Pätynen, Jukka

**Proportion of effects by head-related transfer function and receiver position variation to interaural cross-correlation values**

*Published in:*
Journal of the Acoustical Society of America

**DOI:**
10.1121/1.4985611

Published: 01/06/2017

**Document Version**
Publisher's PDF, also known as Version of record

*Please cite the original version:*
Proportion of effects by head-related transfer function and receiver position variation to interaural cross-correlation values

Jukka Pätynen

Citation: The Journal of the Acoustical Society of America 141, EL579 (2017); doi: 10.1121/1.4985611
View online: https://doi.org/10.1121/1.4985611
View Table of Contents: http://asa.scitation.org/toc/jas/141/6
Published by the Acoustical Society of America

Articles you may be interested in

No correlation between headphone frequency response and retail price
The Journal of the Acoustical Society of America 141, EL526 (2017); 10.1121/1.4984044

Compressive time delay estimation off the grid
The Journal of the Acoustical Society of America 141, EL585 (2017); 10.1121/1.4985612

Preserving spatial perception in rooms using direct-sound driven dynamic range compression
The Journal of the Acoustical Society of America 141, 4556 (2017); 10.1121/1.4984040

A mobility based vibroacoustic energy transmission simulation into an enclosure through a double-wall panel
The Journal of the Acoustical Society of America 141, EL598 (2017); 10.1121/1.4989729

Stability of sonic boom metrics regarding signature distortions from atmospheric turbulence
The Journal of the Acoustical Society of America 141, EL592 (2017); 10.1121/1.4986209

Room acoustic modal analysis using Bayesian inference
The Journal of the Acoustical Society of America 141, 4480 (2017); 10.1121/1.4983301
Proportion of effects by head-related transfer function and receiver position variation to interaural cross-correlation values

Jukka Pätnyen
Department of Computer Science, Aalto University School of Science, FI-00076 Aalto, Finland
Jukka.Patynen@aalto.fi

Abstract: Interaural cross-correlation coefficient (IACC) is an objective room-acoustic parameter associated with the spatial impression and subjective quality of the acoustics. This study investigates the proportion of effects on the IACC variance simultaneously by (1) varied head-related transfer function (HRTF) and (2) locally varied receiver positions. The early IACC80 values are estimated from the binaural room simulations in eight room-acoustic conditions varying in room shape, side wall absorption, and receiver distance. Analysis across the combinations of HRTF and receiver positions indicates that below 2 kHz, around two-thirds of total IACC variance is explained by the receiver displacement, and at higher frequencies the HRTF variations dominate the IACC variance.

© 2017 Acoustical Society of America

1. Introduction

Binaural features of spatial sound are essential in room acoustics perception. While the early reflections arriving from lateral angles have been identified as the principal source for the spatial impression1,2 and overall room acoustic quality,3,4 perception of sound fields in rooms is ultimately based on the binaural cues in the two signals entering the listener’s ears.5 The standardized objective metric of interaural cross-correlation coefficient (IACC)6 is a popular approach for estimating the degree of interaural coherence7 and the perceived spatial impression3,6 from binaural room impulse responses. However, acoustic measurements can rarely be perfectly reproduced. The uncertainties in binaural room acoustic parameters have so far been investigated in several studies,8–11 which concentrate mostly on the sensitivity to measuring position displacement or receiver orientation. Such studies also consider the parameter sensitivity to only one varied factor at a time. Yet another source of uncertainty in binaural analysis methods emerges from the anthropometric differences between human heads’ shapes. The incident sound becomes altered due to individually varying head-related transfer functions (HRTFs). Although the effect of the head and ear geometry to HRTFs has been studied in the particular field of literature,12 the variation of the IACC values due to different HRTFs has not been evaluated closely. Moreover, the proportion of the IACC metric sensitivity to the HRTF and receiver position variations remains uncharted. This study contributes, first, by investigating the influence of different HRTFs to the early interaural cross-correlation coefficient (IACC80) values, and second, by estimating simultaneously the proportion of effects by the HRTF variation and locally selected receiver position to the obtained IACC values. Since binaural room measurements are often applied in reproducing room acoustics, obtained magnitudes of effect are important for evaluating the differences between signals received by individual listeners.

The basis for spatial hearing in rooms is the encoding of single acoustic events in different incident angles into interaural time (ITD) and level (ILD) differences at the two spatially separated ears. The arrival of successive copies of the emitted sound signal follows the pattern of room reflections in different angles, and this induces fluctuations in the ITD cues14 and binaural coherence.15 While the lateral energy fraction (LEF) is a measure for the proportion of lateral energy in the sound field,1 lateral reflections are only instrumental in achieving a desired binaural incoherence. If the decay of the room response is smooth without echo-like artefacts, the spatial perception employs binaural cues instead of observing physical room reflections as distinct events. In this respect, IACC is associated with the auditory perception more closely than LEF parameter.
Earlier research has shown early IACC, calculated from a 0 to 80 ms time window, to correlate inversely with the spatial impression and the apparent source width, but also with the overall subjective room-acoustic preference. Since the differences between the shapes of the ear canals, pinnae, and the torso affect how the incident spatial sound enters the ears, the individual HRTFs influence the exact IACC occurring in human listeners or artificial head receivers. Most research interest towards IACC properties include the effects of level, incident angle, and number of early reflections. In more complex settings, IACC has been shown to vary strongly across a concert hall audience area. The proposed just-noticeable difference (JND) for IACC is 0.075, and measurements with several source-receiver pairs in an existing hall have shown IACC variation ranging over several JNDS.

2. Method

In order to investigate the variation in IACC80, the sound propagation via direct and early reflected paths was modeled with the image-source method. Although the pure image-source method does not take into account wave-based phenomena, the image-source method was selected due to the generalizability and replicability. Advantageously, the image-source approach yields a parametric representation of room-acoustic events for subsequent binaural synthesis through HRTFs.

The interaural cross-correlation coefficient is defined as the absolute maximum of time-limited normalized cross-correlation function between the binaural room impulse response at left ($x_l$) and right ($x_r$) ears,

$$IACC_{t_1} = \max \left| \frac{\int_{t_0}^{t_1} x_l(t)x_r(t+\tau)dt}{\sqrt{\int_{t_0}^{t_1} x_l^2(t)dt \int_{t_0}^{t_1} x_r^2(t)dt}} \right|, \quad |\tau| < 1 \text{ [ms]}.$$  

In this study, IACC80 denotes the integration limit $t_1 = 80$ ms for the early part of the room impulse response beginning from the initial direct sound $t_0$.

2.1 Simulation

Room-acoustic events were simulated in two room geometries. The dimensions of the first, rectangular, room in the length, width, and height axes were $48 \times 20 \times 17$ m, respectively. These dimensions correspond roughly to a typical rectangular concert hall. The raised stage floor was at 1.5 m height, and circling balconies were at 6.5 m height from the floor. Seat rows with 0.9 m height were modeled as an array of consecutive blocks. The second room was modified from the first room into a generic fan shape by skewing the overall geometry so that the side walls splayed into 14.5° angles. Balconies were omitted from the fan shape room model. The surface reflection coefficients were derived from listed values for typical surface materials. In addition to those values presented in Table 1, simulations were repeated with the side walls having twice their respective absorption coefficient. These conditions are referred to as low and high absorption. The geometries and applied materials result in estimated Sabine reverberation times from 1.5 to 2.2 s.

Seven point sources were modeled to span an area corresponding roughly to real and simulated orchestras. The height of the sources was 1 m above the stage floor. Receivers in both rooms included two clusters of positions, each having a 5 x 5 grid of receivers in a 2 x 2 m² area. The centers of the clusters were 2 m off the center line. To summarize, two room shapes, two absorption cases, and two receiver cluster distances yield eight different room-acoustic conditions. The detailed geometry of the simulations is shown in Fig. 1. By default, each receiver was facing towards the center of the source area. Additional simulation runs were repeated, first, with a receiver

<table>
<thead>
<tr>
<th>Surface</th>
<th>Octave band [Hz]</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>125</td>
<td>250</td>
</tr>
<tr>
<td>Front, back wall</td>
<td>0.85</td>
<td>0.88</td>
</tr>
<tr>
<td>Side walls</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>Ceiling, floor</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Seating area</td>
<td>0.63</td>
<td>0.51</td>
</tr>
</tbody>
</table>
orientation along the main longitudinal axis of the room and, second, facing each source separately. The largest angular difference between the center-facing and room-axis aligned receiver was 13°, and 31° between the center-facing and source-facing receivers.

The early room-acoustic events were obtained by simulations of the room models in Odeon software (Odeon A/S, Denmark) up to tenth order of reflection. Based on lists of room-acoustic events imported in parametric format to MATLAB, the binaural room impulse response was synthesized in the complex frequency domain for a convenient application of the HRTF. In essence, the HRTFs for the incident directions of the acoustic events were superpositioned according to the respective levels and time-of-arrivals. The continuous frequency-domain reflection coefficients were obtained with cubic interpolation between octave band data, and the coefficients for frequencies $f=0$ and the Nyquist frequency were fixed to the nearest available value. Applied HRTFs were derived from the Center for Image Processing and Integrated Computing (CIPIC) database that consists of 45 subjects, including a KEMAR artificial head. The angular resolution in CIPIC HRTF measurements is approximately 5°. Using the nearest available CIPIC measurement angle to the actual incident angle resulted in the mean error of approximately 2.4° within the region of CIPIC measurements.

2.2 Analysis

All IACC$_{80}$ values were first estimated for each source-receiver pair in octave bands, and finally averaged over the seven source positions. The grand mean of values averaged over a cluster of receivers in a single room condition is statistically modeled as a normally distributed random variable:

$$\text{IACC} \sim N_{\text{IACC}}(\mu_{\text{IACC}}, \sigma^2_{\text{IACC}}).$$

(2)

In this study, IACC$_{80}$ values in one room condition are assumed to be influenced by the receiver position displacement and receiver HRTF. The effects by dependent variables of receiver position (factor P) and HRTF (factor H) are modeled as independent normally distributed values with zero mean, and the residual term is denoted with R. The corresponding notation for combinations of ($k=1,\ldots,K$) receiver positions, ($m=1\ldots M$) HRTFs and residuals is formulated as

$$\text{IACC} = c_{k,m} = \mu_{\text{IACC}} + p_k + h_m + r_{k,m},$$

(3)

where $p_k \sim N(0, \sigma^2_P)$, $h_m \sim N(0, \sigma^2_H)$, and $r_{k,m} \sim N(0, \sigma^2_R)$.

Given that the sum of the two normal distributions with residual follows a distribution $c_{k,m} \sim N(\mu_{\text{IACC}}, \sigma^2_P + \sigma^2_H + \sigma^2_R)$, the unknown contribution of one factor can be estimated as the subtraction of the means over the other factor from the dataset. For example, the data variation in receiver position $k$ after subtracting the average over all HRTFs yields $c_k = c_{k,m} - \sum_{m=1}^M c_{k,m}/M$. The residual is obtained by subtracting both P and H factor means. Figure 2 demonstrates this approach for one room-acoustic condition with IACC$_{80}$ values averaged over 500–2000 Hz octave bands.

The current balanced experiment design leads to the statistical two-way main effects model for two random effects. The principal interest is in the proportion of magnitudes of the main effects, similarly to the explained variance in regression analysis. The effect of magnitudes can be calculated through two-way analysis of variance. The analysis employs $\omega^2$ statistic for the magnitude of effect (Ref. 21, p. 446). $\omega^2$ is given by the ratios of mean squares (MS) as

$$\omega^2 = \frac{\text{MS} \text{effect}}{\text{MS} \text{effect} + \text{MS} \text{error}}.$$
where

\[
\omega_P^2 = \frac{\hat{\sigma}_P^2}{\hat{\sigma}_{\text{total}}^2} = \frac{\hat{\sigma}_P^2}{\hat{\sigma}_P^2 + \hat{\sigma}_H^2 + \text{MSR}},
\]

for factor P and correspondingly for factor H. If repeated measures were feasible, the residual variation MSR could be further partitioned into the interaction effect P \(* H\) and residual. As the repeated simulations for each position/HRTF combination do not provide additional information, the possible interaction cannot be resolved. Therefore, the residual term MSR takes over the entire remaining variation in the main effects model.

The validity of the statistical model regarding the equal variance is supported by Levene’s test, which suggests that the variances of simulated IACC80 over 45 HRTFs are homogeneous at 25 receiver positions at \(\alpha = 0.05\) confidence level in each condition. The effect by different HRTFs does not deviate significantly from normal distribution, according to Kolmogorov-Smirnov normality test. Also, the variances over the receiver positions within a cluster are homogeneous with each HRTF apart from one condition. The effect of varying receiver position within one cluster also followed the normal distribution with all HRTF in all conditions. Based on these findings, the assumption of normally distributed factors contributing to IACC80 values is reasonable.

The possible bias to IACC80 values by particular HRTFs, or by particular position within a cluster, is explored with correlation tests. This approach also provides information on whether the receiver displacements from the center position of the cluster have a constant effect in different room conditions.

3. Results

The average IACC80 for eight room conditions range between 0.40 and 0.60 (0.5–2 kHz). These values span over 2.5 JNDs, and the range corresponds to IACC values measured in real concert halls.\(^{17}\) The overall IACC80 standard deviation (SD) varies moderately between eight conditions, and SD remains in the same magnitude. The overall SD due to the varied receiver position and HRTF to IACC80 value is approximately 0.04 at the mid-frequencies. Although the data are insufficient for a conclusive correlation analysis, conditions with high IACC80 have a tendency for higher total SD (\(\rho = 0.49\)).

Selection of the local receiver position accounts for 71% of the overall variance, while the effect by different HRTFs is approximately 26% across eight room conditions. The remaining variation (4.2%) is attributed to the possible interaction between position/HRTF combinations, or random effects. The contributions by position variation to the total variance range between 65% and 77%. At single octave bands, the grand mean of IACC80 over room conditions decreases towards high frequencies, but the overall SD does not decrease in a similar proportion. Instead, the magnitudes of effect shift prominently for the HRTF variation (see lower half in Table 2). The \(\omega^2\) statistic for factor H exceeds that of factor P at 2 kHz and above. Also, the magnitude of effect by the
residual term increases substantially. The effect of orienting the receiver along the room axis has only a marginal effect on the results. The receiver orientation separately towards each source cause a minor increase in the overall effect by HRTF variation (see bottom rows in Table 2).

The possible bias by HRTFs to IACC80 values was investigated by Pearson's correlation coefficients with Bonferroni correction. Sets of IACC80 values obtained with 45 HRTFs correlate significantly ($a = 0.05$) between receiver positions. Similarly, sets of position displacement-related IACC 80 values correlate between individual HRTFs. IACC80 values from different HRTFs were correlated also against anthropometric measures collected from the HRTF subjects.20 However, none of the numerous factors alone, including the maximum ITD or ILD value, appear to correlate at a significant level with IACC80.

4. Discussion and concluding remarks

In the room-acoustic simulations, the variation of IACC80 was mostly influenced by the varying receiver position, and the relative effect by the HRTF varied between one-quarter to one-half of that of the position selection. The overall results indicated that over three mid-frequency octave bands, the estimation of IACC80 values is not critically sensitive to the naturally occurring variation in the HRTF. Considering the just noticeable difference (JND) of 0.075,6 the uncertainty due to HRTF variation to IACC80 remains below the JND threshold. At high frequencies, the magnitude of effect by HRTF became prominent as the absolute SD of position displacement decreased. The increase in SD by the HRTF factor can be potentially attributed to the complex small-scale differences between ears’ shapes.20 Further analysis shows that the relative response difference between frontal and lateral incidence begins to vary strongly above 3 kHz across individual HRTF measurements. This outcome offers additional evidence for the increased effect by HRTF to IACC80 at high frequencies in spatial room-acoustic settings.

The correlation analysis showed that certain HRTFs yield systematically different levels of interaural coherence. Consequently, room measurements conducted with particular dummy heads are expected to produce consistently higher or lower IACC estimates. Furthermore, this finding signifies likely that human listeners may experience characteristically higher or lower individual degree of binaural coherence. The perceptual significance of individually varying global bias in interaural coherence remains an open question for future studies. Also, resolving the cause for notable residual variance above 2 kHz stands to be solved with a different research approach.

<table>
<thead>
<tr>
<th>Room condition</th>
<th>IACC80</th>
<th>SD [$\times 10^{-3}$]</th>
<th>$\omega^2$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{c}$</td>
<td>$\bar{c}$ P H R</td>
<td>P H R</td>
</tr>
<tr>
<td>Rectangular</td>
<td>Low</td>
<td>0.40 2.80 2.25 1.00 0.76 72.79 19.6 7.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.41 4.08 3.32 2.05 0.69 69.53 27.6 2.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.43 2.79 2.19 1.14 0.75 69.08 23.5 7.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.45 4.19 3.44 2.09 0.66 70.38 27.0 2.6</td>
<td></td>
</tr>
<tr>
<td>Fan</td>
<td>Low</td>
<td>0.54 3.28 2.80 1.29 0.65 77.09 18.9 4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.58 4.60 3.66 2.50 0.70 66.14 31.5 2.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.56 3.19 2.68 1.31 0.65 75.12 20.6 4.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.60 4.65 3.67 2.60 0.69 64.75 33.0 2.3</td>
<td></td>
</tr>
<tr>
<td>0.5–2 kHz (Average over 8 conditions)</td>
<td>0.50 3.70 3.00 1.75 0.70 70.61 25.2 4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 kHz (8 conditions)</td>
<td>0.59 5.06 4.21 2.38 0.79 71.53 25.8 2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 kHz (8 conditions)</td>
<td>0.48 4.95 3.68 2.38 1.25 62.98 29.6 7.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 kHz (8 conditions)</td>
<td>0.42 3.85 1.61 2.09 1.41 37.58 44.5 17.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 kHz (8 conditions)</td>
<td>0.36 5.06 2.36 2.72 1.71 39.47 46.8 13.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 kHz (8 conditions)</td>
<td>0.28 5.70 0.59 4.77 1.81 8.84 80.2 11.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5–2 kHz, → Length axis</td>
<td>0.50 3.69 3.07 1.66 0.68 72.80 23.2 4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5–2 kHz, → Sources</td>
<td>0.50 3.82 3.05 1.90 0.68 67.09 29.2 3.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overall, these findings suggest that practical estimation of IACC\textsubscript{80} is moderately influenced by the HRTF. Still, the parameter is generally insensitive to changes in both receiver position and HRTF, and that the interaural coherence is a reliable measure regardless small measurement variations in various room acoustic condition or investigated frequency band.

Acknowledgments

The research has been funded by Academy of Finland, Project No. [289300]. The calculations were performed using resources within the Aalto University Science-IT project.

References and links