AUTOMATIC DETECTION OF MICROPHONE COORDINATES

[1] Dirk Döbler
GFaI e.V., Rudower Chaussee 30, 12489, Berlin, Germany
[2] Gunnar Heilmann
gfai tech GmbH, Rudower Chaussee 30, 12489, Berlin, Germany
[3] Marcus Ohm
GFaI e.V., Rudower Chaussee 30, 12489, Berlin, Germany

ABSTRACT

Acoustic imaging methods for sound source localization use microphone arrays. The quality of the results is strongly influenced by an accurate determination of the positions of the array microphones. Assurance of the required precision solely by means of exact design and construction is expensive, and the determination of the geometric sensor coordinates with classical methods is a time-consuming task and it is often not applicable in the field.

The paper gives an estimation of the necessary precision of the positions and it proposes an algorithm which can determine unknown source positions as well as unknown sensor positions simultaneously. There is no need for reference speakers in this method, so it is useful for flexible array designs and their fast calibration in the field. Simulations and application examples are demonstrated, and the limits of the method will be discussed.

1 INTRODUCTION

All beamforming methods need to consider the delays of a sound event travelling to each microphone within the given array. A high spatial resolution of the acoustic map strongly depends on a highly accurate position of each microphone relative to its neighbour. That is needed to identify the location of sound sources exactly. The precise determination of the microphones' coordinates within a 3D-array using common methods (like optical 3D-scanner) poses serious challenges. This also applies for large array constructions, which are used to find sources with mainly low frequency sound characteristics. The different array shapes depend on the respective application, some can be disassembled for the reasons of transport issues. When setting up these kinds of arrays several times, the physical microphone's position is subject to change from the theoretically assumed position. A tremendous effort is needed in order to comply with needed static stability of the array construction tolerances. Therefore the microphones' positions are determined once during production and calibration. The repeat accuracy of the microphone position offset in foldable arrays is still unacceptable. In fact it requires a redeterminiation of the microphones' positions before each measurement session. With the given situation this seems not practicable.

2 REQUIRED ACCURACY

Before developing a method which obtains the 3D microphones' positions automatically, it is necessary to estimate the desired accuracy considering the respective source misplacement. Initially several measurements were simulated to consider different microphone position offsets (Fig1 and Fig2). To determine the quality of the array, the actual maximum available map contrast (difference from the main lobe to the first side lobe) was related to the given map contrast after introduction of the microphone position offset. This was performed for white noise and several different selected frequencies.

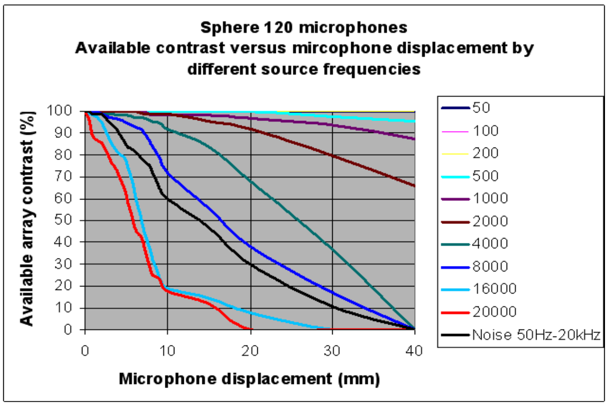


Fig. 1

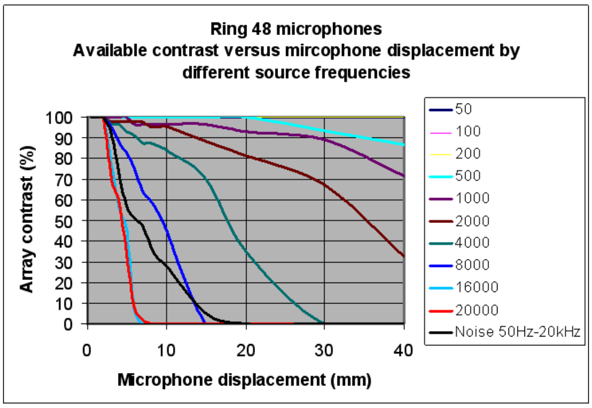


Fig. 2

Already a given microphone position offset of 2mm leads to a reduction of the availably map dynamic of 20% in the higher frequency range. Therefore the aim of the development of such an auto calibration method should be to reach a positioning accuracy less than 2mm.

3 STATE OF THE ART TECHNOLOGY

This issue has been investigated by several authors. [1] points to a method that shows how to obtain the microphone coordinates. Firstly, a tape measure is used to get the relative distances between the microphones. Secondly, the “Multidimensional scaling" method is applied. Reference [2] describes a simplified method which requires less known microphone distances. Considering a given number of base points consisting of a fixed setup of loudspeakers, the distances of the microphones can be calculated by using the delays of the sounds between these speakers and the microphones. Paper [3] describes a method using a speaker which gets physically attached to 5 array microphones in a sequence. The microphone positions are found by processing the delays between the previously synchronised sound impulses from the speaker to the microphones. All the methods have one or more disadvantages:

- The measurement of all microphone distances within one array is very difficult and may cause errors. Considering a 48 channel array, one would need to measure 1128 different microphone distances.
- Creating an additional loudspeaker array with a signal generator presents extra hardware effort to cover. For large arrays such loudspeaker array has to be designed properly to avoid inefficient clipping conditions.
- Attaching loudspeakers to several array microphones also seems not very useful and it is complicated.
- The achievable accuracy reaches +/- 10mm. This is not satisfactory for higher frequencies.

4 THE IDEA

The previously mentioned methods are based on the determination of the distances of the sound sources to the array microphones by processing the delays from the sound impulses considering the speed of sound. The applied algorithm differs from the previous ones in the following manner:

- The system consists of a number of points with unknown coordinates and a number of known distances between them. Hereby the system does not differentiate between microphones and sound sources.
- The number of sources is dramatically increased to 50-100 so the system becomes over-determined. By using so many sources a perfect averaging compensates for disturbed sources, which can eliminate these points from the system.
- In order to reach such a large amount of sound sources, one will use a repetitive sound source moving around the array.
- A constraint dependent numerical optimisation algorithm will be used to obtain the coordinates. The starting parameters of the source coordinates are always random (within reasonable limits) whereas the coordinates of the microphones should be an approximation of the real positions.

5 SOUND SOURCES

The sound source should emit a signal form which allows the calculation of the delay between the source and the microphone. The sound source's signal must propagate in spherical manner in all directions (at least towards the array). To do this procedure in practice one needs to add an additional reference channel, which is available on most of the common beamforming systems. This reference channel will be connected with or without known runtime difference to the sound source. For instance a microphone near the sources can be fed to this channel and serves as origin source time reference.

In practice two different signal types proved to be useful for this procedure:

1. A chirp-signal. The chirp offers the advantage of being resistant to noise combined with a good measurability of the delays using cross correlation. When the cross correlation is done in the frequency domain, one can reach a good interpolation. A disadvantage is the need for an additional sound generator and a speaker. The latter in turn will never be an ideal point like sound source. Furthermore, they cannot move within one chirp.
2. A Dirac impulse. This Dirac impulse can be created simply by creating a spark using two electrodes. The resulting spark creates the almost perfect spherical wave form when the electrodes are not too distant from each other. The only disadvantage of this set up is the sensitivity to louder background noise and to reflections.

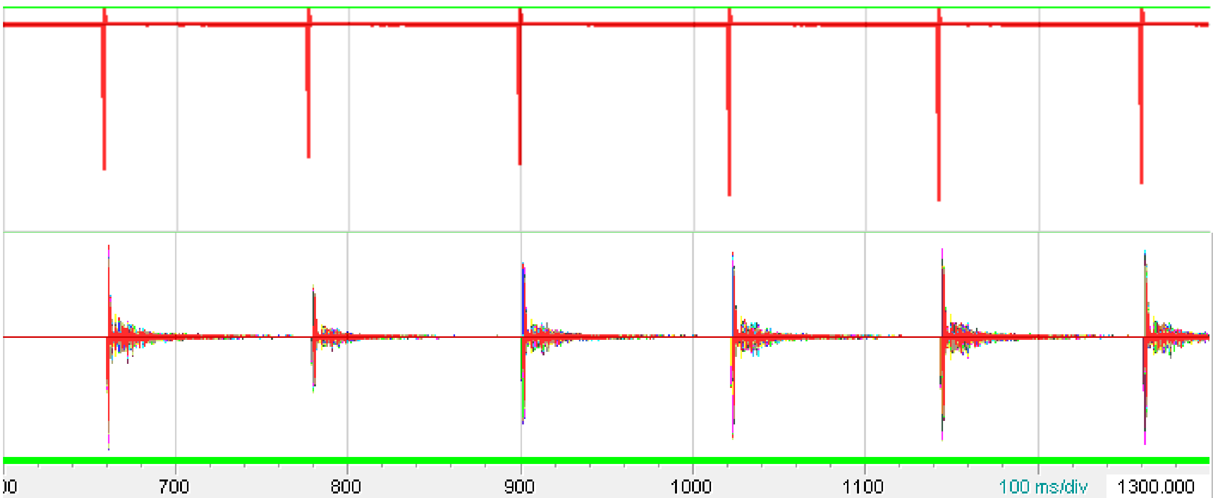


Fig. 3 reference channel with spark impulses (top) and microphone channels with received sound impulses (bottom)

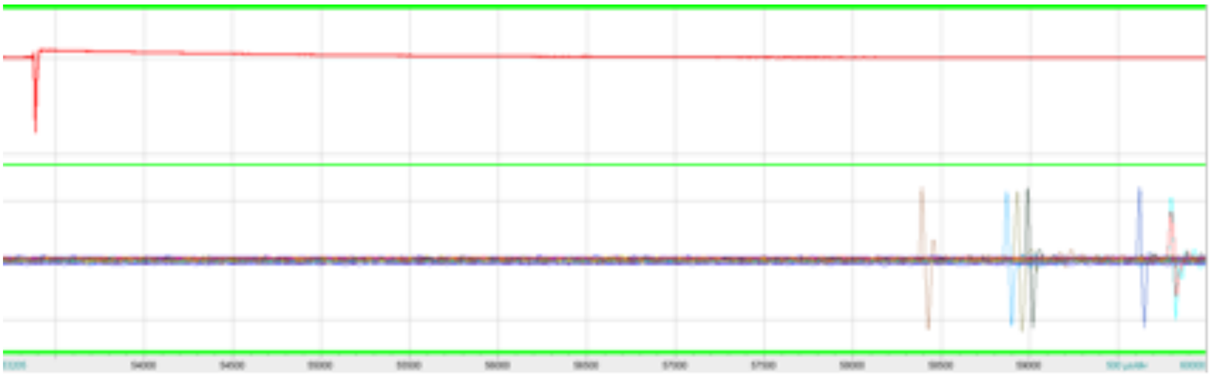


Fig. 4 magnification of Fig. 3 – one spark impulse (top) and the delay of the received sound waves (bottom)

6 ALGORITHM

The result of the measurement and the determination of the delays between the sources and every microphone is a vast number of distances. For example, the total mesh of a sphere array with 120 channels using 100 sources reaches 12.000 different paths. The relation between source coordinates S1(x,y,z), a microphone M1(x,y,z) and the distance S1M1 is due to the Euclidean Distance between two points:

$$D_{S1M1} = \sqrt{\left(\left(x_{S1} - x_{M1}\right)^2 + \left(y_{S1} - y_{M1}\right)^2 + \left(z_{S1} - z_{M1}\right)^2\right)} \qquad \text{Equation 1}$$

Because of the example mentioned above there are 12.000 equations for the system. As this is a matter of a non-linear optimization problem, it is possible to use the known non-linear optimization procedures. But these procedures do only converge with sufficient starting conditions. That is why a simple numeric search algorithm will be used. The convergence criterion is the minimization of the sum of the error square.

1. Initialization of the source coordinates with random numbers. Initialization of the microphones' coordinates with approximated values or random numbers. Consideration of boundary conditions due to the array geometry (e.g. two-dimensional array: all microphones are nearly arranged in-plane and all sources are in front of this plane).
2. Shifting of the first source with step size ss1 in x-, y-, z-direction by considering the boundary conditions until the sum of the error squares reaches its minimum
3. Repeat step 2 for all sources
4. Apply steps 2 and 3 to all microphones
5. Repeat steps 2-4 until the optimization of the errors squares ends
6. Reduce the step size ss1
7. Repeat steps 2 - 6 until termination condition has been reached
8. Verify the convergence criterion
9. If step 8 fails, change starting conditions and repeat the procedure from step 1

A smart selection of starting conditions leads to the fact that the algorithm always converges with the first try by using a randomized distribution of sources and good approximated microphone coordinates (± 0.1 meter). By using a randomized distribution of microphones and a disadvantageous decision of the starting conditions it could happen that the algorithm will not converge. This case can be easily identified through the stagnancy of the sum of error squares at a high level. A restart with new starting conditions will succeed. The outcome of random start positions of the microphones is a random spatial position of the microphone array. The coordinates have to be normalized with help of a coordinate transformation.

7 ANNOYING INFLUENCES

Air temperature

The air temperature has a wide influence on the results accuracy.

$$c_{\mathrm{Luft}} \approx 331{,}5 \frac{\mathrm{m}}{\mathrm{s}} \sqrt{1 + \frac{\vartheta/^{\circ}\mathrm{C}}{273{,}15}}$$

Equation 2

According to equation 2 a change of air temperature of ±1°C with a distance of 1 meter will lead to a change in length of ±1.7 mm. This fact has to be considered for the required accuracy. One solution could be the confirmation of the ambient temperature during the calibration process. But that requires an additional effort of hardware (e.g. a highly precise thermometer ±0.1°C). An easier way is the introduction of fixed distances between two or more microphones into the algorithm. These fixed distances are known for the most arrays and they will not change while building up the array. In all other cases the distances can be determined. By comparing the fixed distances with the distances between the other microphones the air temperature can be used as a free parameter in the algorithm. The algorithm will calculate the temperature itself.

Airflow

Because of the fact that sound propagates through the air, the absolute speed of propagation of sound compared to a rested object can be calculated with equation 2 and the air velocity. An air speed of 3.4 m/s (wind speed 3) during the calibration process causes an error of 1% (worst case), that means 10 mm on a distance of 1 m. Because of that, the calibration process can only be accomplished under conditions of a very low air speed. With the help of chirp-signals and the Doppler effect it is possible to compensate this annoying influence.

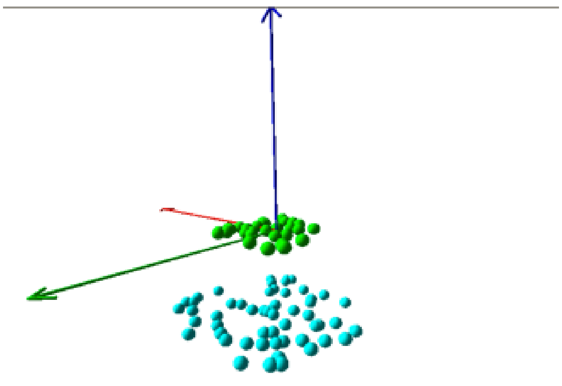
8 EXAMPLES

The signal processing as well as the algorithm will get realized in software. This software tool allows loading and processing the recorded data of microphone signals and reference signals. Furthermore it enables the user to specify the boundary conditions. The results will be graphically presented in 3D. The following 3 examples, which will demonstrate the suitability in practice, will be supported by these screenshots.

In the first example a spiral array (48 microphones, diameter approx. 6 m) mounted on the ceiling is used to localize a speaker. The start coordinates for the microphones are random values. To calibrate the array, a spark generator (a kind of taser) was moved through the room on an elliptical path. It generated a number of 56 sparks within a time of 5 seconds, so there were 56 sound sources.

Fig. 5 spiral array (48 microphones, diameter approx. 6 m)

After the calibration some distances between different microphones were determined and compared to the calculated positions. The maximum deviation was 1.8 mm.



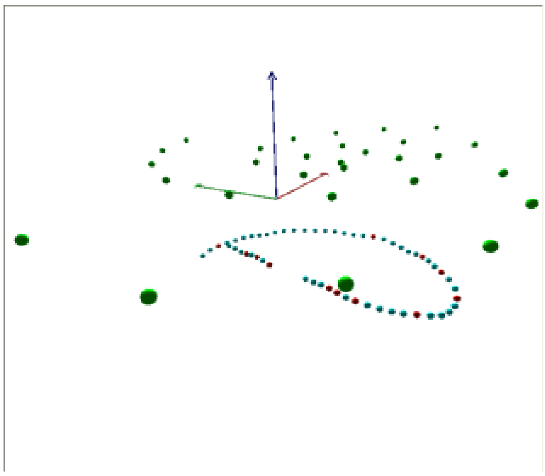


Fig. 6 start configuration and the result after calculation green: microphones blue: sources red: errors

The following example demonstrates the re-calibration of a damaged star array (48 microphones, dimensions ca. 3,8 m x 3,8 m x 2 m). The arms of the array were out of shape after the array toppled over. Before the calibration the array was not applicable. The spark generator was moved on a randomized path and generated 112 sources within a time of 16 seconds.

After the calibration procedure the array is as well applicable as before the damage. The maximum positioning error of the microphones is 1.4 mm.

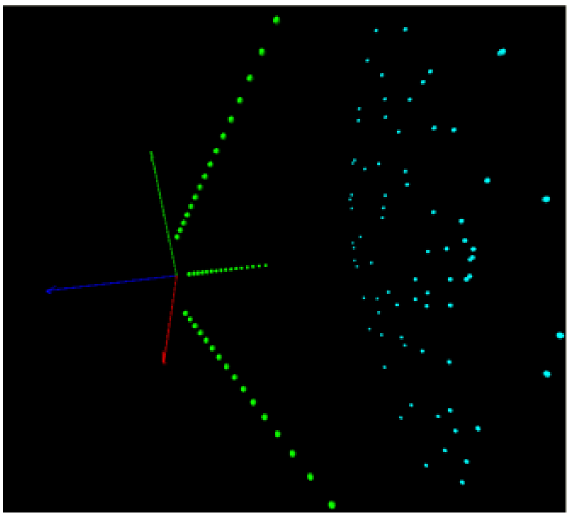


Fig. 7 start configuration for re-calibrating of the star array. Well approximated start coordinates (green) and randomized distribution of sources (blue)

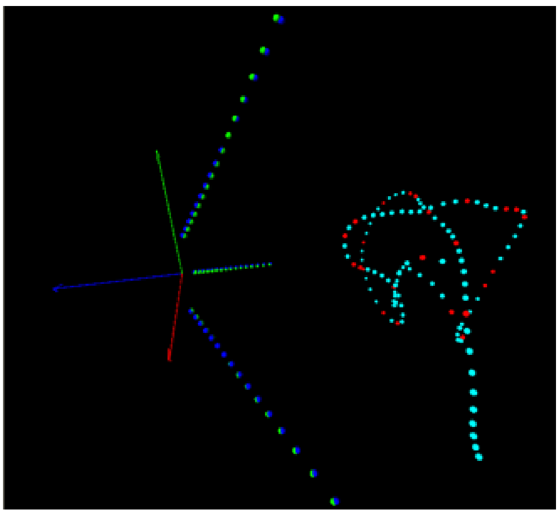


Fig. 8 result after the calibration process.
dark blue: the original microphone coordinates before the damage
green: determined microphone coordinates after calibration
light blue end red: sound sources and errors

The third example shows the determination of the microphone coordinates for a sphere array with 120 channels. From random start coordinates of the microphones as well as the sound sources the algorithm calculates the correct positions. In the case of transparent 3D arrays the sound field will be influenced by diffraction and reflection when entering the array. Furthermore, the acoustic irradiation of the microphones from the backside will cause delays, so that the calculated distances get too long. The maximum difference between the determined and the real microphone coordinates is less than 4 mm.

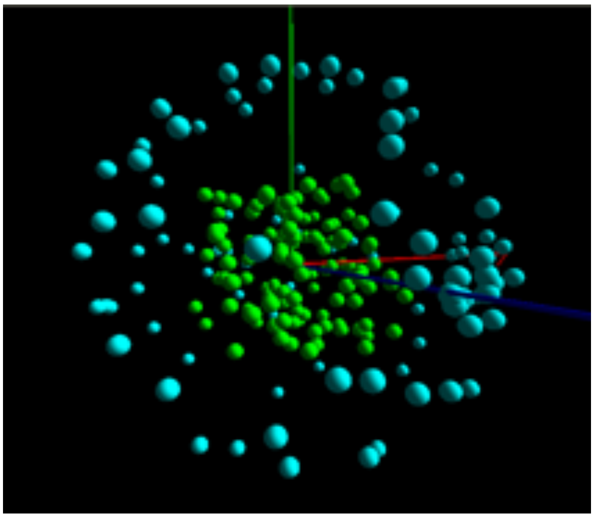


Fig. 9 start configuration for the 120 channel sphere array
random distribution of microphon coordinates and 111 sources

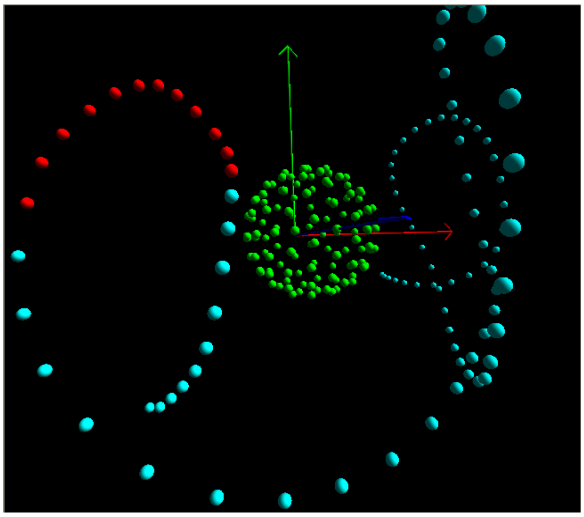


Fig. 10 result after the calibration

9 CONCLUSIONS

With the help of the presented algorithm and a software tool it is possible to determine positions of microphones in 2D and 3D arrays in short time. The advantages of this algorithm over the conventional one are:

- All coordinates of the sources and microphones can be unknown.
- Additional hardware is not necessary, thus loudspeakers, loudspeakers' arrays or other sound sources which have to be connected with the microphone array, are not needed. Thus a signal generator is not necessary. The calibration occurs without contact. The additional hardware consists of a standard taser (costs less than 100 Euro).
- It can be used quickly and easily. The calibration consists of the measurement (16 seconds) and the analysis (about 1 minute). Therefore, also arrays that have to be put together or that can be folded can be even calibrated in the field where the measurement takes place.
- The highly overdetermined equation system allows the identification of incorrect sources (caused by shadowing, airflow and so on), a good mediation and thus a good

error compensation as well as a testimony to the quality of calibration (residual errors, standard deviation)
- The high precision of positions' determination. The deviation in 2D arrays is less than 2 mm, in 3D arrays less then 4 mm

Planned improvements are:

- Use of the chirp signal to improve the interference resistance during the calibration process.
- Determination of the airflow during the calibration process.
- Consideration of the microphones' locus (direction dependent signal delay) in the algorithm.
- Optimizing the algorithm itself (determination of good approximations proceed with the presented method, after that conventional methods, e.g. Gauss-Newton, Levenberg-Marquardt, etc.)

REFERENCES

[1] Stanley T. Birchfield: Geometric Microphone Array Calibration by Multidimensional Scaling,
IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2003, Hong Kong.

[2] Stanley T. Birchfield, Amarnag Subramanya: Microphone Array Position Calibration by Basis-Point Classical Multidimensional Scaling,
IEEE Transactions on Speech and Audio Processing, 2004.

[3] Vikas C. Raykar and Ramani Duraiswami: Automatic Position Calibration of Multiple Microphones,
Perceptual Interfaces and Realities Lab., University of Maryland, CollegePark