



The Sixteenth International Congress on Sound and Vibration

Kraków, 5-9 July 2009

MULTIPLE HARMONIC SOUND SOURCES SEPARATION IN THE UNDER-DETERMINED CASE BASED ON THE MERGING OF GONIOMETRIC AND BEAMFORMING APPROACH

Patrick Marmaroli, Xavier Falourd, Hervé Lissek

EPFL - LEMA, Switzerland

email : {patrick.marmaroli, xavier.falourd, herve.lissek}@epfl.ch

Usually acoustic beamformer or directive microphones are used in order to focus on known potential locations emitting sounds so as to identify and extract their major noise components. However performances of these techniques depend on the number of microphones and the a priori knowledge of the source locations. Based on the measurement of phase delays between sensors and short-time Fourier transform of the sound field, we present an algorithm for localise and extract harmonic sound sources by merging goniometric and beamforming approach. The acoustic goniometer is composed of a small array of microphones and permits to detect and identify more sources than the number of available microphones. In order to evaluate the performance of this system, an experiment is performed with a planar antenna composed by four omnidirectional microphones laid out in front of eight loudspeakers located in the horizontal plane. We describe the experiment performed in an anechoic chamber.

1. Introduction

Sonorous source separation is a signal processing technique which permit to characterize n sonorous sources from m observations. Source separation is mainly used in noise source characterisation (environmental application), extraction of a signal of interest (vocal recognition application), identification of which source gives which sound (motor engine applications) etc. That's why numerous researches have been done this last three decades [HK96]. In this article, we will interest about source separation in the “under-determined” case with instantenous mixtures. A mixture is qualified as under-determined when the number of sources is more important than the number of sensors ($n > m$). The presented algorithm leans of very simple and well-known acoustic properties and turn out be very effective in harmonic and disjoint signals. Separation is done by two step : localisation of the sources by a goniometric approach and extraction of each signal by a adaptive beamforming approach.

1.1 Acoustic goniometry

Acoustic goniometer is here define as a system that measures the direction of arrival of a sound wave [Lan01] by means of an array of sensors and an algorithm. 2D principe of goniometry is represented fig 1 : consider that a ponctual source produce a plane wave from the direction of arrival

θ , the distance between two sensors d_{is} implies a time delay of wave propagation $\tau = \tau_2 - \tau_1$ which is related with the direction θ by the relation :

$$\tau = \frac{d_{is}}{c} \cdot \cos(\theta) \quad (1)$$

where c is the celerity of sound and τ is the time for sound to cover the distance d_{12} .

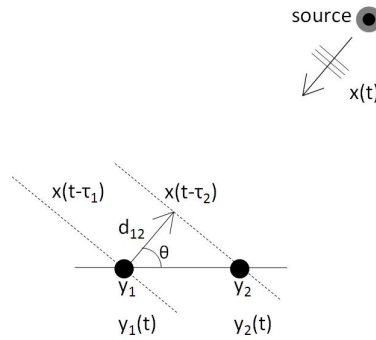


Figure 1. principle of 2D goniometer

1.2 Beamforming

Goal of the adaptive beamforming is to achieve maximum reception in a specified direction while signals of other direction are rejected. Practically, signals of sensors are delayed according to the desired angle of directivity. The beam pattern of an linear array which point at the angle θ_0 is given by : [O.N91]

$$|b(f, \theta)| = \left| \frac{\sin(\pi f m d_{is} (\sin\theta - \sin\theta_0)/c)}{\sin(\pi f d_{is} (\sin\theta - \sin\theta_0)/c)} \right| \quad (2)$$

where m is the number of sensor. Beam power patterns $|b(f, \theta)|^2$ of 4, 10 and 20-sensors array is represented fig. 2 for a frequency $f = c/(4d_{is})$.

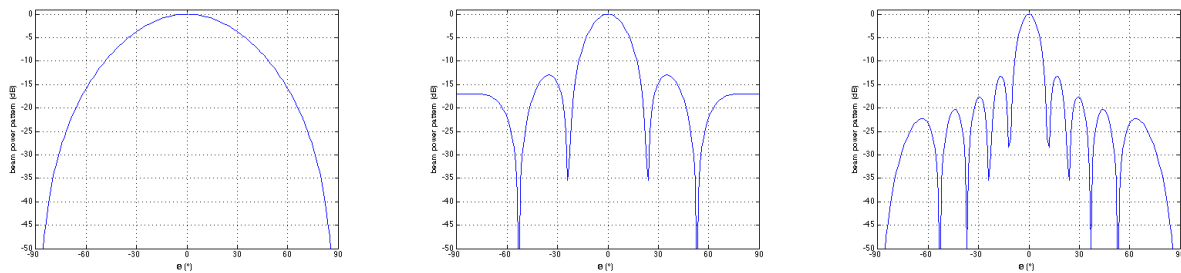


Figure 2. Normalised beam power pattern of a linear array consisting of 4 (left), 10 (middle), 20 (right) sensors without steering ($\theta_0 = 0$)

1.3 Theoretical background of source separation

Consider that we dispose of m observations $y_k(t)$ of the mixtures of n signals $x_i(t)$:

$$y_k(t) = \sum_{i=1}^n x_i(t - \delta_{k,i}), \text{ with } 1 \leq k \leq m \quad (3)$$

where $\delta_{k,i}$ is the time delay between the sensor 1 and the sensor k resulting from the direction of arrival of the signal i . In our application, the attenuation of amplitude is supposed to be negligible because of the little distance d_{is} between sensors. Classically, eq. 3 could be rewritten under a matricial form :

$$Y = AX \quad (4)$$

where Y contains the m observations : $Y \triangleq [y_1(t), \dots, y_m(t)]^T$ and X the n sources signals $X \triangleq [x_1(t), \dots, x_n(t)]^T$. A ($m \times n$) contains all terms of the mixtures. In undetermined-case, equation 4 could not be solved by a matricial inversion of A . So identify this matrix is not sufficient for reconstruct the sources. That's why it's generally assumed that signals had to be sparsely distributed in the time-frequency plane [AJ00], [OB99], [iA99], [RKO06]. We will study in this article the case of n harmonic and disjoint signals.

2. Localisation

2.1 Time delay estimation

Technique of delay estimation are numerous [Car92] and had to be adaptative according to the type of received signal type. In acoustic goniometry, we generally use classical or generalised intercorrelation methods (PHAT, CPSP, modified CPSP, ML...). These methods are as efficient as the ratio B/f_c is high, with B the broadband of the signal, and f_c its central frequency. In our case the n signals sent by loudspeakers are harmonic and disjoint :

$$x_k(t) = \sin(2.\pi.k'.f_0.t) \quad (5)$$

Where $k = 1, \dots, n$ and k' is a multiple of k . Thus the signals received at each sensor is the sum of the n signal but delayed in time because of the distance between sensors. A well-known method to estimate the delay of propagation $\delta_{k,i}$ is to use properties of the Fourier transform [AJ00]. Let call $Y_k(\omega)$ the Fourier transform of the signal $y_k(t)$ recorded at the sensor k :

$$Y_k(\omega) = \int_{-\infty}^{+\infty} y_k(t) e^{-i\omega t} dt \quad (6)$$

On the opposite sensor, we have the recorded signal $y_p(t) = y_k(t-\tau)$. Its Fourier transform is :

$$Y_p(\omega) = \int_{-\infty}^{+\infty} y_k(t-\tau) e^{-i\omega t} dt \quad (7)$$

In calling $u = t - \tau$:

$$Y_p(\omega) = \int_{-\infty}^{+\infty} y_k(u) e^{-i\omega(u+\tau)} du = e^{-i\omega\tau} \int_{-\infty}^{+\infty} y_k(u) e^{-i\omega u} du \quad (8)$$

Hence,

$$Y_p(\omega) = e^{-i\omega\tau} . Y_k(\omega) \quad (9)$$

Thus, for some ω , a time delay could be measured between two sensor p and k :

$$\tau_{pk}(\omega) = \text{Im}(\log(\frac{Y_p(\omega)}{Y_k(\omega)})) / \omega \quad (10)$$

Fig. 3 (left) represents the time delay model between the both pair of the set-up fig. 5 for the frequency of 80 Hz.

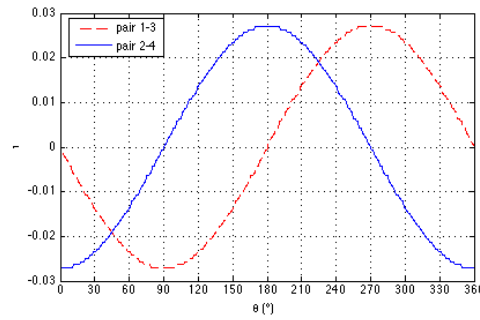


Figure 3. Theoretical time delay between the both pair of sensor in function of the angle θ for an harmonic signal of 80 Hz

2.2 Direction of arrival estimation

For the detected frequency $f = k \cdot f_0$ a signal of simulation $s'_k(t) = \sin(2 \cdot \pi \cdot f \cdot t)$ is generated and numerically walked around the array in order to generate a delay model per each direction of arrival $\tau_f(\theta)$ (figure 3). For each angle, the euclidian distance between the measurement and the theory is computed for both pair :

$$\Delta_{1,\theta} = \sqrt{\delta_1(\theta)^2 - \delta_1^2} \quad (11)$$

$$\Delta_{2,\theta} = \sqrt{\delta_2(\theta)^2 - \delta_2^2} \quad (12)$$

Thus, the association of both differences give a unique solution by minimisation of the criteria Φ :

$$\Phi(\theta) = \sqrt{\Delta_{2,\theta}^2 - \Delta_{1,\theta}^2} \quad (13)$$

Two model are generated for both pair and then compared to the measured phase difference of the both pair. The angle which presents a minimum of difference between the theory and the measurement.

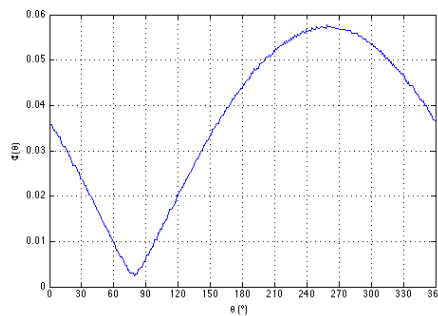


Figure 4. Criteria of localisation, presence of a unique minimum which corresponds to the direction of arrival of the harmonic signal

3. Extraction

Extraction is done by the technique of adaptive beamforming. For given direction θ , a signal of a frequency f obey to a phase difference law :

$$d\phi(f, \theta) = f \cdot \sin(\theta) \cdot \frac{d_{is}}{c} \quad (14)$$

The idea is to compare model and measurement for a given θ_0 , for both pair of the set-up 5

$$\Delta_{p,p+2}(f, \theta_0) = d'_{p,p+2}\phi(f, \theta_0) - d_{p,p+2}\phi(f, \theta_0) \quad (15)$$

where p takes value of 1 and 2 for both pair. A filter is then computed by taking the cosinus of this difference :

$$\begin{aligned} F_{p,p+2}(f, \theta_0) &= 1 \text{ if } \cos(\Delta_{p,p+2}(f, \theta_0)) \geq \text{threshold} \\ F_{p,p+2}(f, \theta_0) &= 0 \text{ if } \cos(\Delta_{p,p+2}(f, \theta_0)) < \text{threshold} \end{aligned} \quad (16)$$

Endly the extracted signal is the which contains only frequencies for which $F_{p,p+2}(f, \theta_0) = 1$.

4. Experiments

Two experiments have been done separately. One for test the localisation algorithm, the other for test the extraction algorithm. The set up is presented in gure 5. The distance between sensor is fixed at $d_{is} = 0.26$ m and the distance between loudspeakers is fixed at $d_{il} = 0.20$ m.

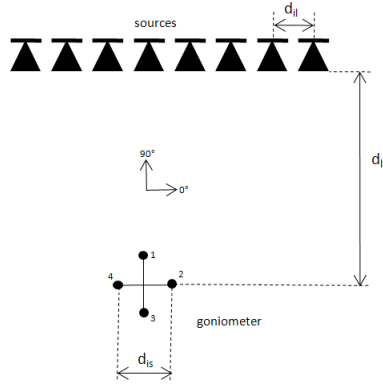


Figure 5. set up of the experiment

4.1 Localisation experiment

With no information about localisation of sources, Nyquist condition had to be respected in order to avoid ambiguities (*spatial aliasing*) [O.N91]

$$\frac{c}{2d_{is}} < f_{max} \quad (17)$$

where f_{max} is the maximal frequency to localize. Considering our array geometry, the frequencies of our eight harmonic sources go from 250 Hz to 600 Hz with a step of 50 Hz. The real and estimated angles are presented in the following table :

freq (Hz)	250	300	350	400	450	500	550	600
real angle	106	74	95	84	101	69	79	90
estimated angle	103	71	96	83	101	67	79	85

The mean error is 1.87 degree so this method of localisation seems to be robustness for harmonic signals.

- OB99 Jean-François Cardoso Olivier Bermond. Méthodes de séparation de sources dans le cas sous-déterminé. *Dix-septième colloque GRETSI*, 1999.
- O.N91 Richard O.Nielsen. *Sonar Signal Processing*. artech house, 1991.
- RK006 Lars Kai Hansen Rasmus Kongsgaard Olsson. Blind separation of more sources than sensors in convolutive mixtures. 2006.