26th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
7-11 July 2019, Montreal

ACOUSTIC IMAGING USING DIFFERENT WEIGHTING FUNCTIONS WITH THE GENERALIZED CROSS CORRELATION BASED ON THE GENERALIZED MEAN

Thomas Padois and Olivier Doutres
École de technologie supérieure, Montréal, Québec, Canada
email: thomas.padois@etsmtl.ca

Hugues Nelisse and Franck Sgard
IRSST, Montréal, Québec, Canada

The generalized cross correlation (GCC) is a standard technique for estimating time delay between microphone signals. A prefiltering operation by a weighting function may be included to whiten the cross spectrum of the microphones signals. The expected result is a narrow cross correlation function and a more accurate estimation of the time delay. Among the classic weighting functions, the most known is the PHAse Transform (PHAT). The ability of the PHAT weighting function to whiten the cross spectrum of the microphone signals can be improved by adding an exponent or the minimum value of the coherence function to the denominator. Both approaches have shown promising results for time delay estimation. In this work, the aforementioned modifications of the PHAT weighting function are considered for performing acoustic imaging with the classic GCC and the GCC based on the geometric mean. Numerical and experimental measurements are carried out in the case of two acoustic sources in front of a regular microphones array.
Keywords: generalized cross correlation, source localization, acoustic imaging, weighting function, geometric mean

1. Introduction

In industry, workers may be exposed to high sound pressure levels. Although, providing hearing protection devices is the easiest solution, the most efficient way is to reduce the noise of the main sources. In this case, the first step is to identify the source positions which can be done with a microphone array. The standard technique for localizing acoustic source positions is the beamforming performed in the frequency or time domain [1]. The main drawback of beamforming is an acoustic image with a large main lobe for low frequency content or high side lobe amplitude for higher frequencies [2, 3]; both preventing from an accurate localization or separation of the sources when they are close to each other.

Many frequency beamforming techniques have been developed and are usually gathered under the name of acoustic imaging techniques. Two recent reviews have presented exhaustive lists of improved acoustic imaging techniques which allow for narrowing the main lobe or removing side lobes [4, 5]. Although both review papers have focused on frequency domain beamforming, Chiarotti *et al.* [5] have noticed an increased interest in time-domain algorithms over the years and Roberto *et al.* [4] have briefly presented time domain techniques for non-stationary sources while noticing the higher computation time for time domain beamforming.

The generalized cross correlation (GCC) [6] is commonly used to estimate the time delay of arrival as noticed in the review paper on source localization [7], but it can also be used for acoustic imaging. It is worth noticing that the source localization review [7] and both acoustic imaging reviews [4, 5] seem to share the same objective but promote different techniques. This work aims to use a source localization technique for performing acoustic imaging.

Previous works have presented acoustic imaging results based on the GCC. For instance, Quaegebeur *et al.* [9] have introduced a spatial criterion with GCC and have shown the efficiency of the proposed technique for imaging loudspeakers in a reverberant environment. Padois *et al.* [8] have used an inverse problem solved with sparsity constraint for enhancing the acoustic images provided by the GCC. The low computation time of the GCC has been demonstrated in [9, 10]. Irs efficiency in narrowing the main lobe and reducing the side lobe amplitude has been shown in [11, 12] where the GCC has been computed with the generalized mean. Finally, the GCC has also been used to image the impulse noise of nail guns [13].

All the previous references have used the GCC without weighting function or with the Phase Transform [6]. However, many alternative weighting functions have been proposed for improving time delay estimation [14]. The objective of this work is to compare the acoustic images obtained with the GCC and the geometric mean for three different weighting functions such as in reference [15].

2. The generalized cross correlation and weighting functions

The generalized cross correlation (GCC) is a standard technique for performing acoustic source localization [16, 17]. The theoretical background has already been presented in reference [6], therefore a quick overview is only provided here.

The goal of the GCC technique is to estimate the time delay between all pairs of a given microphone array. Then, these time delays are interpolated over a scan zone in order to generate an acoustic image where the peak value exhibits the source position.

The time delay between two microphone signals, denoted x_m and x_n, can be estimated by the cross correlation function $\mathrm{R}_{x_m x_n}(\tau)$. Usually, the cross correlation function between two microphone signals is given by the inverse Fast Fourier Transform of the cross spectrum C_{mn}

$$\mathrm{R}_{x_m x_n}(\tau) = \sum_{k=0}^{N_f-1} W(k) C_{mn}(k) \exp\left(j2\pi \frac{k}{N_f}\tau\right), \tag{1}$$

where k is the frequency index, N_f the number of frequency samples, $j = \sqrt{-1}$ and $W(k)$ the weighting function. If the arithmetic mean of the GCC is replaced by the GEOmetric mean, the GCC becomes the GEO [11, 12].

The most known weighting function is the PHAse Transform (PHAT)

$$PHAT = \frac{1}{|C_{mn}|}. \tag{2}$$

The goal of the PHAT weighting function is to whiten the cross spectrum of the microphone signals. When PHAT is used, the GCC becomes GCC-PHAT.

In 1996, Rabinkin *et al.* [18] proposed to partially whiten the cross spectrum by adding an exponent ρ to the weighting function

$$\rho - PHAT = \frac{1}{|C_{mn}|^{\rho}}. \tag{3}$$

When $\rho = 0$ or $\rho = 1$, it corresponds to the cross correlation function or the GCC-PHAT, respectively. If $0 < \rho < 1$, the cross spectrum is partially whiten and the GCC becomes GGC-ρ-PHAT.

In 2009,Shen *et al.* [19] introduced the cross power spectrum phase with coherence function. More specifically, the minimum value of the coherence function γ_{mn}^2 is added to the denominator of PHAT

$$\rho - PHAT - C = \frac{1}{|C_{mn}|^{\rho} + \min \gamma_{mn}^2} \tag{4}$$

in this case, the GCC becomes GGC-ρ-PHAT-C.

3. Numerical acoustic images

Numerical simulations are first considered in order to highlight the performance of the different weighting functions. Two acoustic sources are set in front of a regular 49-microphone array. The source positions are x=-0.25 m and x=0.25 m at y=0. The sources signals are white Gaussian noises. The microphones record the acoustic signal at a frequency sampling of 44,100 Hz. The microphone signals are filtered by a 2^{nd} order bandpass Butterworth in the 1000 Hz octave band. The microphone array aperture is 0.75 m in both directions and the source-array distance is 1.2 m. The scan zone, where the source is searched, is a square with 2 m side with a 40,401 points (201$\times$201).

The acoustic image provided by the GCC is shown in Figure 1.a. The source positions are correctly detected, but the source separation is not apparent due to the merging of the main lobes. In this case, more efficient source localization techniques are required to improve the acoustic image which means decreasing the main lobe widths. When $\rho = 0.8$, the ρ-PHAT and ρ-PHAT-C slightly improve the source separation by decreasing the amplitude in between the sources (Figure 1.b-d) which means that the weighting functions decrease the main lobe width of the cross-correlation function. However, this decrease is not sufficient for separating both sources. With the classic PHAT, the sources are more separated but at the expense of several side lobes especially along x-direction (Figure 1.c). The weighting function ρ-PHAT-C with $\rho = 1$ provides a similar trend although the side lobes are less spread (Figure 1.e). Using the GEO does not allow for separating both sources (not shown here). However, with ρ-PHAT-C ($\rho = 1$) and the GEO, the best source separation is achieved, two distinct spots are present at the source positions without side lobes (Figure 1.f).

4. Experimental acoustic images

Now, the efficiency of the weighting functions for separating two sources is investigated with experimental data. The experiment took place in a hemi-anechoic room. Two loudspeakers were set in front of a regular 16-microphone array. The microphone array aperture and source-array distance are similar to the ones used in the numerical simulations, only the number of microphones being decreased. Again, the 1000 Hz octave band is considered and the scan zone is the same.

The acoustic image obtained with the GCC is similar to the numerical one, the sources are not separated (Figure 2.a). However, side lobes surround the main lobes now. When $\rho = 0.8$, the acoustic images are not improved with the ρ-PHAT and ρ-PHAT-C (Figure 2.b-d). With the classic PHAT, the source positions are no more detected (Figure 2.c). The weighting function ρ-PHAT-C with $\rho = 1$ allows for removing a part of the side lobes surrounding the main lobes, but extend them along the x-direction (Figure 2.e). Again, the best source separation is provided by the GEO and ρ-PHAT-C where two distinct spots are present at the source positions without side lobes (Figure 2.f).

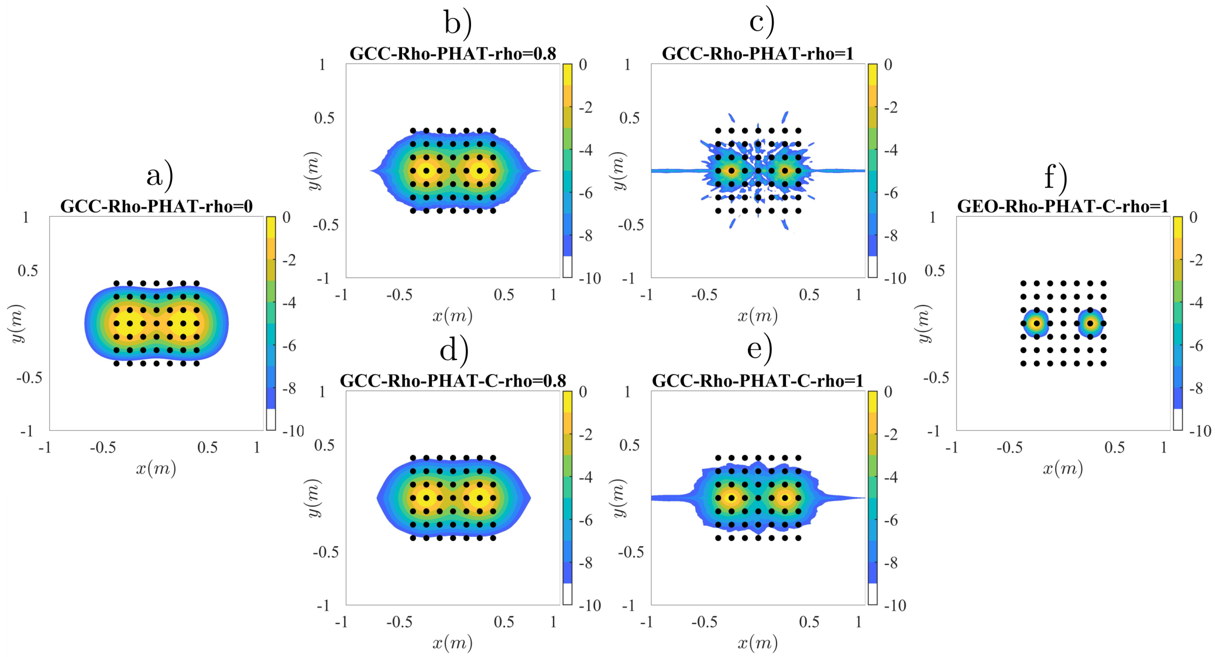

Figure 1: Numerical acoustic images obtained for the 1000 Hz octave band with a) GCC b) $\rho = 0.8$-PHAT, c) $\rho = 1$-PHAT d) $\rho = 0.8$-PHAT-C, e) $\rho = 1$-PHAT-C and f) GEO $\rho = 1$-PHAT-C. The black dots are the microphone positions. The colorbar is in dB.

5. Conclusion

The main objective of an acoustic imaging technique is to narrow the main lobe and to reduce the side lobe amplitude in order to accurately separate multiple sources. In this work, the generalized cross correlation (GCC) was used to detect the source positions. Numerical and experimental acoustic images showed that the GCC is not able to separate two sources closely spaced due to the main lobe width. Weighting functions can be used for improving the estimation of the cross correlation function. The classic PHAT was firstly considered. Although the numerical acoustic image provided by PHAT is slightly better then the acoustic image provided by GCC, the latter is not able to detect the source positions in the case of experimental data. Two others weighting function (denoted ρ-PHAT and ρ-PHAT-C) were then evaluated. While the former adds an exponent to the PHAT, the latter also adds the coherence function. Both weighting functions allows for slightly improving the acoustic images. However, the best results is obtained when the geometric mean is applied with the ρ-PHAT-C. In this case, both sources are clearly separated and no side lobes are present.

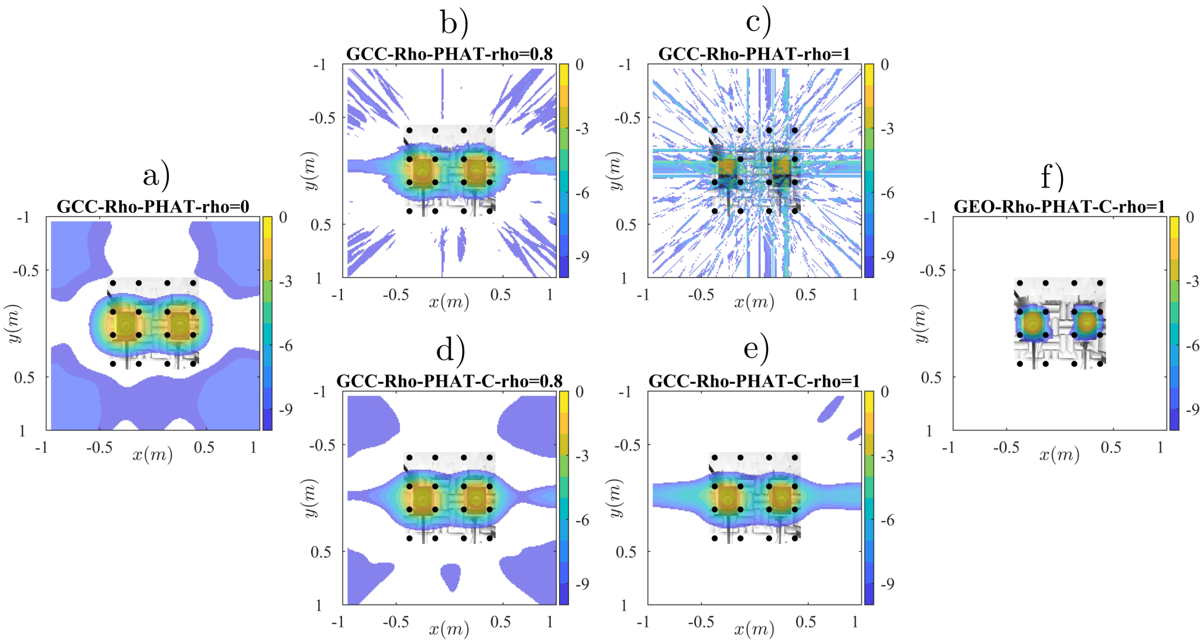

Figure 2: Experimental acoustic images obtained for the 1000 Hz octave band with a) GCC b) $\rho = 0.8$-PHAT, c) $\rho = 1$-PHAT d) $\rho = 0.8$-PHAT-C, e) $\rho = 1$-PHAT-C and f) GEO $\rho = 1$-PHAT-C. The black dots are the microphone positions. The colorbar is in dB.

REFERENCES

1. Johnson, D. H. and Dudgeon, D. E., *Array Signal Processing: Concepts and Techniques*, Prentice Hall,Upper Saddle River,New Jersey,USA, (1993).

2. Camier, C., Padois, T., Gauthier, P-A., Berry, A., Blais, J-F., Patenaude-Dufour, M. and Lapointe, R., Fly-over source localization on civil aircraft, *19th AIAA/CEAS Aeroacoustics Conference*, Berlin, Germnay, 27–29 May, (2013).

3. Padois, T., Laffay, P., Idier, A. and Moreau,. S., Detailed experimental investigation of the aeroacoustic field around a Controlled-Diffusion airfoil, *21st AIAA/CEAS aeroacoustics conference*, Dallas, Texas, 22–26 June, (2015).

4. Merino-Martínez, R., Sijtsma, P., Snellen, M. Ahlefeldt, T., Antoni, J., Bahr, C. J., Blacodon, D., Ernst, D., Finez, A., Funke, S., Geyer, T. F., Haxter, S., Herold, G., Huang, X., Humphreys, W. M., Leclère, Q., Malgoezar, A., Michel, U., Padois, T., Pereira, A., Picard, C., Sarradj, E., Siller, H., Simons, D. G. and Spehr, C., A review of acoustic imaging methods using phased microphone arrays, *CEAS Aeronautical Journal*, 1–34, (2019).

5. Chiariotti, P., Martarelli, M. and Castellini, P., Acoustic beamforming for noise source localization - Reviews, methodology and applications, *Mechanical Systems and Signal Processing*, **120**, 422–448, (2019).

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

7. Cobos, M., Antonacci, F., Alexandridis, A., Mouchtaris, A. and Lee, B., A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks, *Wireless Communications and Mobile Computing*, (2017).

8. Padois, T., Doutres, O., Sgard, F. and Berry, A., Time domain localization technique with sparsity constraint for imaging acoustic sources, *Mechanical Systems and Signal Processing*, **94**, 85–93, (2017).

9. Quaegebeur, N., Padois, T., Gauthier, P-A. and Masson, P., Enhancement of time-domain acoustic imaging based on generalized cross-correlation and spatial weighting, *Mechanical System and Signal Processing*, **75**, 512–524, (2015).

10. Padois, T., Sgard, F. Doutres, O. and Berry, A., Acoustic source localization using a polyhedral microphone array and an improved generalized cross-correlation techniqu, *Journal of Sound and Vibration*, **386** (6), 82–99, (2017).

11. Padois, T., Doutres, O., Sgard, F. and Berry, A., 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.*, **140** (1), EL56–EL61, (2016).

12. Padois, T., Acoustic source localization based on the generalized cross-correlation and the generalized mean with few microphones, *J. Acoust. Soc. Am.*, **15** (2), EL393–EL398, (2018).

13. Padois, T., Gaudreau, M-A, Marcotte, P. and Laville, F. Identification of noise sources using a time domain beamforming on pneumatic, gas and electric nail guns, *Noise Control Engr. J.*, **67** (1), 11–22, (2010).

14. Marinescu, R.-S., Buzo, A., Cucu, H. and Burileanu, C. Applying the Accumulation of Cross-Power Spectrum Technique for Traditional Generalized Cross-Correlation Time Delay Estimation, *International Journal on Advances in Telecommunications*, **6**, 98–108, (2013).

15. Padois, T., Doutres, O. and Sgard, F., On the use of modified phase transform weighting functions for acoustic imaging with the generalized cross correlation, *J. Acoust. Soc. Am.*, **xx**, xx–xx, (2019).

16. Dmochowski, J. P., Benesty, J. and Affes, S. A Generalized Steered Response Power Method for Computationally Viable Source Localization, *IEEE Transactions ON Audio, Speech, And Language Processing*, **15** (8), 2510–2526, (2007).

17. Velasco, J., Pizarro, D. and Macias-Guarasa, J. Source localization with acoustic sensor arrays using generative model based fitting with sparse constraints, *Sensors*, **12**, 13781–13812, (2012).

18. Rabinkin, D. V., Renomeron, R. J., Dahl, A. J., French, J. C., Flanagan, J. L. and Bianchi M., DSP implementation of source location using microphone arrays, *Proc. SPIE 2846, Advanced Signal Processing Algorithms, Architectures, and Implementations VI*, (1996).

19. Shen, M. and Liu, H., A Modified Cross Power-Spectrum Phase Method Based on Microphone Array for Acoustic Source Localization, *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 1286–1291, (2009).