IMPROVEMENT OF SOURCE SEPARATION FOR PHASED MICROPHONE ARRAY MEASUREMENTS

Igor Romenskiy, Olaf Jaeckel

Society for the Promotion of Applied Computer Sciences (GFaI e.V.)
Rudower Chaussee 30, 12489 Berlin-Adlershof, Germany
romenskiy@gfai.de, jaeckel@gfai.de

ABSTRACT

The paper presents a new method for the improvement of microphone array measurement results using an unconventional, relatively simple and very fast iterative post processing algorithm suitable for 2D-beamforming acoustic maps. This procedure is based on the mathematical model of a time reversed diffusion process which is numerically stabilized by means of a spatial low pass filter as well as by an additional amplitude renormalization step.

The method allows the reduction of the mainlobe width of acoustic point sources in the map and can thus be applied for the improvement of the usually very bad image contrast in the lower frequency range and also for a better separation of closely lying sources which can not be clearly resolved by simple delay-and-sum beamforming. An outstanding property of the proposed algorithm is that it does not rely on the knowledge of the point spread function which is needed in complete deconvolution approaches.

The iteration procedure merely utilizes the local curvature information that is already hidden within the acoustic beamforming map to sharpen the individual sources while simultaneously smoothing out disturbing high frequency components that otherwise would lead to numerical instability and unlimited amplification of unwanted small false signal and noise components.

1 INTRODUCTION

The use of microphone arrays (acoustic camera) for the fast localization and visualization of noise sources of machines and equipment of any kind has gained much popularity in real world industrial applications and in acoustic research. The underlying common principle in the farfield approach is the well known delay-and-sum beamforming method [1].

The visual interpretation of the resulting acoustic maps can be quite difficult and it requires lots of practical measurement experience and theoretical knowledge in acoustics and array processing. During the last years, many attempts have been made to generally improve beamforming results, to raise array and image contrast, to increase resolution and to undo the unwanted effects of the spatial convolution between sources and array patterns. Some of these methods originally stem from signal processing or optics, others have been developed as dedicated complete deconvolution approaches for phased array results, especially within the aeroacoustics and airplane research community.

In this paper, we present a new, fast inverse method developed at GFaI to improve the source separation capabilities and image contrast by mere postprocessing of the acoustic 2D-images. While this method can not replace the true deconvolution approaches, it may provide an attractive alternative when computing time and memory ressources are of serious concern or when the specific array pattern is not known. The procedure iteratively increases the local curvature of source peaks within the acoustic photo and it does neither need the original channel data nor the array dependent point spread function.

The paper is organized as follows. In the next section, a short overview of some existing methods for the improvement of source separation is given. In part 3, the basic ideas and the theoretical background of the new postprocessing scheme are explained. Section 4 presents some simulation results to demonstrate the basic behaviour of the proposed iteration scheme, and section 5 gives an application to a real world example. A short summary and an outlook to further research and to possible future applications in part 6 will close the paper.

2 METHODS FOR SOURCE SEPARATION

2.1 General properties of beamforming results

With beamforming, a microphone array (phased array) will be successively focused to the individual points on a measurement plane or on an object's surface by compensating for the relative runtime delays between the microphone channels and adding up the shifted time signals coherently. Normalization by the channel number then allows the determination of the effective sound pressure level at every focus point. This way, a complete mapping of the sound pressure distribution in the measurement plane or on the entire 3D-surface of a distant object can be calculated. This basic method can be computed directly in the time domain or in the frequency domain as well.

Unfortunately, the resulting image (beamforming map, acoustic photo) even of an ideal point source is never a single point, instead this point source appears smeared over a certain region of the image, showing a frequency dependent specific pattern of a broadened mainlobe and several sidelobes.

This happens because beamforming maps are always the result of a spatial convolution between the source signal's spatial spectrum and the so called point spread function (*psf*), which depends on frequency, array size and geometry (sensor count and positions), and on the focus point itself. In practice, the resulting source patterns are even spatially periodic due to the discrete sensor distribution in the array. Depending on the size of the image field and the signal frequencies, certain unwanted effects are the result. For signal frequencies too high, the spatial replica of the mainlobe will become visible in the image field, thus causing high frequency aliasing. For the lowest frequencies, on the other hand, the mainlobe will consume a huge part of the image, thus leading to generally very bad image contrasts for limited array sizes. Generally, the mainlobe width of a standard delay and sum beamformer is proportional to the wavelength and to the focus distance and inversely proportional to the array aperture.

2.2 Improvements - Overview of existing approaches

To overcome the limitations of standard beamforming, several approaches exist. Nearly all of them work in the frequency domain. Basically, there are two main groups of methods. The first group trys to separate between sources without taking array patterns into account. To this group belongs e.g. the quite powerful orthogonal beamforming [2], that is based on an eigendecomposition of the complex cross spectral matrix. Eigensystem based methods are also described in the classic array literature [3] for signal classification purposes etc.

The second group of methods instead trys to separate the individual sources from the influences caused by the specific array pattern. These are inherently deconvolution methods. One well known approach, CLEAN, attempts first to remove the sidelobe pattern of the strongest point source in the acoustic image, thus allowing weaker sources to become visible. This procedure can be used iteratively and has recently been extended to take spatial source coherences into account [4]. While the CLEAN-method could be seen as a pointwise or partial deconvolution approach, the actual state-of-the-art is a complete deconvolution.

Here the goal is to remove the array influence for all sources and all signal frequencies and for all focus points simultaneously. As these are inverse methods, all of them need a form of numerical regularization. Most famous here are DAMAS [5] and various following developments [6], the newer of them also including source coherences [7]. These modern methods allow for a better quantitative analysis and are very powerful, but they are also still very ressource demanding.

2.3 Demands for a more practicable method

For practical use within the Acoustic Camera, we have the need for a method that also allows for a better separation of sources, especially in the lower frequency range. At the same time, this approach should be much less demanding, especially concerning computing time and memory ressources, than the above mentioned methods. Because our system is mainly used as a fast troubleshooting technique, we do not need exact quantitative results in any case.

On the other hand, we need a procedure that is well suited for broadband signals (avoids spectral domain computations), that can postprocess existing acoustic photos (perhaps without the channel data) and that avoids, if possible, the need of simulating the array *psf*. Additionally, it should be possible to extend the results to the more general 3D-case.

3 NEW METHOD

3.1 Basic Ideas

The above mentioned needs are of course not easy to meet. It is clear that a complete deconvolution approach is out of question under these conditions. A partial deconvolution without knowledge of the *psf* is also difficult to imagine. But we can ask to ourselves the following more basic question: Is it possible to undo some of the destructive effects of convolution when we only know *that* our acoustic images are the result of a spatial convolution process, but we do not know exactly *how* this convolution was performed?

Surprisingly enough, the answer to this question is positive, at least partially. In the broadband case that we generally consider here, most of the sidelobes will average out for nonmoving sources, so we will only concentrate on the broadening of the mainlobes in the following. This broadening and smearing of the mainlobes in space has some loose similarity with a physical diffusion or heat conduction process [8]. An initially concentrated parameter will be continually distributed in space under the dynamics of the system. The behaviour of the mathematical model of such a process is also nonlocal, exactly as it is with the convolution kernels of the array *psf*s. After a certain time, every point in the resulting (image) field will contain information from all the formerly concentrated source points. Sources will merge and can not be well separated anymore.

While both processes (diffusion and convolution) are not exactly the same, they share this important common property of nonlocality. So, what happens now when we simply reverse the time parameter t in the more general diffusion model equation? Can we achieve a virtual "deconvolution" this way?

3.2 Background and Algorithmic implementation

The mathematical model used in physics for the description of a diffusion (or also a heat conduction) process is a partial differential equation (PDE) of parabolic type and has the general form:

$$\frac{\partial U(x,t)}{\partial t} = D(x,t,U) \cdot \frac{\partial^2 U(x,t)}{\partial x^2} \quad \text{with} \quad D > 0. \tag{1}$$

Here, the onedimensional form in space is used just for convenience of the following explanations. The parameter U denotes the physical quantity of interest and will be the image function in our case. The coefficient $D(\cdot) > 0$ is a positive definite function of its arguments. In the linear case $D = const > 0$, it can be easily shown by a Fourier transform to the wave number domain that this system dynamics always acts as a spatial low pass filter. The product $D \cdot \partial^2 U/\partial x^2$ in (1) is the analogon to a convolution kernel of an array *psf*.

Inverting the direction of time now will have two effects. First, the system is no more of the parabolic type any longer. Second, it will now act as a nonlinear amplifier for the higher frequency energy components. While the second property is exactly what we must utilize to exploit the hidden information about the fine structure in the acoustic maps, this is also very risky.

The time reversed PDE can be discretized with a finite difference scheme and gives the following simple iteration procedure:

$$U^{(n+1)}(x) = U^{(n)}(x) - \gamma \cdot U^{(n)}(x) \cdot \frac{d^2U^{(n)}}{dx^2} = U^{(n)}(x) \cdot \left(1 - \gamma \cdot \frac{d^2U^{(n)}}{dx^2}\right). \tag{2}$$

Here, the parameter $\gamma > 0$ allows to determine the amount of high frequency amplification introduced into the system. For a (hidden) local maximum point in the image, the local curvature will be equivalent to the second derivative, which is negative then. Since γ must be positive, no matter how small we might choose this amplification constant the system will always be numerically instable, so the direct implementation of (2) is not possible. HF-energy simply will grow to infinity during the iteration cycles, thereby destroying the information content in the lower frequency regions also. Therefore, as it is the case with all inverse methods [9], also this system needs an external regularization. We introduced two different mechanisms for stabilization.

First, an additional low pass filter in the spatial domain is added after each iteration step of Eq. (2) to eliminate the unwanted high frequency components that otherwise would amplify the noise and small but uninteresting false signal components without limits. This filter step is described by

$$U_{LPF}^{(n+1)}(x) = \int F^{(n)}(x'-x) \cdot U^{(n+1)}(x')\, dx'. \tag{3}$$

In (3), the smoohting kernel F is a symmetric function with finite support, and it may also depend on the number n of iteration steps. With this additional filtering, the iteration process now acts as a tunable bandpass type filter in the spatial domain. Nonlinear HF-amplification in the filter's passband is allowed, but disturbing higher frequency components are suppressed. The width of this filter must be choosen according to the mainlobe width of the seeked hidden peaks in the acoustic map. While the use of (3) actually stabilizes the algorithm now, it is important to apply the filter step always after the HF-amplification step (2), in order to allow the system to first nonlinearly amplify the hidden curvature information in the map.

The second mechanism of regularization is needed because the nonlinear behaviour of the local curvature amplification also urgently demands the stabilization of the system's magnitude response. The iterations show maximum efficiency when the image functions and the local curvature values would have the same order of magnitude. For amplitude stabilization a renormalization to unity was introduced into (2), now having the form:

$$U^{(n+1)}(x) = U^{(n)}(x) \cdot \left(\frac{1}{n+1} - \gamma \cdot \left(\sigma^{(n)}\right)^2 \cdot \frac{d^2U^{(n)}}{dx^2}\right). \tag{4}$$

In this modified form of the amplification step, the parameter $\sigma^{(n)}$ is an (unknown) average width of the image function. This renormalization step can only guarantee that the peak maximum level will be stable, but the metrics in the rest of the image can not be conserved. So, the method is only a qualitative one. Generalization of (4) and (3) to the 2D-case is relatively straightforward and will therefore not be shown in this paper.

4 SIMULATION RESULTS

In Fig. 1, the resulting image function of two closely lying peaks is shown in the 1D-case. While the image function itself has just one local maximum (the two real sources are "merged" together), the analysis of the first (red curve) and second derivative (blue curve) clearly shows the hidden peak maxima. This situation could well stem from a cross section through an acoustic 2D-photo. Increasing the local curvature allows separation of the two peaks, which works perfectly in the undistorted case.

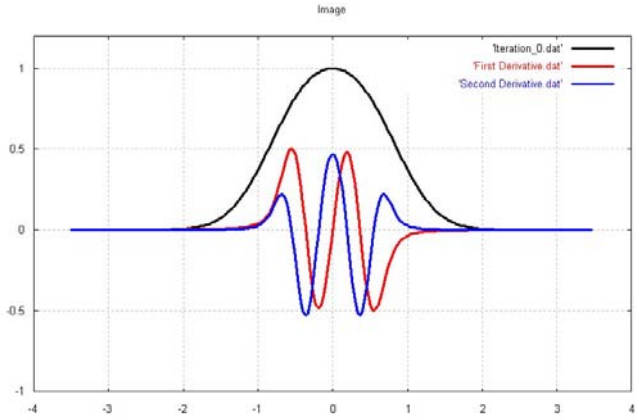

Fig. 1. Simulation of two hidden peaks in the 1D-case. The higher moments allow for separation.

The more realistic case of a noisy image function is shown in Fig. 2 a and b. Note the high amount of noise especially in the second derivative (blue curve). Fig. 2b shows the result of the local curvature enhancement by applying Eq. (4) and (3) for just 6 iteration steps. The HF-disturbances outside strongly curved regions are very effectively filtered out. The first and second derivative now clearly show the two peaks again, which are now also visible in the reconstructed image function. Since both peaks had the same strength initially, the noise only introduced a slight asymmetry into the end result here. In this case, the amplitude renormalization scheme proved to be effective. During the simulations, the method began to fail for noise levels greater than approximately 4% to 5% of the maximum amplitude.

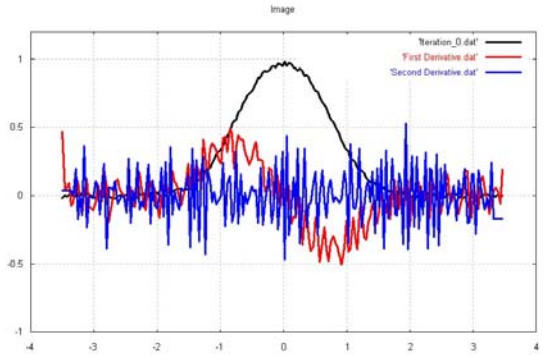

Fig. 2a. Same function as in Fig. 1 but with 2% noise added to the image function.

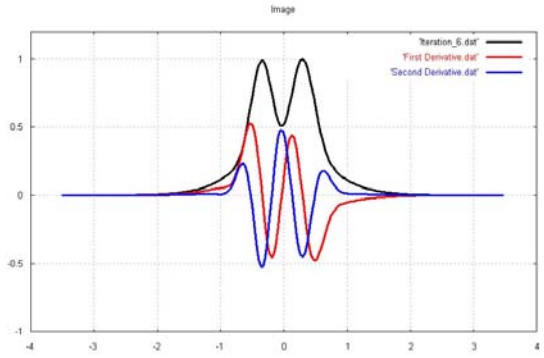

Fig. 2b. Result after 6 iterations of the separation method. Local curvature has been increased.

5 APPLICATION EXAMPLE

The method was applied to real acoustic maps of the low frequency emissions of a car door. The array used was our new standard symmetric ring array with 48 channels at 2m measurement distance. The sampling rate was 192 kHz. A white noise source was active inside the car. While the original task to detect high frequency leakages at the door sealings around 5 kHz to 6 kHz is no problem for the Acoustic Camera, the lowpass filtered signals below 1200 Hz yielded very bad image contrasts and did not allow for source separation (Fig. 3a. and 4a.). In Fig. 3a., the user could surmize that there is more than one source due to the slight asymmetry of the emission, but there is only one visible local maximum.

The new postprocessing method was convergent after only 2 iterations, because this map is much smoother than the example above. Manual tuning of the spatial low pass filter and amplification parameters was still necessary. The results clearly show the application potential of the proposed method. The right emission was an actual lower frequency door leakage, while the left emissions in Fig. 3b. and 4b. show the true speaker position within the door.

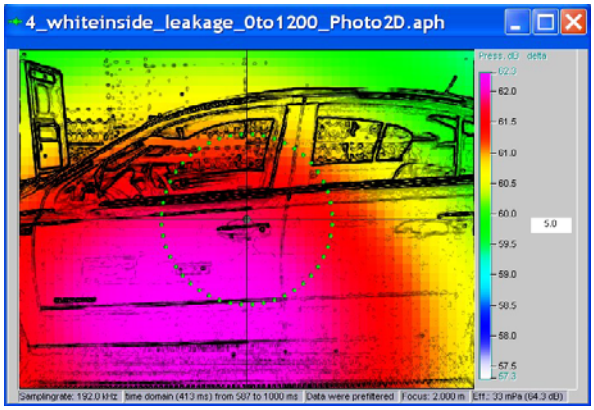

Fig. 3a. LF emissions from 0 to 1200 Hz at a car door. Image contrast is 5dB.

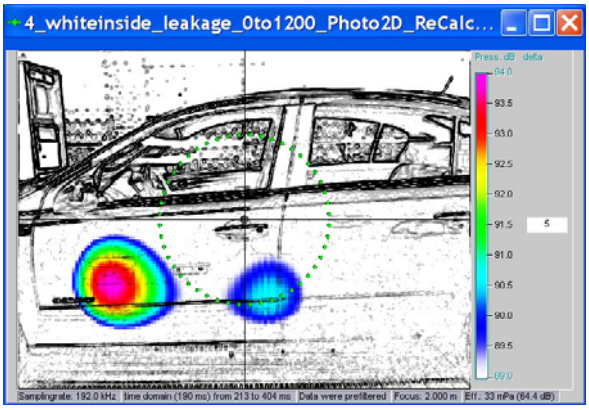

Fig. 3b. Result after 2 iterations, contrast 5dB.

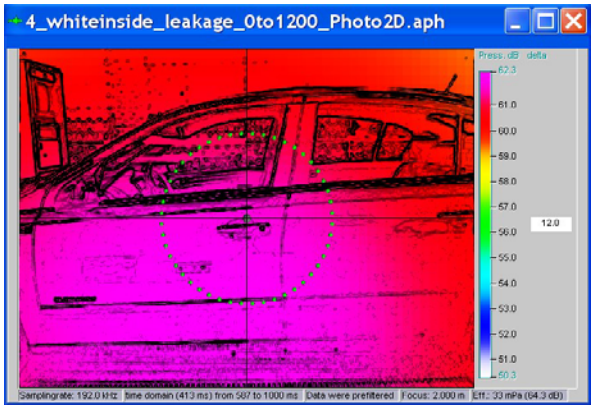

Fig. 4a. LF emissions from 0 to 1200 Hz at a car door. Image contrast is 12dB.

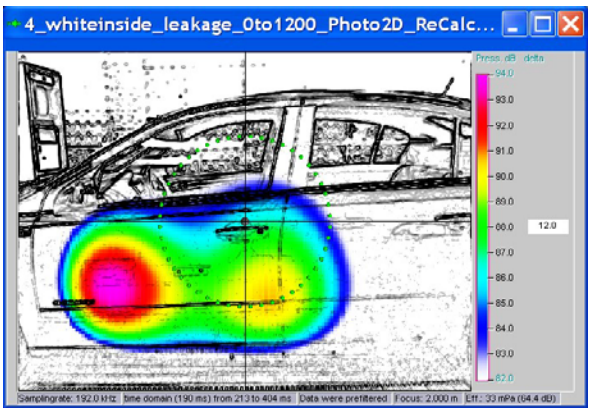

Fig. 4b. Result after 2 iterations, contrast 12dB.

6 SUMMARY

The paper presented a new and promising method for the improvement of image contrast and a better source separation in acoustic images. It is a mere postprocessing method and it does not rely on the knowledge of array point spread functions. The procedure has low computational and memory ressource demands and is very well suited for broadband signals. For typical acoustic maps, the algorithm usually converges within only two or three iteration steps.

As with any method with limited a priori information, it clearly has its drawbacks. It can not replace the more advanced full deconvolution methods, and it can fail when noise level is too high, sources are too close or when there are many sidelobes in the map. In the future, we will optimize the spatial low pass filter, perform much more simulations and practical applications, integrate the method into the NoiseImage software and try a generalization to the 3D-case. In addition, inclusion of the array *psf* for simple geometries (ring) can be an option.

ACKNOWLEDGEMENT

This research work has been funded by the German **BMWi** under project registration number **IW - 06 1211**.

REFERENCES

[1] Johnson, D.H.; Dudgeon, D.E.: Array Signal Processing.Concepts and Techniques. Prentice Hall PTR: Englewood Cliffs, NJ, 1993.

[2] Ennes Sarradj, Christian Schulze, Andreas Zeibig: Einsatz eines Mikrofonarrays zur Trennung von Quellmechanismen. DAGA, München 2005.

[3] H. van Trees: Optimum ArrayProcessing. J. Wiley & Sons, 2002.

[4] Sijtsma, P.: CLEAN Based on Spatial Source Coherence. Proc. of 13th AIAA/CEAS Aeroacoustics Conference, Rome, Italy, 2007.

[5] Brooks, T.F.; Humphreys, W.M.: A Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS) Determined from Phased Microphone Arrays. Proc. of 10th AIAA/CEAS Aeroacoustics Conference, Manchester, UK, 2004.

[6] Dougherty, R.P.: Extensions of DAMAS and Benefits and Limitations of Deconvolution in Beamforming. Proc. of 11th AIAA/CEAS Aeroacoustics Conference, Monterey, Cal., 2005.

[7] Brooks, T.F.; Humphreys, W.M.: Extension of DAMAS Phased Array Processing for Spatial Coherence Determination (DAMAS-C). AIAA Paper 2006-2654, May 2006.

[8] Beck, J.V.; Blackwell, B.; St. Clair, C.R.,Jr.: Inverse Heat Conduction: Ill-posed Problems. John Wiley & Sons, Inc., 1985.

[9] Lattès, R.; Lions, J.-L.: Méthode de quasi-réversibilité et applications. Dunod, Paris, 1967.