Comparison of acoustic centering maps for radiation capture of musical instruments with spherical microphone arrays

Daniel Deboy[1], Franz Zotter[2]
[1] *Student, KU-/TU-Graz, Email: daniel.deboy@student.tugraz.at*
[2] *Institut für Elektronische Musik und Akustik, KU-Graz, Email: zotter@iem.at*

Introduction

Recently, the capture of sound radiated from musical instruments has been topic of research, using surrounding spherical microphone arrays. Astonishingly, spatial aliasing errors of these arrays depend on accurate acoustic centering of the sound source as shifting yields higher-order spherical harmonic components. Acoustic centering, however, is nontrivial for musical instruments if not impossible. Therefore using different measures like weights of spherical harmonic components can be used to roughly track the acoustical center for each partial of the musical sound after the recording has been made. Eventually, the paper investigates different strategies for tracking with respect to the uniqueness of the positions by visualizing their localization map. The suggested method shall give a more compact description of radiation-patterns extracted from spherical microphone array recordings, which can be more efficient in further analysis and synthesis processes.

Spherical exterior problem

The spherical base-solution (math./phys.) are capable of fully expanding sound fields that radiate from a radial point of origin $r = 0$ with the coefficients c_{nm}

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{c_{nm} h_n^{(2)}(kr)}_{:=\psi_n^m(kr)} Y_n^m(\boldsymbol{\theta}), \qquad (1)$$

the spherical Hankel-functions[1] $h_n^{(2)}(kr)$ depending on the wave-number $k = \omega/c$, ($\omega = 2\pi f$, $c \approx 343\text{m/s}$), and the radius r, and the spherical harmonic $Y_n^m(\boldsymbol{\theta})$ depending on the cartesian unit vector $\boldsymbol{\theta} = (\cos(\varphi)\sin(\vartheta), \sin(\varphi)\sin(\vartheta), \cos(\vartheta))^{\mathrm{T}}$ the direction of which depending on azimuth φ and zenith ϑ, cf. [1, 2].

As the coefficients c_{nm} fully describe the radial and angular propagation of the field, e.g. sound-pressure, they can be called the *wave-spectrum*. The term $\psi_n^m(kr) = c_{nm}\, h_n^{(2)}(kr)$ marked in eq. (1) is called *spherical wave-spectrum* as defined in [1]. It describes the sound-field, e.g. pressure, expanded into spherical harmonics at any sphere of constant radius kr and is defined by the transform integral

$$\psi_n^m(kr) = \int_{\mathbb{S}^2} p(kr, \boldsymbol{\theta}) Y_n^m(\boldsymbol{\theta}) d\boldsymbol{\theta}. \qquad (2)$$

[1]It is a matter of convention which kind $h_n^{(1)}(kr)$, or $h_n^{(2)}(kr)$ of Hankel function to use. This paper uses the spherical Hankel function of the second kind.

Assuming the L discrete observations of the sound-pressure by a spherical array to be a linear combination of spherical harmonics

$$\boldsymbol{p}_{\mathrm{L}} = \boldsymbol{Y}_{\mathrm{N}}\, \hat{\boldsymbol{\psi}}_{\mathrm{N}}, \qquad (3)$$

$$\text{with } \boldsymbol{Y}_{\mathrm{N}} = \begin{pmatrix} Y_0^0(\boldsymbol{\theta}_1) & Y_1^{-1}(\boldsymbol{\theta}_1) & \dots & Y_{\mathrm{N}}^{\mathrm{N}}(\boldsymbol{\theta}_1) \\ \vdots & \vdots & \ddots & \vdots \\ Y_0^0(\boldsymbol{\theta}_{\mathrm{L}}) & Y_1^{-1}(\boldsymbol{\theta}_{\mathrm{L}}) & \dots & Y_{\mathrm{N}}^{\mathrm{N}}(\boldsymbol{\theta}_{\mathrm{L}}) \end{pmatrix}, \qquad (4)$$

an angularly band-limited version $\psi_n^m(kr) = 0 : n > \mathrm{N}$ of the *spherical wave-spectrum* $\boldsymbol{\psi}_{\mathrm{N}}$ is obtained by matrix inversion

$$\hat{\boldsymbol{\psi}}_N(kr) = \boldsymbol{Y}_N^{-1} \boldsymbol{p}_L. \qquad (5)$$

The *wave-spectrum* $\boldsymbol{c}_N$ is calculated from eq. (5) by a diagonal matrix $\boldsymbol{H}_{\mathrm{N}}$ containing the associated radial propagation terms

$$\boldsymbol{c}_N = \boldsymbol{H}_N^{-1} \hat{\boldsymbol{\psi}}_N(kr) = \boldsymbol{H}_N^{-1} \boldsymbol{Y}_N^{-1} \boldsymbol{p}_L, \qquad (6)$$

$$\boldsymbol{H}_N = \mathrm{diag}_{\mathrm{N}}\{h_n^{(2)}(kr)\}. \qquad (7)$$

Coordinate transforms: Re-expansion into spherical base solutions

Translations described by $\boldsymbol{r}' = \boldsymbol{r} + (d_x, d_y, d_z)^{\mathrm{T}} = \boldsymbol{r} + \boldsymbol{d}$ express what happens if the sources inside the spherical array is shifted. A sound-field described by the *wave-spectrum* c_{nm} at $\boldsymbol{r}$ can be re-expanded into a *wave-spectrum* c'_{nm} at $\boldsymbol{r}'$. Translations yield a mapping $c_{nm} \xrightarrow{\boldsymbol{d}} c'_{nm}$, cf. [2, 3]

$$c'_{nm} = \sum_{n'} \sum_{m'} c_{n'm'}\, T_{n'n}^{m'm}(\boldsymbol{d}) \qquad (8)$$

Translation tends to create components of higher orders $c'_{nm} \neq 0 : n > \mathrm{N}$. That also means, that the assumption of a band-limited *spherical wave-spectrum* $\psi_n^m(kr)$ in eq. (5) may not always hold for arbitrarily shifted sources. Nevertheless, this assumption is kept to allow for the calculation of a band-limited *spherical wave-spectrum* by eq. (5), using hyperinterpolation [4, 5].

Band-limitation of misaligned sources

A correct representation of a source of known angular band-limit N after a displacement by d requires the re-expansion to have wave-spectral coefficients of a higher band-limit, i.e. $\mathrm{N}' > \mathrm{N}$. For the present surrounding spherical array analysis this means off-center sound sources cause higher-order components that technically need to be resolved by the array without spatial misinterpretations, i.e. aliasing [5]. A rough rule-of thumb could be given as

$$\mathrm{N}' \geq \mathrm{N} + kd. \tag{9}$$

However, it is hard to know or estimate the band-limit N of an unidentified source. Even more, the misalignment d of its acoustic center from the center of the measurement array is hard to know or estimate in advance.

Therefore, as a working hypothesis, let us *assume* the correctness of the *wave-spectrum* c_{nm} calculated in eq. (6).

Cost functions

Two approaches have been examined for the static tracking of the acoustic center. One is the complex-squared sum of the discrete sound pressure distribution

$$\begin{aligned} J_{ssc}(\boldsymbol{d}) &= 1 - \left| \frac{\int p(\boldsymbol{\theta})^2 \, d\boldsymbol{\theta}}{\int |p(\boldsymbol{\theta})|^2 \, d\boldsymbol{\theta}} \right| \qquad (10) \\ &= 1 - \left| \frac{\boldsymbol{\psi}'^{\mathrm{T}}_{\mathrm{N}'} \boldsymbol{\psi}'_{\mathrm{N}'}}{\|\boldsymbol{\psi}'_{\mathrm{N}'}\|^2} \right|, \end{aligned}$$

the other is the center of mass of the spherical harmonics components

$$J_{mc}(\boldsymbol{d}) = \frac{\boldsymbol{c}'^{H}_{\mathrm{N}'} \mathrm{diag}_{\mathrm{N}'}\{w_n\} \boldsymbol{c}'_{\mathrm{N}'}}{\boldsymbol{c}'^{H}_{\mathrm{N}'} \boldsymbol{c}'_{\mathrm{N}'}}, \tag{11}$$

where w_n is a penalizing weight increasing with the order n. A detailed examination of the proposed cost functions can be found in [6].

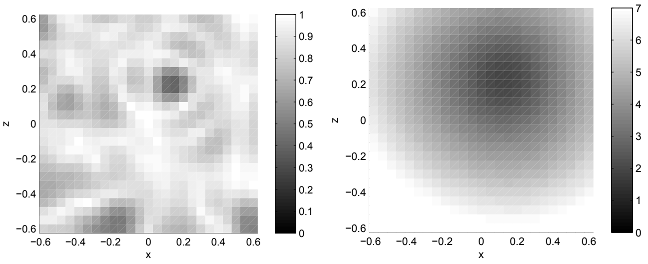
Optimization and centering maps

Both cost functions can be evaluated within the volume of the array. A global minimum should give an approximated dislocation of the acoustic center of the sound source. The global minimum can be determined using search methods like the simplex-search algorithm.

Fig. 1 and Fig. 2 show the centering maps of different musical instruments recorded with the 64 channel microphone array at IEM Graz. The cost functions are evaluated at slices through the spherical volume. Slicing is done at the assumed acoustic centers given by the optimization algorithm.

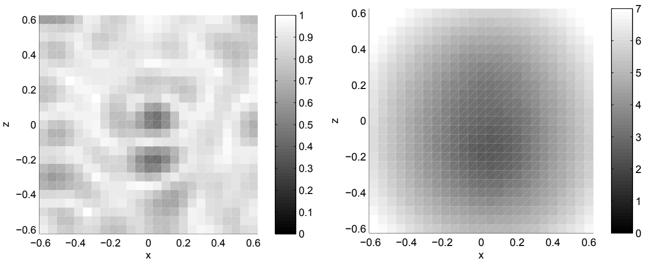
Conclusion

A cost function utilizing the center of mass of the spherical harmonic components is a robust measure for roughly tracking the acoustic center of an musical-instrument whereas the complex-squared sum criterion evokes sharper minima but fails for complex-valued wave-spectrum coefficients. A combination of both measures may result in a robust an precise localization of an acoustic center. The reduction of higher-order components to a minimum is beneficial when using playback systems which have a low order resolution.



(a) J_{ssc} criterion (a) J_{mc} criterion

Figure 1: Example of centering maps for a trombone recording, 2nd partial at 529 Hz. The spherical volume is sliced at the optimization result of the J_{mc} criterion. Both maps show the x-z-plane of the volume.



(a) J_{ssc} criterion (a)J_{mc} criterion

Figure 2: Example of centering maps for a bassflute recording, 2nd partial at 526 Hz. The spherical volume is sliced at the optimization result of the J_{mc} criterion. Both maps show the x-z-plane of the volume.

References

[1] E. G. Williams, *Fourier Acoustics - Sound Radiation and Nearfield Acoustical Holography.* Academic Press, 1999.

[2] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*, 1st ed. Elsevier, 2004, no. 0-08-044371-0.

[3] B. Rafaely, "Spatial alignment of acoustic sources based on spherical harmonics radiation analysis," 2010.

[4] F. Hohl, "Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten," Master's thesis, Institut für Elektronische Musik und Akustik, Universität für Musik und darstellende Kunst, Graz, Österreich, November 2009.

[5] F. Zotter, "Sampling strategies for acoustic holography/ holophony on the sphere," in *NAG-DAGA, Rotterdam*, 2009.

[6] D. Deboy, "Acoustic centering and rotational tracking in surrounding spherical microphone arrays," Master's thesis, Institut für Elektronische Musik und Akustik, Universität für Musik und darstellende Kunst, Graz, Österreich, 2010.