inter.noise 2010
noise and sustainability
13-16 June 2010 . Lisbon . Portugal
www.internoise2010.org

STUDY OF AN OCTAHEDRAL ANTENNA FOR BOTH SOUND PRESSURE LEVEL ESTIMATION AND 3D LOCALIZATION OF MULTIPLE SOURCES

Patrick Marmaroli; Xavier Falourd; Hervé Lissek

EPFL Acoustic and Electromagnetic Laboratory (LEMA)
patrick.marmaroli@epfl.ch

Abstract

This study deals with localization and characterization of multiple sources by mean of an octahedral array processing. We show how correlations enable to both localize more sources that the number of sensors and estimate their sound level pressure simultaneously. The theoretical basis of the algorithm, based on analytical geometry and time differences of arrival, is explained for the broad-band sources in the far field and experimented in anechoic conditions. The localization error due to uncertainty of some physical measurements and numerical recording is studied. Results and limitations of the proposed method are discussed.

Keywords: 3D Microphone Array, Acoustic Source Localization, Time Difference Of Arrival, Generalized Cross-Correlation Methods, Uncertainty of Measurements

1 Introduction

The so-called passive sound sources localization using time differences of arrival (TDOA) estimation techniques is very popular in acoustic. But to well describe a sonorous environment, a simple location of noise sources is often not sufficient and it could be necessary to classify the competitive sources involved according to predefined criteria. In this paper we present our first results in terms of location and energy ranking of broadband, stationary, uncorrelated and spatially separated noises. Each source is considered as a monopole located in the far field. This paper is divided as follows: in Section 2 we explain the method to identify the directions of arrival of multiple broadband sources simultaneously. A study on the direction of arrival estimation accuracy relative to numerical noise and physical measures is done in Section 3. We introduce in Section 4 how the cross-correlations functions could be used for estimate the sound pressure level of one or more sources. Finally an experiment in an anechoic chamber is described in Section 5 and results and limitations of the proposed method are described in Section 6.

2 Localization algorithm for one or more broadband sources

2.1 The classical cross-correlation

The popular used method for locate sound source in space consist in estimate, in a first step, the time delay between pairs of sensors of a array with known geometry, and in a second step, to estimate position of source according these obtained values. Consider a broadband signal *s(t)* recorded by two microphones p_i and p_j, we can modelling each recorded signal as :

$$\begin{aligned} p_i(t) &= s(t) + n_i(t) \\ p_j(t) &= \alpha_{ij} s(t-\tau_{ij}) + n_j(t) \end{aligned} \tag{1}$$

where n represents a noise measurement, the amplitude ratio α_{ij} is symbolizing the attenuation of sound and τ_{ij} the time propagation of the wave field between the both sensors. A quick method for estimating τ_{ij} consists in compute the cross-correlation [2] expressed by:

$$R_{ij}(\tau) = \int_{-\infty}^{+\infty} p_i(t) p_j(t-\tau) dt \tag{2}$$

The cross-correlation function admits a maximum value related to the delay between both signals such as:

$$\tau_{ij} = \arg\max_{\tau} R_{ij}(\tau) \tag{3}$$

2.2 Case of one source in the far field

Let *R(O,x,y,z)* be an orthonormal basis (figure 1a) in which six microphones form an octahedral array of diameter *d* centered in (0,0,0) (figure 1b). We will consider the three pairs related to the microphones {1,2}, {3,4} and {5,6}. Let *S* be a sonorous source located at coordinates (X,Y,Z) in the antenna basis. Times of arrival differences between each pairs enable to estimate, in the near field, the Cartesian position of *S* (hyperbolic approach) and in the far field, its direction of arrival in azimuth and elevation (hyperbolic or conic approach).

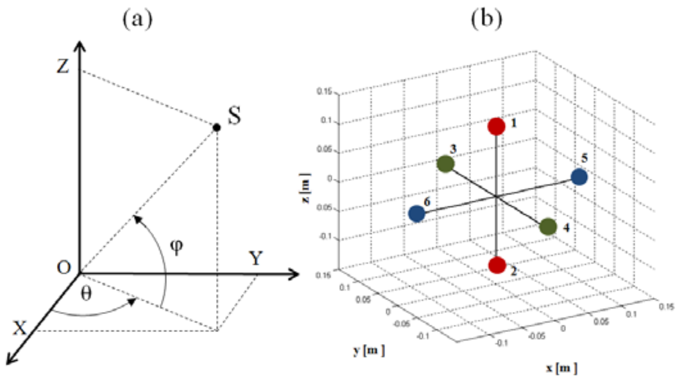

Figure 1

In a first time, we consider the problem of one sound source localization. Without a priori on its position, possible solutions are supported by a sphere of radius R large enough to respect the far field hypothesis such as:

$$x^2 + y^2 + z^2 = R^2 \tag{4}$$

For each pair {i,j}, the possible solutions with respect to a delay τ_{ij} is a circular conical surface vertex O and aperture 2 α_{ij} such as:

$$\alpha_{ij} = Arc\cos\left(c\tau_{ij} / d\right) \tag{5}$$

Where c is the speed of sound. The equation of the three cones supported by the three pairs is:

$$\begin{aligned} x^2 + y^2 &= z^2 \tan^2 \alpha_{12} \\ z^2 + x^2 &= y^2 \tan^2 \alpha_{34} \\ y^2 + z^2 &= x^2 \tan^2 \alpha_{56} \end{aligned} \tag{6}$$

Coordinates of the point which satisfy the equations (4) and (6) are:

$$\begin{aligned} X &= R\cos\alpha_{56} \\ Y &= R\cos\alpha_{34} \\ Z &= R\cos\alpha_{12} \end{aligned} \tag{7}$$

If R is unknown, only the source direction of arrival (azimuth θ and elevation φ) which is independent of R is known:

$$\varphi = Arc\tan\left(\frac{\cos\alpha_{12}}{\sqrt{\cos^2\alpha_{34}+\cos^2\alpha_{56}}}\right)$$
$$\theta = Arc\tan\left(\frac{\cos\alpha_{34}}{\cos\alpha_{56}}\right) \tag{8}$$

2.3 Case of multiple sources in the far field

Theoretically, when N uncorrelated and broadband sources are present, cross-correlations computed on each pair of sensors give a number of detected peaks less or equal to N. In the noiseless and non reverberant case, figure 2 gives a schematic example of what could give the cross-correlations in presence of three sources. The three cross-correlations contain each three peaks. The problem consist in associate to each peak those of the two other pairs which give a physical coherent solution for a source direction of arrival.

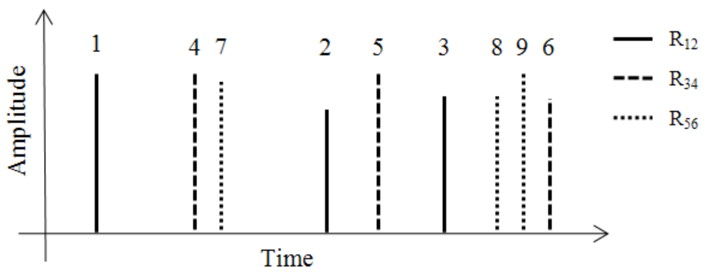

Figure 2

In the example of the figure 2, several combinations are possible for the peaks associations, for example : {1,4,7}, {1,4,6}, {1,4,9}, {2,4,7}, {2,4,8} etc… We proposed a combinatory treatment to select the only physical coherent solutions. By replacing (7) in (4), we obtain a relation which links all aperture angles, independently of the source position:

$$\cos^2\alpha_{12} + \cos^2\alpha_{34} + \cos^2\alpha_{56} = 1 \tag{9}$$

In term of inter-sensor delays, relation (9) is equivalent as:

$$\sqrt{\tau_{12}^2 + \tau_{34}^2 + \tau_{56}^2} = \frac{d}{c} \tag{10}$$

In other words, if a triplet of delays does not verify this equality, then its association does not give a coherent source position. This criteria permit to select only the combinations which verify this equality and then localize each source. In practice, knowing the number K of sources to locate, we chose the K triplets of delays which maximize the equality (10). Thus, the number of sources could be higher as the number of sensors: the figure 3 shows a simulation where seven sources are simultaneously localized.

Notice that equation (9) is independent of the source direction of arrival. This property implies that an angle can be found knowing the two others. For example, elevation estimation is possible even if we dispose of a planar array (semi-tridimensional localization).

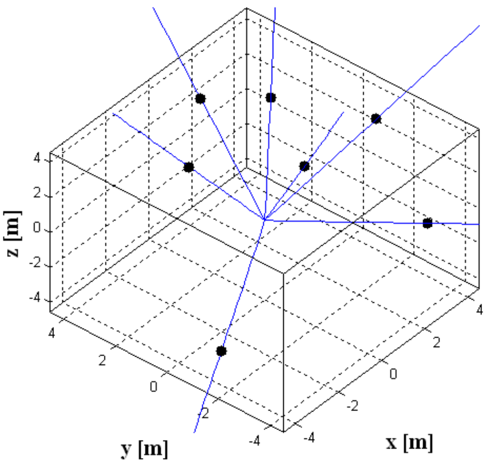

Figure 3

3 Angular accuracy

In this section we focus on the influence of each parameter on the localization accuracy for a two microphones array, the simplest case.
According to the equation (5), the direction of arrival estimation is dependent on physical criteria like the distance between microphones, sound speed and on one digital parameter i.e. the frequency sampling for the delay estimation. Let d be the distance between both (different of the measured distance $\hat{d}$), T_e the temperature different of the measured temperature $\hat{T}_e$ and f_s the known and real sampling rate.
The linear array is oriented in such a way that a source comes from the direction of arrival 0°, i.e. the end-fire scenario.

3.1 Angular accuracy due to an inter-sensor spacing error

The real and measured inter-sensor spacing between both microphones d are linked according to the relation:

$$d_{\inf} \leq d \leq d_{\sup} \tag{11}$$

Where $d_{\inf} = \hat{d}(1-\delta)$ and $d_{\sup} = \hat{d}(1+\delta)$ and δ is a tolerance of measurement expressed in %. We deduce from (5) and (11) the following relation between the measured and real directions of arrival.

$$\alpha_d^{\inf} \le \alpha \le \alpha_d^{\sup} \quad \text{for} \quad 0 \le \alpha \le \frac{\pi}{2} \tag{12}$$

Where $\alpha_d^{\inf}=Arc\cos\left(c\tau / d_{\inf}\right)$ and $\alpha_d^{\sup}=Arc\cos\left(c\tau / d_{\sup}\right)$. Let $\varepsilon_{\alpha,d}$ be the maximal possible angular error we can have with an inter-sensor spacing measurement, we deduced from (12):

$$\varepsilon_{\alpha,d} = \alpha_d^{\sup} - \alpha_d^{\inf} \quad \text{for} \quad 0 \le \alpha \le \frac{\pi}{2} \tag{13}$$

The accuracy for considered antenna and for different tolerances is represented on Figure 4 a.

3.2 Angular accuracy due to an error of temperature measurement

The relation between speed of sound c and temperature T_e (in Celsius degree) is [4]:

$$c = 331.4 + 0.607 T_e \tag{14}$$

If an error of γ degree is done between T_e and $\hat{T}_e$, the measured celerity $\hat{c}$ and real celerity c are linked according to the relation:

$$c_{\inf} \le c \le c_{\sup} \tag{15}$$

Where $c_{\inf} = 331.4 + 0.607(\hat{T}_e - \gamma)$ and $c_{\sup} = 331.4 + 0.607(\hat{T}_e + \gamma)$. We deduce from (5) and (15) that:

$$\alpha_c^{\inf} \le \alpha \le \alpha_c^{\sup} \quad \text{for} \quad 0 \le \alpha \le \frac{\pi}{2} \tag{16}$$

Where $\alpha_c^{\inf}=Arc\cos\left(c_{\sup}\tau / d\right)$ and $\alpha_c^{\sup}=Arc\cos\left(c_{\inf}\tau / d\right)$. Let $\varepsilon_{\alpha,c}$ be the maximal possible angle error we can have with a temperature error measurement, we deduce from (16) :

$$\varepsilon_{\alpha,c} = \alpha_c^{\sup} - \alpha_c^{\inf} \quad \text{for} \quad 0 \le \alpha \le \frac{\pi}{2} \tag{17}$$

The accuracy for considered antenna and for different errors of measurement is represented on Figure 4 b.

3.3 Angular accuracy due to the sampling frequency

By definition, we can't distinguish two peaks if the gap between both is less than two samples. Then a measured delay $\hat{\tau}$ is linked to the real delay τ by the relation:

$$\tau_{\inf} \leq \tau \leq \tau_{\sup} \qquad (18)$$

Where $\tau_{\inf} = \max(\hat{\tau} - 2/f_s, -\tau_{\max})$ and $\tau_{\sup} = \min(\hat{\tau} + 2/f_s, \tau_{\max})$. Indeed, the absolute value of a delay can't be upper the time for sound for propagating between both microphones, this limit is $\tau_{\max} = d_{\sup}/c_{\inf}$. We deduce from (5) and (18):

$$\alpha_{f_s}^{\inf} \leq \alpha \leq \alpha_{f_s}^{\sup} \quad \text{for } 0 \leq \alpha \leq \frac{\pi}{2} \qquad (19)$$

Where $\alpha_c^{\inf} = Arc\cos\left(\tau_{\sup}/d\right)$ and $\alpha_c^{\sup} = Arc\cos\left(\tau_{\inf}/d\right)$. Let ε_{α,f_s} be the maximal possible angular error we can have with a given sampling frequency, we deduce from (19) :

$$\varepsilon_{\alpha,f_s} = \alpha_{f_s}^{\sup} - \alpha_{f_s}^{\inf} \quad \text{for } 0 \leq \alpha \leq \frac{\pi}{2} \qquad (20)$$

The accuracy for considered antenna and for different sampling rates is represented on Figure 4 c.

Figures below illustrate by comparison that sampling rate is the preponderant parameter for the angular accuracy. Indeed the tools for measuring distance and temperature are relatively precise nowadays, errors from these parameters are well known. On the other hand, high sampling rate is not easy to obtain in the case of embedded real-time system; this last parameter is the more penalizing considering the proposed algorithm.

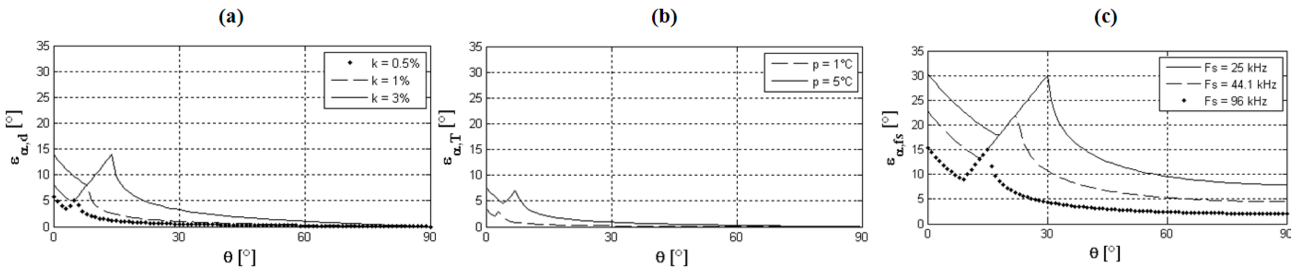

Figure 4

We notice that for all three studied parameters, errors are higher in the end-fire than in the broadside case. That's why a cross geometry in the plane, and by extension, a cubic geometry in 3D, provides the most robust compensation with an equal number of microphone. In the Figure 5, the obtained angular accuracy is showed by taking the same situation of the Figure 3; this gives an overview of possible problem to separate and to locate two very close sources.

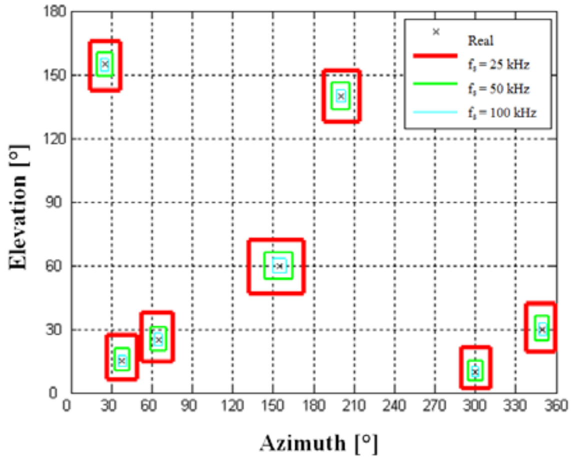

Figure 5

4 Sound pressure level estimation

For all detected and localized sources we search now to sort them according to its noise level. The proposed method for this involves analysis of the maxima of the cross-correlation function. Indeed, this value, noted $R_{ij}^m = \max(R_{ij})$ is related to the sonorous level of the source by the relation [1]:

$$L_{dB} = 20\log\frac{\sqrt{R_{ij}^m}}{p_0} \tag{21}$$

Where p_0 is the reference sound level pressure equal to 20 µPa. For a given delay τ_0, the real value of $R_{ij}(\tau_0)$ is within a confidence interval of 95% such as $R_{ij}(\tau_0) = \hat{R}_{ij}(\tau_0) \pm 2\sigma_{ij}(\tau_0, T)$ [1] where:

$$\sigma_{ij}^2(\tau_0, T) = \frac{2\beta - 1 + 2e^{2\beta} + [(2\alpha + 1)(2\beta - 1) - 2\alpha^2]e^{-2\alpha}}{2\beta^2} \text{ for all } T > \tau_0 \geq 0 \tag{22}$$

Where $\alpha = \pi B\tau_0$, $\beta = \pi BT$, B is the bandwidth of the signal and T is the time of observation. Some simulations to measure the absolute error in dB between the real and the estimated value in function of the time of observation are investigated. For a fixed bandwidth of 5 kHz, and in function of the bandwidth for a fixed time of observation of 100 ms, results are represented on the Figure 6. As predicted by the theory, a good estimation of sound pressure level needs both high B and T , moreover error is dependent on the real value, so the louder the source the lower the error.

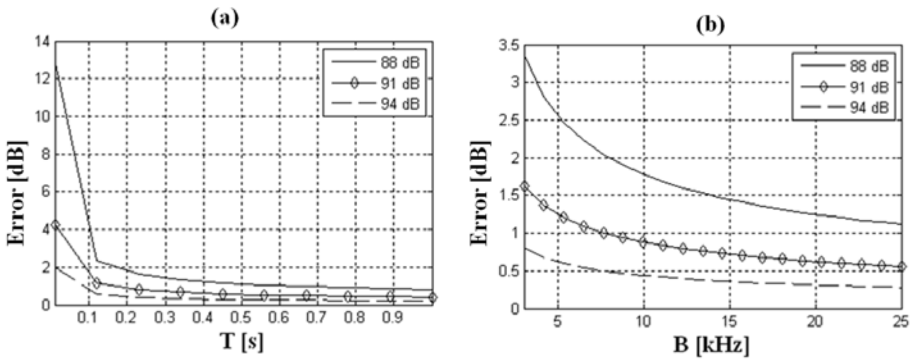

Figure 6

5 Anechoic room experiment

An octahedral antenna, with an inter-sensor spacing of 18 cm, is placed in front of two loudspeakers with angles $\{\theta,\varphi\}_1$ = {25°,20°} and $\{\theta,\varphi\}_2$ = {7°,30°}. Sound sources are centered and normally distributed white noise with 25 kHz bandwidth, the sampling rate of the record is 50 kHz and the time of observation is 5 sec.
Experiments are done in anechoic conditions. The both emitted white noises are uncorrelated and their sonorous levels are fixed at 85 dB.

In first step each loudspeaker is switched on alternatively to detect their contribution through the peak obtained by cross-correlation between signals recorded on two microphones composing a pair {i,j}: Figure 7(a, b). The position of the sound source and its angular accuracy area are then computed: Figure 7(d,e). Concerning the sound pressure level, values of peak give a mean of 84 dB for both sources, so the link with cross-correlators is verified.

When both sources are switched on at the same time, we get the cross-correlations represented on Figure 7c. The peaks p_{11}, p_{12}, p_{13} and p_{21}, p_{22}, p_{23} are found at the same time in the mixture. For the pair {56}, the position of loudspeakers implies peaks in the same delay. That's why we can't distinguish two different peaks in the mixture. However we note that the energy of this peak has doubled as expected. In applying the physics criterion of the equation (10) we find the two physical coherent triplets and positions of sources are well estimated: Figure 7f. However, the resolution of correlations decreases as the number of sources increases. We can overcome this problem by applying filters on signals before computing the correlations [2], or use some "high resolution" methods [3], but the energetic peaks values risk to be lost. We currently work on how use these techniques to increase our resolution without loose the sound level pressure information.

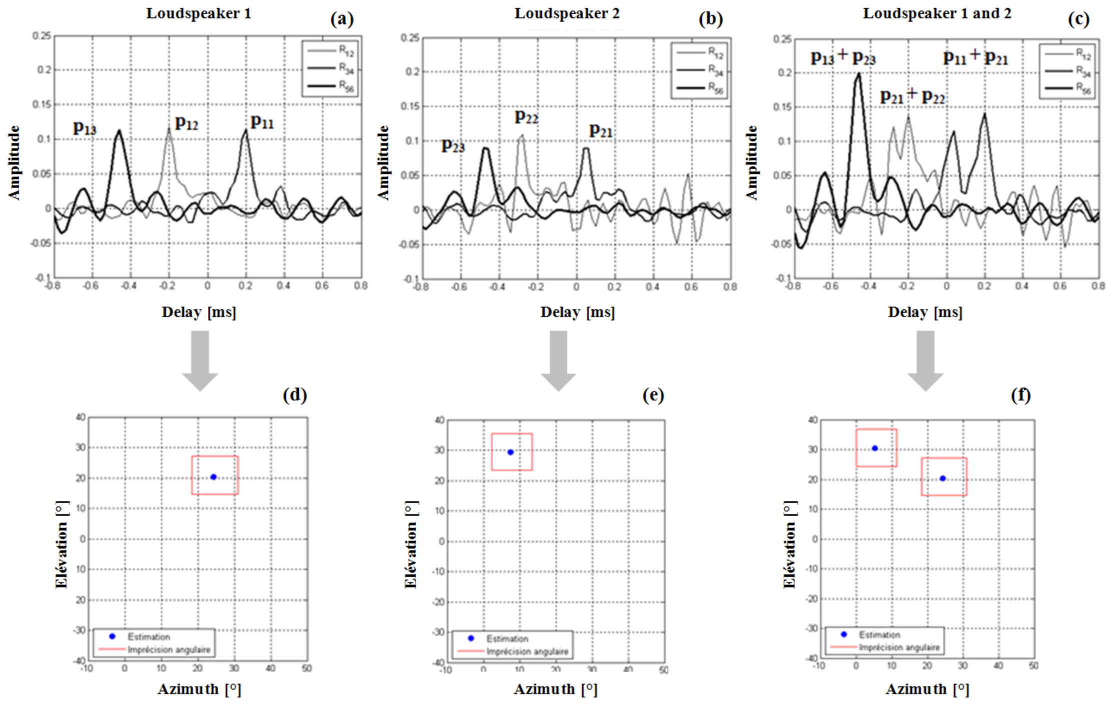

Figure 7

6 Conclusion

In this paper, methods and experiments are presented in order to localize and characterize simultaneously multiple sources by an octahedral array processing. The used geometry is justified by studying the angular accuracy due to physical measurements or set-up. It has been shown how classical cross-correlators could be used for both localize and estimate the sound pressure of sources, however, the best performances are obtained for very precise case and the forthcoming work will consist in working on our algorithm to wider applications.

References

[1] Bendat. Interpretation and application of statistical analysis for random physical phenomena. *I.R.E Transactions on Bio-Medical Electronics*, 9(1):31–43, January 1962.

[2] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 24(4):320–327, Aug 1976.

[3] Marie Agnes Pallas. *Identification active d'un canal de propagation à trajets multiples.* PhD thesis, INP Grenoble, 1988.

[4] Mario Rossi. *Audio*. Presses Polytechniques et Universitaires Romandes, 2007.