega}) \\ \vdots \\ \mathcal{S}_{n_s}(\omega; \bar{\omega}) \end{bmatrix} = \mathbf{A}\mathcal{S}(\bar{\omega}). \quad (7)$$

Therefore, after omitting the dependence on $\bar{\omega}$ for notational simplicity, one has

$$\mathbf{p} = \mathbf{G}\mathbf{s} = \mathbf{G}\Phi\mathbf{A}\mathcal{S}, \quad (8)$$

where parameters \mathbf{p} , \mathbf{G} and Φ are all known, but parameters \mathbf{A} and \mathcal{S} are still undetermined. The next subsection explains how to find them.

2.2 Pre-whitening of measurements

In order to find the unknown mixing matrix \mathbf{A} , let us first consider its singular value decomposition (SVD)

$$[\mathbf{A}]_{SVD} = \mathbf{U}\mathbf{D}\mathbf{V}^H, \quad (9)$$

where \mathbf{U} and \mathbf{V} are unitary matrices, \mathbf{D} is a nonnegative diagonal matrix, and superscript ' H ' stands for the transposition conjugate operation.

Next, because by definition the n_s acoustic sources are uncorrelated to each other, the latent source $\mathcal{S}_i(\omega; \bar{\omega})$ are seen to have a diagonal correlation matrix,

$$\mathbb{E}\left\{SS^H\right\} = \mathbf{I}, \quad (10)$$

where \mathbb{E} is the expectation operator of $\bar{\omega}$. Note that the power of each latent sources is assumed unitary by convention and the actual intensity is captured in coefficients $a_{ki}(\omega)$ of the matrix \mathbf{A} .

Therefore, according to Eq. (8) and Eq. (10), one has

$$\mathbb{E}\left\{\mathbf{p}\mathbf{p}^H\right\} = \mathbb{E}\left\{\mathbf{G}\Phi\mathbf{A}SS^H\mathbf{A}^H\Phi^H\mathbf{G}^H\right\} = \mathbf{G}\Phi\mathbf{U}\mathbf{D}^2\mathbf{U}^H\Phi^H\mathbf{G}^H, \quad (11)$$

which uniquely returns \mathbf{U} and \mathbf{D} from the eigenvalue decomposition (EVD) of the correlation matrix $\mathbb{E}\left\{\mathbf{p}\mathbf{p}^H\right\}$ provided that $\mathbf{G}\Phi$ is known. However, matrix \mathbf{V} which entered the SVD in Eq. (9) is still missing. This characterises the non-uniqueness of PCA (i.e. EVD of the data correlation matrix) as a solution of the BSS problem^[6, 7]. The next subsection now addresses its recovery.

2.3 The principle of least spatial complexity

To search for the matrix \mathbf{V} , the principle of least spatial complexity is introduced. Inspired from the definition of Shannon entropy in information theory^[8], a novel concept, the spatial entropy H , is proposed. The authors propose to define the spatial entropy H as

$$H = \sum_{i=1}^{n_s} \pi_i H_i = - \sum_{i=1}^{n_s} \pi_i \int_{\Gamma} P_i(\mathbf{r}) \ln P_i(\mathbf{r}) d\Gamma(\mathbf{r}), \quad (12)$$

where H_i , $P_i(\mathbf{r})$ and π_i are the spatial entropy, the intensity and the averaged power of the i -th source, respectively. Upon defining the intensity as

$$P_i(\mathbf{r}) = \mathbb{E} \left\{ \left| s_i(\mathbf{r}, \omega; \bar{\omega}) \right|^2 \right\}, \quad (13)$$

and the averaged power as

$$\pi_i = \int_{\Gamma} \mathbb{E} \left\{ \left| s_i(\mathbf{r}, \omega; \bar{\omega}) \right|^2 \right\} d\Gamma(\mathbf{r}), \quad (14)$$

one has

$$H = - \sum_{i=1}^{n_s} \int_{\Gamma} \mathbb{E} \left\{ \left| s_i(\mathbf{r}, \omega; \bar{\omega}) \right|^2 \right\} \ln \frac{\mathbb{E} \left\{ \left| s_i(\mathbf{r}, \omega; \bar{\omega}) \right|^2 \right\}}{\int_{\Gamma} \mathbb{E} \left\{ \left| s_i(\mathbf{r}, \omega; \bar{\omega}) \right|^2 \right\} d\Gamma(\mathbf{r})} d\Gamma(\mathbf{r}). \quad (15)$$

In words, the spatial entropy is constructed such that, the more complexity the source distribution, the higher the value of H and vice versa. This means that minimising the spatial entropy with respect to \mathbf{V} will favor the recovery of sources with as simple a spatial structure as possible, i.e. with some sort of compactness. This corresponds very well to our physiological perception of sound sources.

Let us now rewrite the expression of H explicitly as a function of \mathbf{V} in order to undertake its minimisation with respect to the later. The discrete acoustic source can be rewritten as

$$s_i(\mathbf{r}_l, \omega; \bar{\omega}) = \mathbf{e}_l^T \Phi \mathbf{A} \mathbf{E}_i \mathcal{S}, \quad (16)$$

where \mathbf{e}_l is a vector whose elements are all zero except the l -th one, and $\mathbf{E}_i = \mathbf{e}_i \mathbf{e}_i^T$ is a matrix whose elements are all zero except the i -th diagonal one (superscript ' T ' stands for the transposition operation). Therefore the acoustic source $s_i(\mathbf{r}, \omega; \bar{\omega})$ has intensity

$$P_i(\mathbf{r}) = \mathbf{e}_l^T \Phi \mathbf{A} \mathbf{E}_i \mathbb{E} \left\{ \mathcal{S}_j \mathcal{S}_j^* \right\} \mathbf{E}_i \mathbf{A}^H \Phi^H \mathbf{e}_l = \mathbf{e}_l^T \Phi \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \Phi^H \mathbf{e}_l, \quad (17)$$

with \mathbf{v}_i the i -th row of the matrix \mathbf{V} . After substituting Eq. (17) into Eq. (15) and discretising the fields, the discrete spatial entropy reads

$$H_d(\mathbf{V}) = - \sum_{l=1}^N \sum_{i=1}^{n_s} \left(\mathbf{e}_l^T \Phi \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \Phi^H \mathbf{e}_l \right) \ln \left(\frac{\mathbf{e}_l^T \Phi \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \Phi^H \mathbf{e}_l}{\text{tr} \left\{ \Phi \mathbf{U} \mathbf{D} \mathbf{v}_i^H \mathbf{v}_i \mathbf{D} \mathbf{U}^H \Phi^H \right\}} \right). \quad (18)$$

This cost function $H_d(\mathbf{V})$ is now to be minimised with respect to matrix \mathbf{V} ; the optimal solution will then return a unique set of compact source distributions. The difficulty here is to minimise $H_d(\mathbf{V})$ under the constraint of unitarity of \mathbf{V} . A recent and efficient conjugate gradient (CG) algorithm under unitary constraint (optimised within the Stiefel manifold) is used yet not detailed here due to lack of space – see Refs [9, 10].

3 Experiment

3.1 Experimental setup

To assess the performance of the proposed algorithm, an experiment was conducted in a semi-anechoic room. The experimental setup is shown in Fig.1. The microphone array is parallel to the loudspeaker membranes. The distance between the microphone array and the loudspeakers is represented by Z . The spacing between each loudspeaker is defined by D . The loudspeakers were taken as the actual acoustic sources. When the loudspeakers were turned on, an acoustic field was radiated from the loudspeakers to the microphone array. The pressures of the acoustic field were recorded by 60 microphones located on the array and sampled at 44.1 kHz.

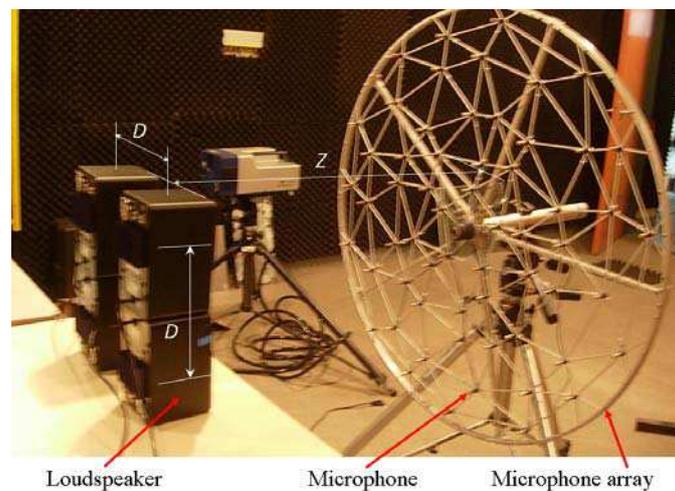


Fig.1 Experimental setup for blind source separation

3.2 experimental results

The experimental data were then processed with the proposed algorithm. First the source field on the loudspeakers is reconstructed according to the backpropagation technique described in Ref. [5]. Then it is decomposed into four distinct and uncorrelated sources by using the principle of least spatial complexity described above.

Various experimental parameters were tested. Those illustrated hereafter are for a frequency band of 1320 Hz, a distance Z equal to 10 cm and a spacing D set to 24 cm. The final results are displayed in Fig.2. The first upper left picture shows the positions of the four loudspeakers; the first lower left one presents the microphone array; the second lower left one is the acoustic pressure interpolated on the surface of the microphone array; the second upper left one is the acoustic pressure on the surface of the loudspeakers and the other four pictures are the acoustic velocity distribution separated on the surface of every single uncorrelated acoustic source. Because the spatial resolution of the acoustic velocity is much better than that of the acoustic pressure, the final results are displayed in terms of the former. It can be seen that the location of the separated acoustic sources match very well with the expected positions of the loudspeakers.

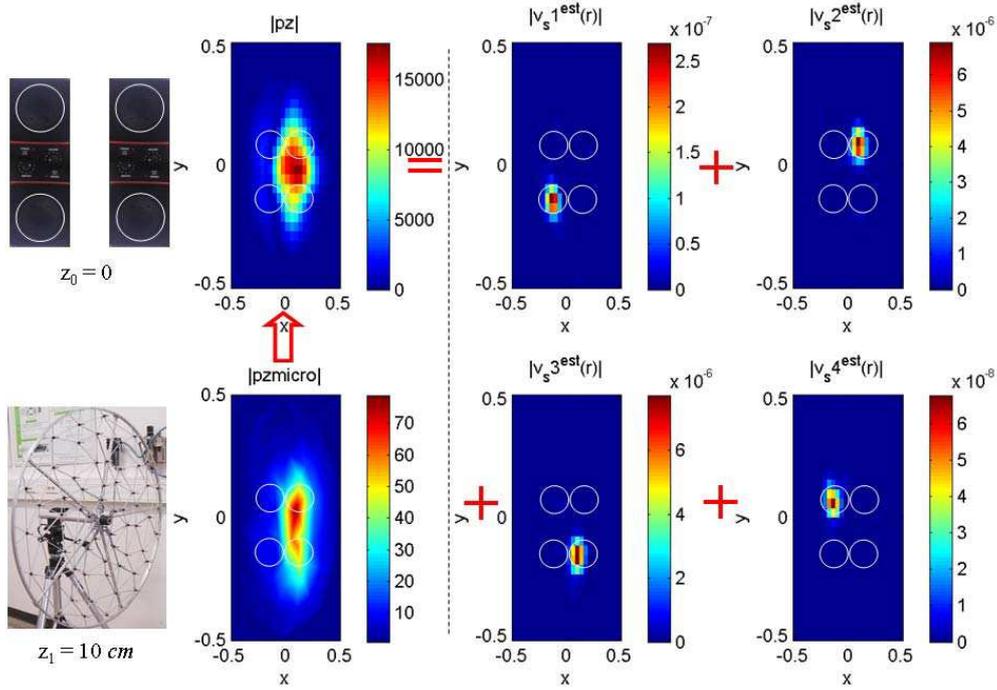


Fig.2 Separation result returned by the proposed algorithm

3.3 Discussion

A large number of similar experiments with various parameter settings demonstrated the validity of the proposed separation algorithm in most cases. However, the following limitations of the algorithm can be identified.

First, it is seen in Fig. 2, that the order of the separated acoustic sources does not correspond to the actual one. That is a fundamental indeterminacy of any BSS algorithm. However, this is not a practical limitation. In addition, in this experiment attention was paid only to the resolvability of the source field reconstruction and separation.

Second, the proposed approach is based on the fundamental assumption of uncorrelated sources, which is likely to be met in many configurations where sound sources are stemmed from different physical phenomena. This means the method will fail to decompose the correlated acoustic sources, as may be found in case of reflections for instance.

Third, the quality of the separation results is directly impacted by the performance of the backpropagation (acoustic inversion) which is a preliminary step in the proposed approach. The latter is restricted by three experimental parameters. The first one is the analysis frequency of the acoustic signal. As well-known, the lower the analysis frequency is, the more difficult the inversion is due to large wavelengths. The second one is the distance Z . The larger the distance Z , the less information is captured by the array. A critical case is the filtration of evanescent wave at low frequencies and large distances. Hence, the proposed algorithm will have excellent performance at a small distance Z to the sources and at high frequencies. The last critical parameter is the spacing D . When D tends to zero, it means the sound sources are overlapping. They will be regarded as a whole and large source, in which case source separation will no longer make sense.

4. Conclusion

The proposed source separation algorithm is based on two simple assumptions. The first one is the uncorrelation of the sound sources of interest and the second one is their spatial compactness or low “complexity”. It is believed that these two assumptions may be easily satisfied in many industrial applications. The other limitations of the method are similar to classical source reconstruction by backpropagation.

A future direction of research is to generalise the separation to the case of correlated sound sources.

Acknowledgements

First of all, we would like to acknowledge Dr. Erliang Zhang who assisted us in generating acoustic source signals and offered us many technical advices. We also acknowledge Jean-Marc Gherbezza and Félix Foucart for setting up the experiment. At last we thank Fedia Maalej for her technical help.

References

1. B. Lafon, J. Antoni. "Cyclic Sound Intensity and Source separation from NAH measurements on a Diesel engine", *The Journal of the Acoustical Society of America*, 123 (5), 2008, p. 3387.
2. Pierre Comon. "Independent Component Analysis: a new concept?", *Signal Processing*, 36(3), 1994, p. 287 - 314.
3. A. Hyvärinen, E. Oja . "Independent Component Analysis: Algorithms and Application", *Neural Networks*, 13(4-5), 2000, p. 411 - 430.
4. A. Hyvärinen, J. Karhunen, E. Oja. "Independent Component Analysis", New York: Wiley, 2001, p. 147 – 163.
5. J. Antoni. "Focalisation bayésienne: une approche unifiée du problème inverse en acoustique", CFA2010, 10eme Congres Français d'Acoustique - 12-16 Avril 2010 - Lyon, France.
6. Pearson, K. "On Lines and Planes of Closest Fit to Systems of Points in Space" , *Philosophical Magazine*2 (6), 1901, p. 559 - 572.
7. Jonathon Shlens. "A Tutorial on Principal Component Analysis", April 22, 2009, Version 3.01.
8. Lin, J. "Divergence measures based on the Shannon entropy", *Information Theory*, 37 (1), Jan. 1991, p. 145 - 151.
9. T. Abrudan, J. Eriksson, V. Koivunen. "Conjugate Gradient Algorithm for Optimization Under Unitary Matrix Constraint", *Signal Processing*, 89 (9), Sep. 2009, p. 1704 - 1714.
10. Traian Abrudan. "Advanced Optimization Algorithms for Sensor Arrays and Multi-Antenna Communications", Doctor Thesis of Helsinki University of Technology, 2008, p. 33 - 50.