Analysis of the robustness of various advanced beamforming algorithms in comparison to the classical beamforming method when applied in reactive sound fields

Clemens Nau[1], Johannes Klein[2],
Werner Moll[1], Michael Vorländer[2]

[1] *Daimler AG, Béla-Barényi-Straße 1, 71063 Sindelfingen, Germany*
Email: Clemens.Nau@daimler.com
[2] *Institute of Technical Acoustics, RWTH Aachen University, Kopernikusstr. 5, 52074 Aachen, Germany*
Email: Clemens.Nau@akustik.rwth-aachen.de

Motivation

The conventional beamforming (CBF) is a valid method for the spatial localization of acoustic signals in free field. In order to use this method as a tool for the detection of acoustic signals in reactive sound fields, a modification of the applied signal processing is necessary. This is realized in the form of advanced or adaptive beamforming algorithms and a generalized correlation technique used as a preprocessing of the cross spectral matrix (CSM). This study is a comparison of the localization performance of advanced algorithms and the CBF, and their advantages when used in reactive or modal sound fields.

Beamforming and Processing

The output of the CBF can be expressed in the frequency domain, its performance can be formulated as:

$$b(\boldsymbol{g}) = \boldsymbol{g}'\,\boldsymbol{C}\,\boldsymbol{g} \qquad (1)$$

at which $\boldsymbol{g}$ is the array steering vector ($\boldsymbol{g}'$ conjugate-complex) and $\boldsymbol{C}$ the cross spectral matrix. Due to the formulation of the CBF it is readily apparent that any components which are correlated to sound source contained in the CSM are considered equally [1]. Adaptive beamforming algorithms such as the Robust Capon algorithm, the Functional Beamforming (FUBF) or the MUSIC algorithm show significant advantages over the CBF as concerns dynamics and resolution. Furthermore, studies give evidence that especially the MUSIC-algorithm is able to provide accurate localization results under the influence of a reactive sound field [2]. From the structure-borne sound acoustics and speech processing methods are known which may detect the presence of a radiating source and estimate the signal travel time difference at physically separated sensors, when energy of this source is received at the sensors. This generalized correlation technique, called Smoothed Coherence Transform (SCOT) is formulated as [3]:

$$\boldsymbol{C}_{nk\,SCOT} = \frac{\boldsymbol{C}_{nk}}{\sqrt{(\boldsymbol{A}_k * \boldsymbol{A}_n)}} \qquad (2)$$

with the cross correlation C, row vector A with the diagonal entries of C (thus the auto-correlation). Originally developed for two sensors the transformation applied to the cross spectral matrix acts as a kind of correlation filter. By appropriate weighting of the matrix entries with regard to their correlation to the desired signal, this technique suppresses the correlated components [4].

Experimental Scenarios

In the following sections the robustness of advanced beamforming algorithms in comparison with the CBF based on two experimental scenarios is studied. The first scenario "defined positions of a reflection face" represents the simple case of a single reflecting surface in the free field. The second scenario, however, represents the most difficult case, a distinctive modal sound field in an enclosure.

Defined positions of a reflection face

In the first scenario, a sweep signal (20 Hz to 20.000 Hz) is emitted in the direction of a spherical microphone array (32-channel sphere array) through a loudspeaker. The loudspeaker and the array are located at a height of 1.24 m from the reflecting floor and are placed at 1 m distance from each other. The studies compared the localization results of the various beamforming algorithms in an acoustic semi-anechoic room are divided into the following variants:

- Without reflection face (not shown) (Case 1)
- Reflection face behind the array (Case 2)
- Reflection face laterally to the speaker-array-axis (Case 3)

Fig. 1 represents the measurement configurations described.

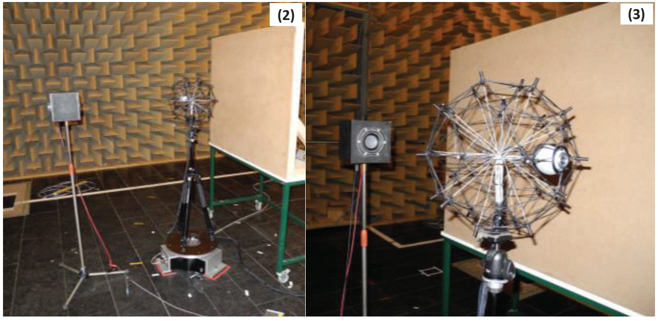


Fig. 1: Experimental setup with different positions of the reflection face

Using the examples of greatly simplified reactive sound fields (2) and (3) the behavior of the CBF compared to advanced beamforming algorithms (here the MUSIC algorithm) is to be shown. In the case of the CBF, it is expected that the reactive sound field caused by the introduced reflecting surface has an influence on its localization result and performance. The plots show the relative sound pressure distribution (the so-called beamforming map) with a dynamic range of 3 dB shown in 360^0 azimuth and elevation 180^0.

Beamforming results: Reflection face behind the array (2)

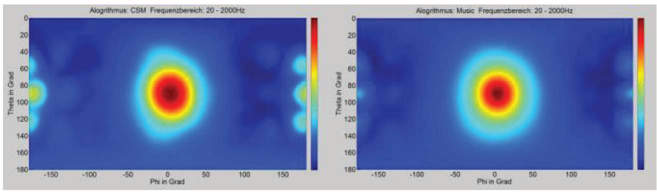


Fig. 2: Beamforming map (20 – 2000 Hz), reflection face behind the array (CBF left, MUSIC right)

It can be seen that with both the standard beamformer, the CBF, as well as the MUSIC algorithm, the source can be easily identified. The result differs only by the higher dynamic range and the associated lower artifact formation, which can be seen on the right and left edges of both plots.

Beamforming results: Reflection face laterally to the speaker-array-axis (3)

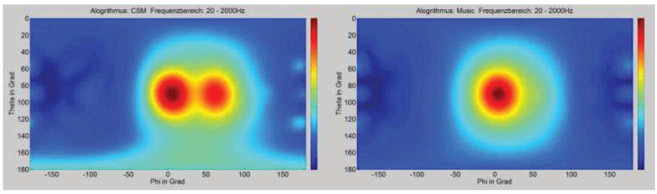


Fig. 3: Beamforming map (20 – 2000 Hz, dynamics 3 dB), reflection face laterally to the speaker-array-axis (CBF left, MUSIC + SCOT right)

This example illustrates the sensitivity of CBF towards reflections, particularly in the direct environment of the detected source. Although in this case again the same sweep signal is emitted via the same speaker, the CBF detects the reflection, which is radiated from the lateral reflection face in the direction of the array as a second source. The MUSIC algorithm, combined with the SCOT technique is, however, capable of the incident and suppresses the reflection to detect the location of the real source.

Modal sound field of a scale model room

The sound field of an enclosed space is characterized by its distinctive modal sound field. Below the so-called Schroeder frequency (f_s) the modal influence dominates the sound field, thus ensuring a strong, dynamic sound pressure distribution [5]. The conventional beamforming is particularly disturbed by these distinct modes below the Schroeder frequency, since the free-field assumption, which it required for the CBF, has no validity [1].

The figure shows a scale modal room with a volume of 1.6656 m^3, in which the subsequent series of investigations is performed.

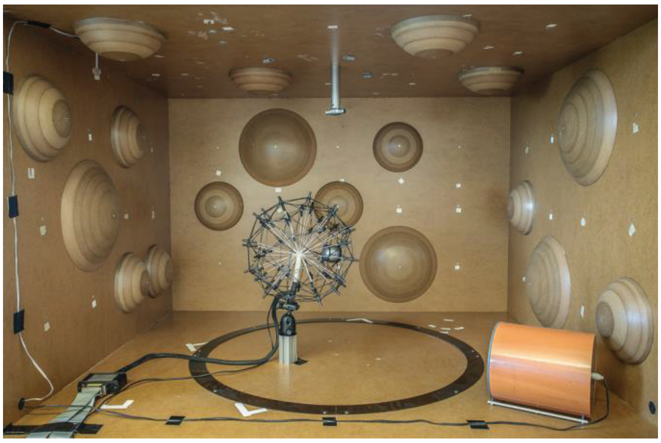


Fig. 4: Scale model room

Via the cylindrical loudspeaker box shown in the front (Fig. 4) a sweep signal between 20 Hz and 20,000 Hz over a period of 23.78 s is radiated again. This signal impinges at the channels of microphone array showing a spectral sound pressure distribution which can be characterized by the reverberation time of this room. For the analyzed room, a Schroeder frequency of 1463 Hz (marked in the figure) can be calculated from the measured reverberation time of 0.892 s.

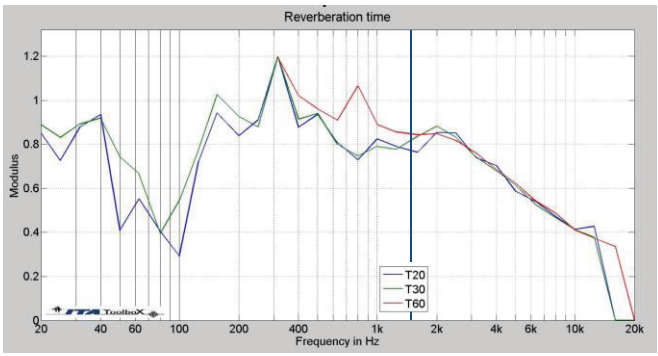
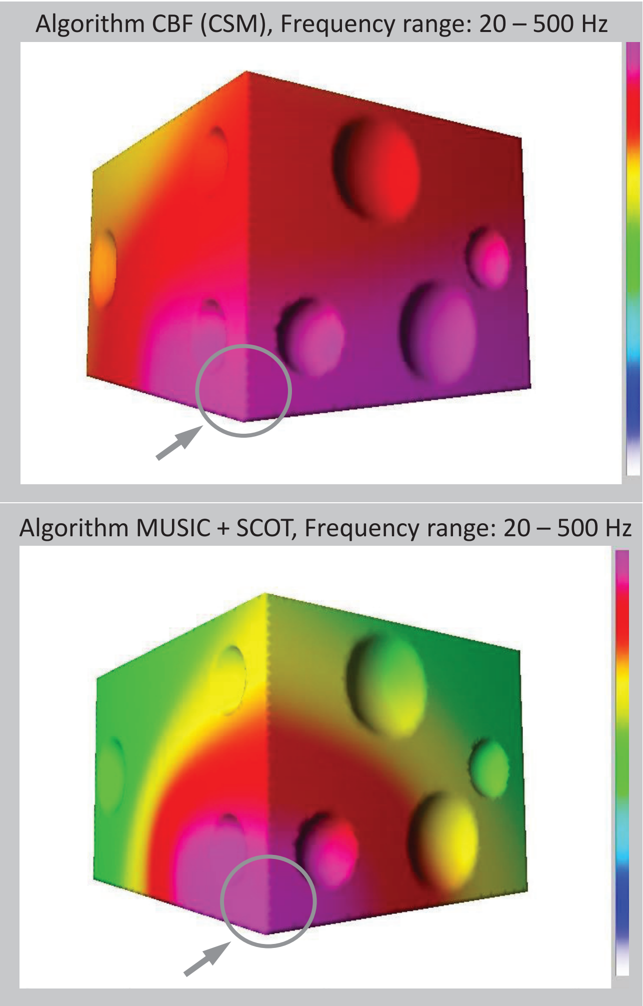


Fig. 5: Reverberation time of the scale model room

Under these conditions, similar to the scenario with defined reflection faces, the performance with respect to the localization accuracy of various beamforming algorithms compared to the CBF is investigated. For this purpose, the shown scale model room is transferred to a CAD model and adjusted to the measurements previously performed on this model. The algorithms implemented in Matlab are then applied to the model.

Fig. 7 shows the results of this comparison. The sound source is located at the selected position. Here the three-dimensional mapping (relative sound pressure distribution) of the CBF or CSM and the mapping of the MUSIC algorithm in combination with the SCOT method between 20 Hz and 500 Hz and a dynamic range of 3 dB is exemplary compared.



Algorithm CBF (CSM), Frequency range: 20 – 500 Hz

Algorithm MUSIC + SCOT, Frequency range: 20 – 500 Hz

Fig. 6: 3D-beamforming-map (20 – 500 Hz, dynamics 3 dB), Localization result of the CBF (top), Localization result of the MUSIC + SCOT (bottom)

Results

The advanced beamforming algorithms show far better results in terms of source localization compared to CBF when applied in reactive respectively modal sound fields. This improvement is shown on the one hand in terms of simple reflections as in the example defined reflecting surfaces, on the other hand also in the much more demanding environment of a closed space. The benefits of advanced beamforming algorithms with respect to their localization performance, in this study illustrated by the example of the MUSIC algorithm can be attributed exclusively to the evaluation of the cross spectral matrix. This demonstrates the eigenvector analysis of the MUSIC algorithm to be significantly more robust against noise field influences than the CBF. Furthermore, the preprocessing using the SCOT technique proves to be an effective tool for suppression of correlated, so modal, sound field effects. Although the preprocessing with the SCOT technique (Generalized Correlation Method) is basically applicable to almost every Beamforming algorithm, the combination of the MUSIC algorithm with the SCOT technique has shown to be particularly robust with respect to the studied sound field effects.

References

[1] Van Trees H. L., "Optimum Array Processing – Part IV of Detection, Estimation, and Modulation Theory" Wiley-Interscience, John Wiley & Sons Inc. Publication, 2002

[2] Nau C., Vorländer M., "Comparison and evaluation of localization results of synthetic and real acoustic excitations using various beamforming algorithms in a vehicle interior", Aachen Acoustics Colloquium, 2014

[3] Kuhn J. P., "Detection Performance of the Smooth Coherence Transform (SCOT)", Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP, 1978

[4] Benesty J., Chen J. "Microphone Array Signal Processing", Springer Topics in Signal Processing, Vol.1, Springer-Verlag Berlin, 2008

[5] Kuttruff H., "room acoustics" Fourth edition, Institute of Technical Acoustics, RWTH Aachen University, Spon Press London, 2000