AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

INTERAURAL CROSS CORRELATION AND SPATIAL CORRELATION IN A SOUND FIELD REPRESENTED BY SPHERICAL HARMONICS

Amir Avni [1], Boaz Rafaely[2]

[1] Department of Electrical and Computer Engineering,
Ben-Gurion University of the Negev,
Beer-Sheva 84105 ,
Israel
(amiravni@yahoo.com)
[2] (br@ee.bgu.ac.il)

Abstract: *Spaciousness is an important psychoacoustic feature in room acoustics, with the interaural cross correlation (IACC) an accepted parameter for its measure, the latter employing the head head-related transfer functions (HRTF). Recently, spherical microphone arrays have been studied for room acoustics analysis and music recordings. As these arrays typically use a finite number of microphones, they may not be able to capture the spatial information required for complete spatial analysis or for sound reproduction with realistic spaciousness. This study employs spherical harmonics representations for both the HRTF data and the sound field data, facilitating IACC analysis for sound fields represented by a finite order in the spherical harmonics domain. As diffuse sound fields, often used to model sound in reverberant rooms, are characterized by spatial correlation, the relation between IACC and spatial correlation for diffuse sound fields is studied in the spherical harmonics domain. The effect of limited spherical harmonics order on the spaciousness of diffuse and other sound fields is presented using simulated and measured data.*

Key words: LEV, IACC, diffuse field, spherical harmonics, spatial-temporal correlation

1 INTRODUCTION

Two important measures that affect spatial impression, or spaciousness, in concert hall are apparent source width (ASW) and listener envelopment (LEV). ASW relates to the spatial width of the perceived sound source, and is affected by the degree of dissimilarity of the musical sound reaching the two ears in the first 80-100 ms after the direct sound [1]. LEV relates to the density and the spatial distribution of reflections reaching the ears, and is mostly affected by the sound arriving 80-100 ms after the direct sound [2] [3]. Although ASW and LEV are subjective measures, some objective measures have been developed to relate to these measures, among these an important measure is the interaural cross correlation (IACC)[4]. IACC is measured using a human or a dummy head and is computed from the time correlation between the two ears. IACC is affected by the head related impulse response (HRIR) or head related transfer function (HRTF) which are the time and frequency response functions between a source and the ears of a listener.

Spherical microphone arrays have been studied recently for a broad range of applications, in particular analyzing room acoustics in three-dimensional sound fields, using, for example, plane-wave decomposition of the sound field [5] [6]. Spherical array data combined with HRTF data can be used to measure acoustic parameters in rooms such as IACC. [7]

The diffuse sound field model is often used when analyzing late reflections in halls, and can therefore be useful when calculating IACC with late reflections i.e. IACC$_L$. This paper presents a relation between spatial-temporal correlation in a diffuse field and IACC. Analysis of IACC and diffuse fields are made in the spherical harmonics domain in order to investigate the effect of sound field order in the spherical harmonics domain on IACC$_L$.

2 PLANE WAVES AND SPHERICAL FOURIER TRANSFORM

Consider a sound pressure function $p(k, r, \theta, \phi)$, with (r, θ, ϕ) the standard spherical coordinate system, which is square integrable over $\Omega \equiv (\theta, \phi)$, with k the wavenumber. Its spherical Fourier transform (SFT), $p_{nm}(k, r)$ and the inverse spherical Fourier transform (ISFT) are defined [8] by:

$$p_{nm}(k, r) = \int_{\Omega \in S^2} p(k, r, \Omega) Y_n^{m^*}(\Omega) d\Omega \tag{1}$$

$$p(k, r, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_{nm}(k, r) Y_n^m(\Omega) \tag{2}$$

where $\int_{\Omega \in S^2} d\Omega \equiv \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta\, d\phi$, and the spherical harmonics are defined by:

$$Y_n^m(\theta,\phi) \equiv \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi} \quad (3)$$

with n the order of the spherical harmonics and P_n^m is associated Legendre function. We consider a plane wave with unit amplitude arriving from (θ_l, ϕ_l) , the pressure at the position (r,θ,ϕ) due to the plane wave is $p_l(kr,\theta,\phi)$ and its SFT is given by [9]:

$$\begin{aligned} p_l(kr,\theta,\phi) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} b_n(kr) Y_n^{m*}(\theta_l,\phi_l)\, Y_n^m(\theta,\phi) \\ &= \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} b_n(kr) P_n(\cos\Theta) \end{aligned} \quad (4)$$

where Θ is the angle between (θ_l,ϕ_l) and (θ,ϕ), and the equality in (4) is derived using the spherical harmonic addition theorem [10]. $b_n(kr)$ is defined for open sphere and rigid sphere as follows:

$$b_n(kr) = \begin{cases} 4\pi i^n \left(j_n(kr) - \frac{j_n'(ka)}{h_n'(ka)} h_n(kr) \right) & \text{rigid} \\ 4\pi i^n j_n(kr) & \text{open} \end{cases} \quad (5)$$

where j_n is the spherical Bessel function and h_n is the spherical Hankel function and j_n', h_n' are derivatives. Analyzing sound fields in the spherical harmonics domain usually requires a spherical microphone array, configured around an open or rigid sphere. One advantage of analyzing the sound field in the spherical harmonics domain is the ability to decompose the pressure function into plane waves [11] therefore estimating the number and amplitudes of the plane waves composing the sound field. Furthermore, a rigid sphere may provide an approximation for a human head, and therefore can be useful when analyzing the expected sound field in the presence of a listener. When using a spherical microphone array, a finite order of the spherical harmonics will be used, leading to measurement errors that depend on the spherical harmonics order, the number and locations of the microphones (samples over the sphere) and the maximum frequency of the sound field [12].

3 IACC CALCULATION FROM HRTF AND SOUND FIELD DATA

HRTF is the frequency response between a source and the left or the right ear ,$H^l(k,\Omega)$ and $H^r(k,\Omega)$ respectively, with $k = \frac{2\pi f}{c}$ and c the speed of sound. Combining the HRTFs with $a(k,\Omega)$, the amplitude density of the plane waves from all directions, we get the Fourier transform of the pressure at the left ear [7]:

$$p_l(k) = \int_{\Omega \in s^2} a(k,\Omega) H^l(k,\Omega) d\Omega \quad (6)$$

Denoting $\tilde{a}(k,\Omega) \equiv a^*(k,\Omega)$ where "$*$" denotes complex conjugate , substituting Eq. (2) in Eq. (6) and using the orthogonality property of the spherical harmonics, Eq. (6) can be simplified, written here for both ears:

$$\begin{aligned} p_l(k) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \tilde{a}_{nm}^*(k) H_{nm}^l(k) \\ p_r(k) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \tilde{a}_{nm}^*(k) H_{nm}^r(k) \end{aligned} \quad (7)$$

Calculating the inverse Fourier transform of the pressure will produce the pressure in the time domain, which can be used to calculate the interaural cross correlation function [13]:

$$\rho_{t_1,t_2}(\tau) = \frac{\int_{t_1}^{t_2} p_l(t) p_r(t+\tau) dt}{\sqrt{\int_{t_1}^{t_2} p_l^2(t) dt \cdot \int_{t_1}^{t_2} p_r^2(t) dt}} \quad (8)$$

IACC is then defined as the maximum of the absolute value over τ:

$$IACC_{t_1,t_2} = \max_{\tau} |\rho_{t_1,t_2}(\tau)|, \quad \tau \in (-1,1)\,\text{msec} \quad (9)$$

4 SPATIAL CORRELATION IN A DIFFUSE FIELD

A method to determine the spatial correlation in a diffuse field has been presented by Cook et al. [14] and developed by Nicholas et al. [15]. A diffuse sound field is defined as consisting of plane waves with equal magnitude and random phases. The time correlation between two measurement points due to a single harmonic plane wave, can be shown to equal $\cos(\omega\tau + kr\cos(\theta))$, when θ is the angle between the wave incident direction and the line connecting the two points. This term can be integrated over a sphere with equal weights to derive the spatial-temporal correlation in a diffuse field, given by [16]:

$$\begin{aligned} \rho(k,r,\tau) &= \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} \\ &\quad \cos(\omega\tau + kr\cos(\theta)) \sin\theta d\theta d\phi \\ &= \frac{\sin kr}{kr}\cos(\omega\tau) \end{aligned} \quad (10)$$

If the sound field has a band of frequencies of equal weights, from ω_1 to ω_2 , then the spatial-temporal correlation can be calculated using the integration over all bands:

$$\rho_{BB}(r,\tau) = \frac{1}{\omega_2 - \omega_1}\int_{\omega_1}^{\omega_2} \frac{\sin(\omega\frac{r}{c})}{\omega\frac{r}{c}} \cos(\omega\tau) d\omega \quad (11)$$

Computing the spatial-temporal correlation in the spherical harmonics domain can be useful for generalization to the case of a measurement around a rigid sphere. Assuming a harmonic plane wave, the measured acoustic pressure at a given measurement point will be defined as the real part of the complex pressure $p_i = \mathrm{Re}\left\{p_{\phi i} e^{j\omega t}\right\}$, when $p_{\phi i}$ is the complex amplitude. Computing time correlation between two different measurement points gives:

$$R(\tau) = \frac{1}{2}\mathrm{Re}\left\{e^{j\omega\tau} p_{\phi 1} p_{\phi 2}^*\right\} \quad (12)$$

Assuming two measurement points at coordinates $(r, \frac{\pi}{2}, 0)$ and $(r, \frac{\pi}{2}, \pi)$, where r is constant, Eq. (12) can be integrated over the sphere to represent a diffuse field. By using Eq. (4) and the orthogonality property of the Legendre polynomials, the integration will give:

$$R(k,\tau) = \frac{\cos(\omega\tau)}{2} \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)}{4\pi} |b_n(kr)|^2 \quad (13)$$

The autocorrelation for the same measurement point would be $\frac{1}{2}|p_{\phi i}|^2$, and integrating over the sphere will give:

$$R(k) = \frac{1}{2} \sum_{n=0}^{\infty} \frac{(2n+1)}{4\pi} |b_n(kr)|^2 \quad (14)$$

and overall the spatial-temporal correlation coefficient will be:

$$\rho(k,\tau) = \frac{\sum_{n=0}^{\infty} (-1)^n (2n+1) |b_n(kr)|^2 \cos(\omega\tau)}{\sum_{n=0}^{\infty} (2n+1) |b_n(kr)|^2} \quad (15)$$

For a broadband sound field a similar definition would apply as in Eq. (11):

$$\rho_{BB}(\tau) = \frac{\int_{\omega_1}^{\omega_2} \sum_{n=0}^{\infty} (-1)^n (2n+1) \left|b_n(\omega \frac{r}{c})\right|^2 \cos(\omega\tau) d\omega}{\int_{\omega_1}^{\omega_2} \sum_{n=0}^{\infty} (2n+1) \left|b_n(\omega \frac{r}{c})\right|^2 d\omega} \quad (16)$$

Notice that when choosing $b_n(kr)$ as defined in Eq. (5) for an open sphere, Eq. (16) reduces to Eq. (11).

IACC$_L$ is the IACC which is measured for sound arriving approximately 80-100 ms after the direct sound, and is considered to be related to LEV. Reverberant sound can usually be modeled by the diffuse sound field model and therefore this model can be useful in estimating IACC$_L$. Furthermore, calculating IACC$_L$, can be useful in estimating the level of diffuseness of the sound field. Computing the spatial-temporal correlation of a diffuse field for the selected measurement points as in Eq. (16), may be useful representation for the measurement of IACC$_L$.

5 SIMULATION STUDY

In this section, a simulation study has been performed aiming to investigate the relations developed above. A sound field composed of harmonic plane waves was simulated, where the distance r was chosen as 0.09 meters to represent the radius of an average head. For a single frequency sound field, a range up to 10KHz was analyzed and for broadband sound field, octave bands from 125Hz to 8KHz were analyzed. Figure 1 presents an analysis of a diffuse sound field with a single frequency in free space. A comparison has been made of Eq. (15), when $\tau = 0$ for a maximum value, with different number of spherical harmonics coefficients, when b_n was chosen to represent the sound field in free space. The value $N = 20$ was chosen to construct $\rho(\tau)$ without errors in the given frequency range. Figure 1 shows that a smaller number of coefficients is required to construct $\rho(\tau)$ without errors for lower frequencies. The symbols X_N is marked on the figure, N denoting the number of coefficients, indicate the point where $N = kr$. It seems that up to the point where $N = kr$ the errors are smaller and the number of coefficients is sufficient for approximating $\rho(\tau)$. Computing using sound field with larger frequencies will cause noticeable errors. Figure 2 presents the same analysis as Fig. 1, for a b_n which represent the sound field measured on a rigid sphere. It seems that here $N = kr$ also indicates the maximum frequency where a given order N will be sufficient for approximating $\rho(\tau)$.

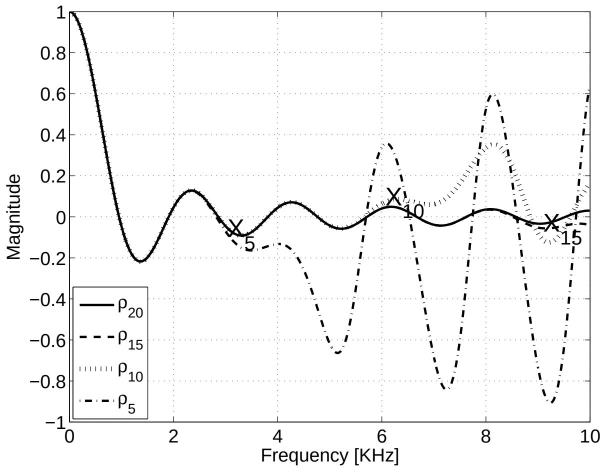

Figure 1: The magnitude of ρ_N for N coefficients, where b_n is defined for an open sphere. The X_N marks indicate the frequency where $N = kr$.

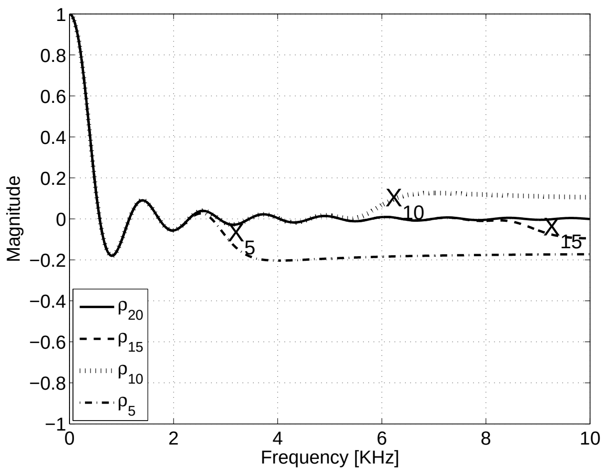

Figure 2: The magnitude of ρ_N for N coefficients, where b_n is defined for a rigid sphere. The X_N marks indicate the frequency where $N = kr$.

Figure 3 presents the difference between using b_n for an open and rigid sphere in Eq. (15), using $N = 20$ as the number of coefficients constructing both functions. As expected, the correlation between two measurement points on a rigid sphere will drop faster as frequency increases, due to bending of sound fields around the rigid sphere causing plane waves to travel a larger distance between the points. The noticeable difference emphasizes the need for using the rigid sphere b_n especially when comparing the computations of IACC$_L$, where a listener's head is present and affecting the sound field.

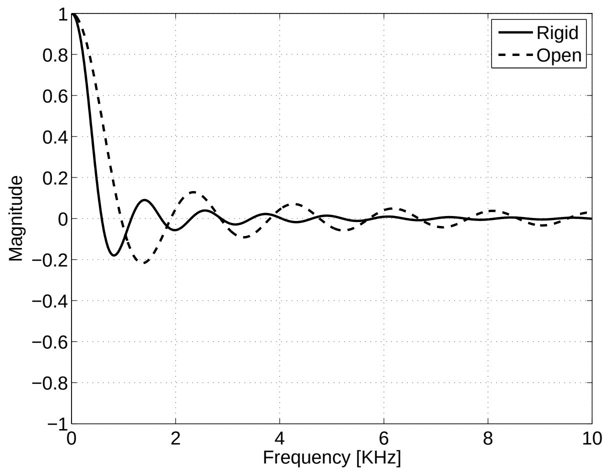

Figure 3: The magnitude of Eq. (15), comparing b_n with an open and a rigid sphere

Table 1 presents several numerical computations of the spatial-temporal correlation in a diffuse field of an octave band sound field. The computations were made using Eq. (16) where the limits of the integration were chosen according to different octave bands. The values presented are the maximum absolute value of $\rho(\tau)$, where $\tau \in (-1, 1)\,\mathrm{msec}$. The computations were made using both b_n as open and rigid sphere as indicated in the table. N, which was calculated as $\lceil max(kr) \rceil$, represents the minimum coefficients needed in order to achieve a good estimation of the spatial-temporal correlation. The results shown in table 1 may be a good estimation for the real values of IACC$_L$, measured in a diffuse sound field.

Octave [Hz]	Open $max\,|\rho|$	Rigid $max\,|\rho|$	N
125	0.9673	0.9250	1
250	0.8737	0.7029	1
500	0.5584	0.1097	2
1000	0.1213	0.0747	3
2000	0.0772	0.0279	5
4000	0.0375	0.0095	10
8000	0.0179	0.0029	19

Table 1: Correlation in a diffuse sound field for different octave bands

6 CONCLUSION

IACC is an important measure for estimation of spaciousness in concert halls. IACC$_L$ refers to IACC measurements of late reflections, usually in reverberant sound fields. Since the sound field can be modeled as a diffuse sound field, IACC$_L$ can be estimated by this model. Spatial-temporal correlation of a diffuse field may therefore provide a good representation for IACC$_L$ in a diffuse field, although the fine attributes of the HRTF may not be modeled. Analyzing this model in the spherical harmonics domain can give a better understanding of the behavior of IACC$_L$, when instead of using a dummy head, the rigid sphere model is used. Furthermore, this study may help understand the relations between the order of the spherical harmonics and the spatial perception of sound.

7 ACKNOWLEDGMENT

This research was supported by THE ISRAEL SCIENCE FOUNDATION (grant No. 155/06).

REFERENCES

[1] L. L. Berank. concert hall acoustics - 2008. *J. Audio Eng. soc.*, 56(7/8):532–544, July/August 2008.

[2] J. S. Bradley, R. D. Reich, and S. G. Norcross. On the combined effects of early- and late-arriving sound on spatial impression in concert halls. *J. Acoust. Soc*, 108(2):651–661, August 2000.

[3] Kazuhiro Iida Masayuki Morimoto, Koichi Nakagawa. The relation between spatial impression and the law of the first wavefront. *Applied Acoustics*, 69:132–140, 2008.

[4] T. Okano, L. L. Berank, and T. Hidaka. Relations among interaural cross-correlation coefficient $(\mathrm{IACC_E})$, lateral fraction $(\mathrm{LF_E})$, and apparent source width (ASW) in concert halls. *J. Acoust. Soc. Am.*, 105(1):255–265, July 1998.

[5] M. Park and B. Rafaely. Sound-field analysis by plane-wave decomposition using spherical microphone array. *J. Acoust. Soc. Am.*, 118(5):3094–3103, November 2005.

[6] B. Rafaely, B. Weiss, and E. Bachmat. Spatial aliasing in spherical microphone arrays. *IEEE Trans. Sig. Proc.*, 55(3):1003–1010, March 2007.

[7] B. Rafaely and A. Avni. Inter-aural cross correlation in a sound field represented by spherical harmonics (a). *J. Acoust. Soc. Am.*, 125(4):2545, April 2009.

[8] J. R. Driscoll and D. M. Healy Jr. Computing fourier transforms and convolutions on the 2-sphere. *Adv. App. Math.*, 15:202–250, 1994.

[9] E. G. Williams. *Fourier Acoustics. Sound radiation and nearfield acoustic holography*. Academic Press, New York, 1999.

[10] G. Arfken and H. J. Weber. *Mathematical methods for physicists*. Academic Press, San Diego, 5 edition, 2001.

[11] B. Rafaely. Plane-wave decomposition of the pressure on a sphere by spherical convolution. *J. Acoust. Soc. Am.*, 116(4):2149–2157, October 2004.

[12] B. Rafaely. Analysis and design of spherical microphone arrays. *IEEE Trans. Speech Audio Proc.*, 13(1):135–143, January 2005.

[13] ISO 3382:1997, International Organization for Standardization, Geneva, Switzerland. *Acoustics – measurement of the reverberation time of rooms with reference to other acoustical parameters*, 1997.

[14] Richard K. Cook, R. V. Waterhouse, R. D. Berendt, Seymour Edelman, and Jr. M. C. Thompson. Measurement of correlation coefficients in reverberant sound fields. *J. Acoust. Soc. Am.*, 27(6):1072–1077, November 1955.

[15] H. Net'lisse and J. Nicolas. Characterization of a diffuse field in a reverberant room. *J. Acoust. Soc. Am.*, 101(6):3517–3524, June 1997.

[16] Heinrich Kuttruff. *Room acoustics*. Taylor & Francis, New York, 4 edition, 2000.