

Advanced Source Localization Techniques Using Microphone Arrays

2011-01-1657

Published
05/17/2011Sandro Guidati and Roland Sottek
Head acoustics GmbH

Copyright © 2011 SAE International

doi:[10.4271/2011-01-1657](https://doi.org/10.4271/2011-01-1657)

ABSTRACT

Microphone arrays used in vehicle acoustics are mainly designed for fast setup and basic evaluation (e.g. using delay-and-sum beamforming) resulting in a restriction to free field environments. Applications in vehicle interiors require advanced source localization techniques taking into account the reflections at the different panels appearing as mirror sources.

Coherence filtering techniques allow for the detection of these mirror sources. An additional sensor is placed as a reference close to the main source. This reference signal is used to filter the array signals increasing the overall dynamic range of the acoustic source mapping. The discrimination of the original source and the reflections is obtained by manipulating the impulse responses between the reference signal and all microphone signals.

Other advanced signal processing techniques can be used to increase the limited dynamic range of conventional beamforming such as principal component analysis removing the dominant sources. An implementation allowing for online application based on spatial filters will be presented.

In addition the production requirements of large scale arrays, for example wind tunnel measurements and pass-by or fly-over measurements, are discussed. Manufacturing a large scale array with an appropriate precision for the microphone positions and the video system would be very expensive. A very efficient solution turns out to be a combination of a small calibrated camera module (consisting of a video system plus a small number of microphones with high-precision positioning) and a large scale array containing microphones with low-precision positioning. Array measurements using a standard loudspeaker for excitation serve as input for

nonlinear optimization techniques to identify the positions of all microphones based on the orientation of the camera module. The techniques also allow for combining small arrays to increase the spatial resolution.

This paper describes the theoretical background of different source localization techniques based on microphone arrays. Application examples are used to illustrate not only their advantages but also their potential drawbacks.

INTRODUCTION

The basic theory of array technology is known and has been used since the beginning of the 20th century¹. At the introduction, collectors were used to enhance the sound coming from a single direction; the steering had to be done manually. Later the first arrays using transducers were built, principally for submarines, here using hydrophones. The superposition of the sensor signals was done electronically. With the progresses in computer technology, array technology also evolved. Nowadays, standard PCs and laptops without the need for specialized hardware allow for the measurement of a large number of microphones as well as the processing and visualization of high-resolution acoustic images online and with very low latency.

BEAMFORMING

One of the basic algorithms used for microphone array source localization is called delay-and-sum or beamforming¹. The sound signal radiated by the sources of interest is measured by an array of microphones with nominal omnidirectional measurement characteristic (other measurement characteristics have been used, e.g. spherical beamforming). By superposing the microphone signals the array is steerable electronically towards any direction or point of interest. In the

basic delay-and-sum algorithm the steering is performed simply by delaying the microphone signals individually to compensate the propagation time between the assumed source position (or direction) and the individual array microphone. The algorithm itself has no constraint for the assumed radiation characteristic used for propagation time calculation. Typically for near-field sources a monopole radiating in free field is assumed. For sources in the far field a plane wave assumption is made. In general, the assumed radiation characteristic should match the radiation characteristic of the real sources as well as possible². For example, the noise radiated by the trailing edge of an airfoil is best modeled by a dipole³. The algorithm also has no constraint on source position. Typically a scan-plane parallel to the array plane is defined on which the strength of potential sources is calculated, but three-dimensional scan surfaces are possible without changing the algorithm. The resulting acoustic image is finally superposed with a video image to allow for the identification of the sources at the observed measurement object.

FLEXIBLE ARRAYS

The majority of commercially available beamforming systems are highly integrated with a fixed microphone pattern and a small number of cables to ensure a fast and easy setup, since time-to-result is one of the key issues for industrial applications. The array pattern is a compromise between size and portability, between a large number of microphones and the resulting cost.

For the best results, the positions of the microphones should be optimized for the current measurement situation. For example, a highly directive source requires an array with a small aperture. Otherwise the comparison with monopoles will not lead to correct results. On the contrary, the aperture must be large to get good resolution for low frequencies. The same applies for sources that are far away. To cover a wide range of measurement situations the array should be made of a large number of individual microphones positioned in some kind of modular structure. Such a system consists of many cables and requires significant time for setup. One of the basic principles of beamforming is that the position of the microphones, the pattern, could be almost arbitrary following some basic rules, but must be known very precisely referenced to the system's video device. Providing this precision also for large arrays leads to high manufacturing costs.

One solution to this problem is to divide the system into two parts: first, a camera module consisting of a video device and a small number of microphones (figure 1). These microphones are positioned with high precision. The alignment with the video device is calibrated. The video device defines the center of the coordinate system. Second,

the remaining microphones are positioned with lower precision, and optimized for the measurement situation. The exact position of these microphones in the coordinate system is calculated using basic measurements and non-linear optimization algorithms.

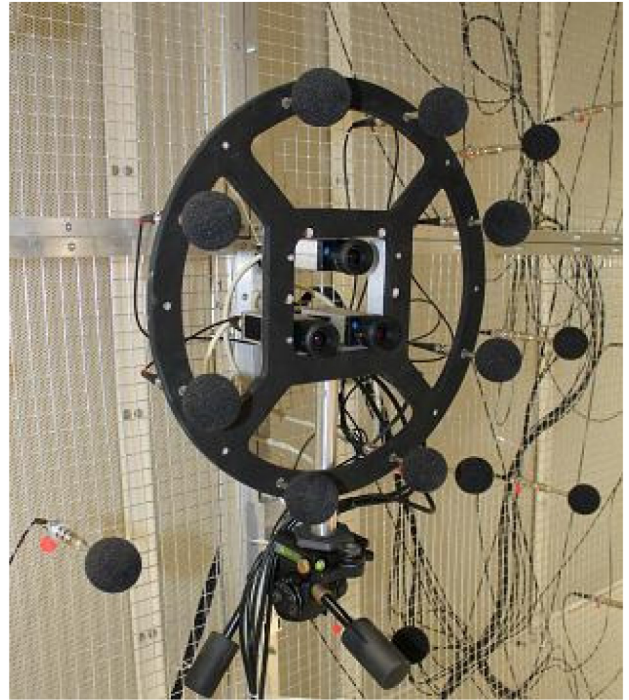


Figure 1. Camera module with three video cameras for automated distance detection and eight fixed microphones.

DISTANCE MEASUREMENTS

The basic idea is to measure the distances between the known and unknown microphone positions and a reference source moved randomly in front of the array (figure 2). The source should have almost monopole directivity, but a small loudspeaker has proven to be good enough. The loudspeaker is excited with a known signal. The best results were achieved with a bandlimited pulse (2-7 kHz) emitted synchronized with the array's data acquisition system. Most data acquisition devices have at least one digital-to-analog (D/A)-converter as a monitor output that could be used for this purpose. The synchronization between D/A and A/D conversion allows for very simple detection of the time delay between the output signal and the recorded signals, and finally the distance between the source and sensors. The measurement is executed until approximately 50-80 linear independent measurements have been acquired (the number should be larger than the number of microphones that must be localized in one calculation step). Linear independency is detected by calculating the scalar product between the current measurement vector (all calculated distances) and the prior measurements. To reduce the computation time it proves to

be sufficient to check the linear independency between the current and the last measurement. Thereby, the preparation time, when the loudspeaker is not moved at all, is ignored.

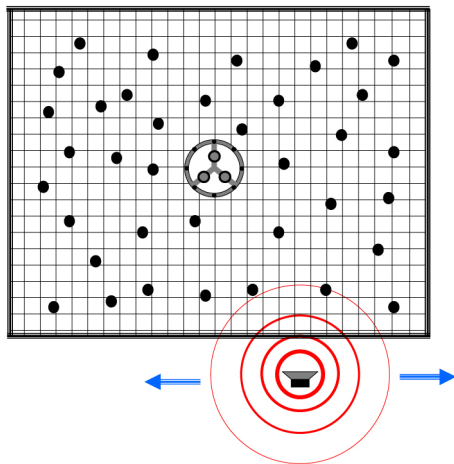


Figure 2. Measurement with a loudspeaker.

NON-LINEAR OPTIMIZATION

In the next step the unknown microphones and source positions must be calculated using the measured distances. The distance is a second-grade function of the position. Therefore it is not possible to calculate the positions directly by solving a set of linear equations. Instead non-linear optimization is applied. A good compromise between accuracy and time-to-result can be achieved using the Levenberg-Marquardt algorithm⁴. The already-known positions of the microphones integrated in the camera module are included in the optimization strategy. Usually there is a fixed time delay in the measurements caused by the computer latency as well as delays caused by the amplifier, leading to a distance offset. This value is the same for all measurements. It can be modeled as an additional variable in the optimization. The same applies for the often unknown air temperature. Because the positions of the fixed microphones in the camera module are precisely known, all unknown coordinates will converge to their correct values. The problem can be divided into subsets of microphones taking advantage of multi-processor architectures. A standard quad-core PC takes approximately 5-15 seconds to calculate 100 microphone positions. The resulting accuracy depends on the quality of the measurements and first of all on the directivity of the source. Typical results are in the range of ± 5 mm.



Figure 3. Grid array (3 m × 3 m) with integrated camera module.

ARRAY DESIGN

The procedure described above is not based on any assumptions about the array pattern. Typical patterns used in industrial systems are logarithmic spirals, circles and spheres. All have well-known characteristics and are optimized for a variety of applications, but it is sometimes difficult to alter these patterns to fit a specific measurement situation due to limited space or the effort needed to build extremely large arrays. One of the basic rules of array design is to avoid redundancies. Here the important quantity is not the position of the array microphones but the vector between two microphones. The spatial distribution of all vectors is called the coarray. Each vector is directly correlated with the phase delay between two sensor signals in the spatial sampling process. This is the reason why a regular microphone grid is not a good choice for beamforming. Here the resulting vectors are highly redundant. Then, to avoid spatial aliasing one must increase the number of channels (increasing the cost). A circular array e.g. has almost no redundancy. The same applies for the logarithmic spiral. Another good way to avoid redundancies is to use random patterns. A random pattern may have a limited performance for single frequencies (increased side lobes) but the overall performance will be in the same range as e.g. the logarithmic spiral.

The array used for the measurements presented in this paper is built using steel grid panels (figure 3). One panel has a dimension of 1.5 m × 1.5 m. Multiples of these panels can easily be combined to build larger arrays. The grid surface is