TIME DOMAIN SOURCE LOCALIZATION TECHNIQUE BASED ON GENERALIZED CROSS CORRELATION AND GENERALIZED MEAN

Thomas Padois *[1], **Olivier Doutres**†[1], **Franck Sgard**‡[2], **and Alain Berry** §[3]

[1]Département d'ingénierie mécanique, École de Technologie Supérieure, Montréal, Québec, H3C 1K3, Canada

[2]Institut de Recherche Robert-Sauvé en Santé et en Sécurité du Travail, Montréal, Québec, H3A 3C2, Canada

[3]Groupe d'Acoustique de l'Université de Sherbrooke, Département d'ingénierie mécanique, Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1, Canada

1 Introduction

Microphone arrays are an efficient tool for localizing acoustic sources [1–3]. The acoustic field is recorded by a set of distributed microphones. Then the microphone signals are processed in order to build a noise source map which highlights the source positions. The standard technique is the Delay And Sum (DAS) beamforming which can be performed in the frequency domain via the cross-spectral matrix [1, 2] or in the time domain via the cross correlation of the microphone signals [3]. In the frequency domain, DAS beamforming is known to be less efficient at low frequencies, therefore several techniques based on inverse methods have been developed to improve its performances [1, 3]. In this study, DAS beamforming is implemented with the generalized cross correlation in the time domain. The arithmetic, geometric or harmonic means of the spatial likelihood functions are used similarly to Padois *et al.* [4]. The influence of the number of microphones is investigated with numerical data in the case of a standard circular microphone array.

2 Source localization method

Generalized cross correlation of the microphone signals

The cross correlation $\mathrm{R}_{x_m x_n}(\tau)$ is a useful function to estimate the time delay between two microphone signals x_m and x_n. Commonly in microphone array processing, the cross correlation function between two signals is obtained by the inverse Fast Fourier Transform of the cross-spectrum $C_{x_m x_n}$ which allows prefiltering operation

$$\mathrm{R}_{x_m x_n}(\tau) = \sum_{k=0}^{N_f-1} W(k) C_{x_m x_n}(k) \exp\left(j 2\pi \frac{k}{Nf}\tau\right), \quad (1)$$

where $W(k)$ is the prefilter and k is the frequency index. The technique used to prefilter is the PHAse Transform (PHAT) which allows removing the magnitude of the cross-spectrum in order to only retain the phase information. This technique is named the Generalized Cross-Correlation (GCC) or GCC-PHAT when the PHAT prefiltering operation is used (in the following the PHAT prefilter is used but the term PHAT is omitted for simplicity).

*Thomas.Padois@etsmtl.ca
†olivier.doutres@etsmtl.ca
‡franck.sgard@irsst.qc.ca
§alain.berry@usherbrooke.ca

Spatial Likelihood Function of a microphone pair

The peak value of the cross correlation provides the time delay between the microphone signals. This time delay corresponds to a set of potential source positions. If the cross correlation function is interpolated over spatial locations $\mathbf{r}_l$ where the source is searched, the resulting map is typically a hyperbola in two dimensions, this map is called Spatial Likelihood Function (SLF). However, the source position cannot be accurately detected with a single hyperbola (i.e a single microphone pair). If three microphones are used, three hyperbolas can be defined. If the SLFs are summed, the maximum value corresponds to the intersection of the hyperbolas and provides the source position. Commonly, the microphone array output signal $y^{AM}(\mathbf{r}_l)$ is obtained by computing the Arithmetic Mean (AM) of the SLFs determined for M_p microphone pairs as

$$y^{AM}(\mathbf{r}_l) = \frac{1}{M_p} \sum_{p=1}^{M_p} \mathrm{R}_p(\tau_{\mathbf{r}_l}), \quad (2)$$

where R_p is the cross correlation function of microphone pair p and $\tau_{\mathbf{r}_l} = \Delta t_{ml} - \Delta t_{nl}$ is the time lag for the spatial source location l.

Generalized mean of the SLFs

In the case of the GCC and M_p microphone pairs, the geometric mean (GM) and harmonic mean (HM) [4] are given by

$$y^{GM}(\mathbf{r}_l) = \left(\prod_{p=1}^{M_p} \left|\mathrm{R}_p(\tau_{\mathbf{r}_l})\right|\right)^{\frac{1}{M_p}}. \quad (3)$$

$$y^{HM}(\mathbf{r}_l) = M_p \left(\sum_{p=1}^{M_p} \frac{1}{\left|\mathrm{R}_p(\tau_{\mathbf{r}_l})\right|}\right)^{-1}. \quad (4)$$

where $|\cdot|$ denotes the absolute value. The influence of the number of microphones, therefore the number of microphone pairs, is investigated with numerical data in the following section.

3 Influence of the number of microphones on the noise source map

Numerical set-up: case of a standard circular microphone array in front of an acoustic point source

An acoustic point source, arbitrary located at $(x = -0.25, y = 0, z = 0.25)$ m, generates a white noise signal.

A standard circular array with a radius of 20 cm is set in front of the source 1 m away. The center of the circular array is located at $(x = 0, y = 1, z = 0)$ m. Therefore, the distance source-array is 1 m and the source is outside the inner circle of the array. This kind of array is commercially used and the geometry is convenient for investigating the decreasing of number of microphones. The time delay between each microphone pair signal is computed via Eq.1. The noise source maps are obtained by using the AM (Eq.2), GM (Eq.3) and HM (Eq.4). The scan zone is a square with side equal to 1 m and discretized with 2601 points (51×51). The only difference between the noise source maps is the combination of the SLFs.

Noise source maps: case of a 20-microphone circular array

The noise source maps provided by the three different means are shown in Fig 1. With the AM, the source position is correctly detected. However, the number of side lobes is very large due to the number of microphones. Moreover, the side lobes level is high which may prevent the localization of a source with a weaker level. When the GM and HM are used, all the side lobes are removed and the main lobe is narrower. The best noise source map is provided by the HM with a sharp main lobe.

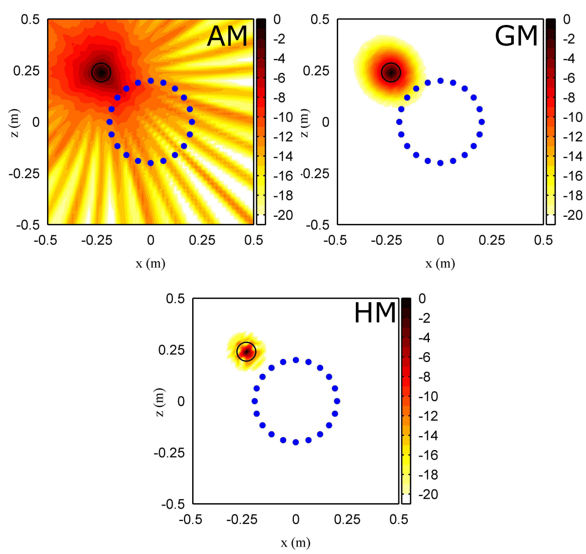

Figure 1: Noise source maps obtained with the AM, GM, and HM in the case of a 20-microphone circular array. The black circles is the source position. The blue dots are the microphone positions. The colorbar is in dB.

Noise source maps: case of a 5-microphone circular array

Now, the number of microphones is reduced to 5. The noise source maps provided by the three different means are shown in Fig 2. With the AM, the number of side lobes is smaller due to the decreasing of microphone pairs. However, the side lobes level is higher because there is less destructive interferences. With the GM and HM, the noise source maps are similar to Fig 1 without side lobes and with a sharp main lobe.

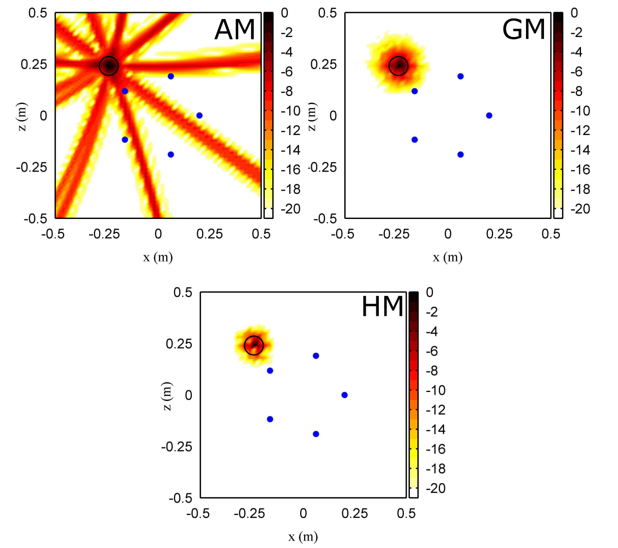

Figure 2: Noise source maps obtained with the AM, GM, and HM in the case of a 5-microphone circular array. The black circles is the source position. The blue dots are the microphone positions. The colorbar is in dB.

4 Conclusions

This paper deals with a source localization technique based on the generalized cross correlation in the time domain. Commonly, the Arithmetic Mean (AM) is used to average the spatial likelihood functions. In this study, the AM has been replaced by the Geometrical and Harmonic Mean (GM and HM). Two numerical cases involving a large or a small number of microphones were investigated. In each case, the GM and HM outperform the noise source map provided by the AM by removing the side lobes and narrowing the main lobe. Moreover, the GM and HM are less sensitive to the decrease of number of microphone as compared to the AM.

References

[1] T. Padois, P-A. Gauthier and A. Berry. *Inverse problem with beamforming regularization matrix applied to sound source localization in closed wind-tunnel using microphone array*, Journal of Sound and Vibration, Vol.**330(25)**, pages 6858-6868, 2014.

[2] T. Padois and A. Berry. *Orthogonal matching pursuit applied to the deconvolution approach for the mapping of acoustic sources inverse problem*, J. Acoust. Soc. Am., Vol.**138(6)**, pages 3678-3685, 2015.

[3] T. Padois, F. Sgard, O. Doutres and A. Berry. *Comparison of acoustic source localization methods in time domain using sparsity constraints*, Proceedings of Internoise, 9-12 August,San Francisco, CA, USA, pages 1-10, 2015.

[4] T. Padois, O. Doutres, F. Sgard, and A. Berry. *On the use of geometric and harmonic means with the generalized cross-correlation in the time domain to improve noise source maps*, J. Acoust. Soc. Am, Vol.**140(1)**, pages EL56-EL61, 2016.