Influence of auralization methodology of musical pieces in the subjective evaluation

LAMBERTO TRONCHIN, RYOTA SHIMOKURA, VALERIO TARABUSI and ILARIA DURVILLI

DIENCA – CIARM, University of Bologna, Italy

http://www.ciarm.ing.unibo.it

Abstract: -
Generally acoustical qualities of concert halls and opera houses are evaluated in base of BIRs (binaural impulse responses) measured in them, and the anechoic music convoluted with BIR realizes the virtual sound as it were played in the sound field. Interaural Cross Correlation Coefficient (IACC) calculated from BIR is one of parameter to evaluate the spaciousness of hall. However the value of IACC is changed by the convolution technique in accordance with the kind of musical motif. For example, one binaural impulse response measured in the concert hall in Tsuyama (Japan) presents 0.16 for the value of IACC, while the IACC of the symphony "Royal Pavane" (Orlando Gibbons) convoluted with it is 0.39, and the IACC of the symphony "Symphonietta n14 the forth movement" (Malcolm Arnold) convoluted with it is 0.32. The aim of this study is to predict the value of IACC after convolution using the acoustical characteristic extracted from the musical motif

Key words: - Auralization, Binaural recordings, Stereo-dipole, Listening room, Room acoustics
.

INTERAURAL CROSSCORRELATION

When sound is propagated from a sound source, the signals received at left and right ears of a listener are different. Interaural cross-correlation function (*IACF*) represents the interdependence between left (right) signal at the origin and the right (left) signal at a delay within 1 *ms*. *IACC* is one maximum value in the *IACF* like:

$$IACC = \left| \frac{\int_{-T}^{T} p_l'(t) p_r'(t+\tau) dt}{\sqrt{\int_{-T}^{T} p_l'^2(t) dt \int_{-T}^{T} p_r'^2(t) dt}} \right|_{\max} \quad |\tau|<1 \quad (1)$$

where $2T$ is the integral interval, τ is the time delay, and $p'(t)$ is obtained after passing through the A-weighting filter, which corresponds approximately to the sensitivity of the ear.

ACOUSTIC EVALUATION OF SIGNALS

Acoustical parameters of ACF

For the blending of sound field and performance, Ando proposed τ_i and τ_e to determine temporal acoustical characteristics of sound sources [1]. The τ_i and τ_e are extracted from a normalized autocorrelation function (*ACF*) as shown in

$$\phi(\tau) = \frac{\Phi(\tau)}{\Phi(0)} \quad (2)$$

where:

$$\Phi(\tau) = \frac{1}{2T} \int_{-T}^{T} p'(t) p'(t+\tau) dt \quad (3)$$

In (3) 2T is the integral interval, τ is the time delay, and *p'(t)* is an original acoustical signal after passing through the A-weighting filter; τ_i is a delay time of the first positive peak, and τ_e

is an effective duration of *ACF*, defined by the delay time at which the envelope of the normalized *ACF* becomes and then remains smaller than 0.1. Value of τ_1 indicates pitch of the signal, and value of τ_e is repetitive feature that corresponds to kinds of musical instrument, tempo of the motif and pattern of playing like *legato* or *staccato*.

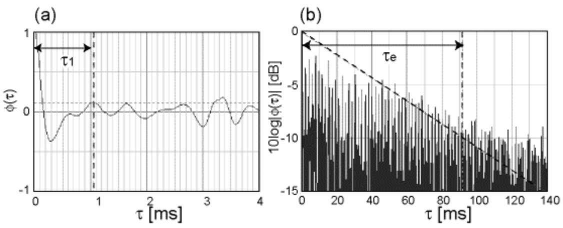

Fig. 1 - Definition of (a) τ_1 and (b) τ_e in normalized ACF

Practical use of ACF for music

During a performance of music, the acoustical characteristics (e.g. pitch and tempo) are varied as a function of time. Running ACF is necessary to observe the fluc-tuation of acoustical characteristic [1]. Running ACF is defined by:

$$\phi_p(\tau;t,T) = \frac{\Phi_p(\tau;t,T)}{\left[\Phi_p(0;t,T)\Phi_p(0;\tau+t,T)\right]^{1/2}} \quad (4)$$

where:

$$\Phi_p(\tau;t,T) = \frac{1}{2T}\int_{t-T}^{t+T} p'(s)p'(s+\tau)ds \quad (5)$$

After passing through the A-weighting filter, normalized ACF of $p'(t)$ is calculated in the range of integral interval $2T$ that is sliding along the duration of motif. Figure 2 explains the structure of running ACF.

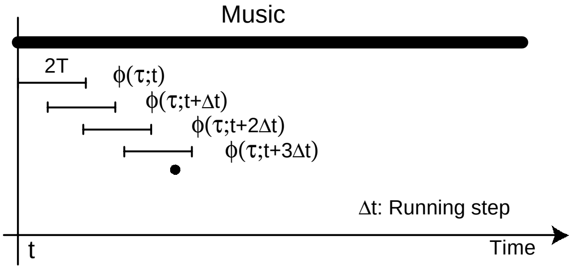

Fig. 2 Running ACF of musical motif

τ_1 and τ_e extracted from each running ACF indicate varied acoustical characteristic as a function of time. This section shows the examples of running acoustical parameters. Two kinds of rhythmic musical melody, "Melody A" and "Melody B" are used (see Figure 3). Melody A is played by trumpet and piano, and Melody B is played by organ and piano. Totally 4 kinds of anechoic musical motif are generated by MIDI.

Fig. 3 - Scores of (a) melody A and (b) melody B

Figure 4 shows the τ_1 and τ_e of these motifs as a function of time when the integral interval ($2T$) is 1 s and the running step is 0.1 s. It can be observed that τ_1 is affected by the difference of musical instruments and τ_e is mainly affected by the difference of melody.

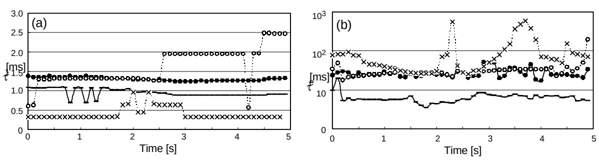

Fig. 4 - Acoustical parameters for 5 s. (a) τ_1 and (b) τ_e. Different symbols indicate different musical motifs: (-): Melody A by trumpet; (●): Melody A by piano; (×): Melody B by organ; (○): Melody B by piano.

IACC OF VIRTUAL SOUND

Convolution technique and IACC

A musical motif after convolution with BIR of a hall creates a virtual sound as if the music were played in the hall. For the sake of convenience, IACC of BIR is termed by "$IACC_{IR}$" and the IACC of the virtual music is termed by "$IACC_{AC}$" (After Convolution).
For example, when a white noise (monaural) is convoluted with one BIR, the $IACC_{AC}$ agrees with the $IACC_{IR}$ of the all-passed impulse response. And when a pure tone (monaural) is convoluted with one BIR, the $IACC_{AC}$ is always 1 independently from the frequency of the pure tone and from the difference of the $IACC_{IR}$

$IACC_{AC}$ in case of musical motif

This section shows examples of $IACC_{AC}$ in the case that a sound signal is a musical motif.

BIR: We used 6 kinds of BIR that are measured in Kirishima International Concert Hall (Japan), Tsuyama Concert Hall (Japan) and ancient theatre of Delphi (Greece). $IACC_{IR}$ and $IACC_{E3}$ of these BIR are listed in Table 1. Since these values of $IACC_{IR}$ are calculated by Equation (1), they take the A-weighting filtering into account.

BIR	$IACC_{IR}$	$IACC_{E3}$
Kirishima1	0.07	0.19
Kirishima2	0.23	0.29
Kirishima3	0.32	0.38
Tsuyama1	0.50	0.70
Tsuyama2	0.68	0.80
Delphi1	0.86	0.90

Table 1 $IACC_{IR}$ and $IACC_{E3}$

Musical motif: 8 kinds of anechoic musical motifs (monaural) used in this examination are listed in Table 2. The duration is 10 s. All signals are generated by MIDI. To find the temporal characteristics of the musical motifs, running ACF of each motif is calculated and τ_1 and τ_e as a function of time are extracted like Figure 4. As a result, 100 values of either τ_1 or τ_e are obtained after the running ACF calculation (integral interval: 1 s; running step: 0.1 s). Since both τ_1 and τ_e change dynamically as a function of time, unique representative value to express the characteristic of each motif is determine by the 50 % probability of cumulative frequency when the 100 values of either τ_1 or τ_e are converted into a histogram. These values are also shown in Table 2.

Musical motif	τ_1 (50%)	τ_e (50%)
Melody A by glockenspiel	0.56	423.4
Melody A by harpsichord	0.96	145.4
Melody A by piano	1.33	246.5
Melody A by trumpet	0.88	54.9
Melody B by harpsichord	0.81	208.0
Melody B by organ	0.46	526.7
Melody B by piano	1.94	308.8
Melody B by strings-ensemble	1.26	79.9

Table 2 Anechoic musical motifs and their τ_1 (50%) [ms] and τ_e (50%) [ms].

Convolution: As a result of convolution of 6 BIRs and 8 musical motifs, 48 kinds of virtual music can be created. Since these virtual sounds last for 10 s, running IACF is calculated with same manner as running ACF. The integral interval and the running step are 1 s and 0.1 s respectively. Figure 5 shows the examples of $IACC_{AC}$ as results of the running IACF.

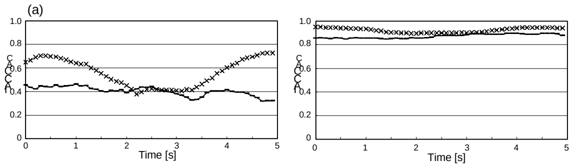

Fig. 5 – $IACC_{AC}$ for 5 s. BIRs that followed the convolution is (a) Kirishima3 and (b) Delphi1. Different symbols indicate different musical motifs: (-): Melody A by trumpet; (×): Melody B by organ.

Estimation of $IACC_{AC}$

100 values of $IACC_{AC}$ can be calculated for each virtual sound. These 100 values are converted into histogram, and the represented value of $IACC_{AC}$ for each virtual sound is determined by 75 % probability of the cumulative frequency [2]. Figure 6 shows the relationship between $IACC_{AC}$(75%) and $IACC_{IR}$ according to each musical motif. These relationships are fitted by power regression curves, such that:

$$IACC_{AC} = IACC_{IR}{}^{\alpha} \qquad (6)$$

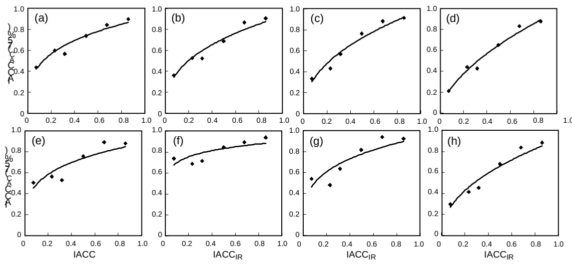

Fig. 6 – $IACC_{AC}$ (75%) of virtual sound as a function of $IACC_{IR}$. (a) Melody A by glockenspiel, (b) Melody A by harpsichord, (c) Melody A by piano, (d) Melody A by trumpet, (e) Melody B by harpsichord, (f) Melody B by organ, (g) Melody B by piano and (h) Melody B by strings-ensemble.

Table 3 shows the exponent α and the correlation coefficient for the fitting R^2. Although $IACC_{AC}$ could be estimated exactly by $IACC_{IR}$ for the most musical motifs, "Melody B by organ" is unsuitable for the estimation by Equation (6). And it has a tendency that the estimated $IACC_{AC}$ of Melody A is more accurate than that of Melody B. Since Melody B is composed by a series of chord (see Figure 3 (b)), $IACC_{AC}$ is changed largely according to the moment sounding or maintaining the chords as shown in Figure 5 (a). Melodies including *legato* and *tenuto* playing are difficult to determine the one represented value of $IACC_{AC}$, and the estimation of $IACC_{AC}$ by Equation (6) tends to lead the errors.

Musical motif	α	R^2
Melody A by glockenspiel	0.38	0.90
Melody A by harpsichord	0.46	0.95
Melody A by piano	0.47	0.96
Melody A by trumpet	0.61	0.97
Melody B by harpsichord	0.36	0.77
Melody B by organ	0.19	0.38
Melody B by piano	0.33	0.80
Melody B by strings-ensemble	0.57	0.96

Table 3 Exponent α in Equation (6) and correlation coefficient for the fitting R^2

Figure 7 shows α in Equation (6) as a function of τ_e (50%) of musical motifs. As a result of asymptotic curve regression, the relationship between α and τ_e (50%) is expressed by:

$$\alpha = \frac{21.7}{\tau_e(50\%)^{0.6} + 21.7} \qquad (7)$$

When τ_e of musical motif is short, the value of α approaches to 1. For the signal of white nose, the τ_e is theoretically 0 and α becomes 1. On the other hand, when τ_e of musical motif is long, the value of α approaches to 0. This means that the α value of pure tones ($\tau_e = \infty$) becomes 0. The values of α calculated by Equation (7) are highly correlated with the measured values ($R^2 = 0.82$).

The relationship between $IACC_{AC}$ (75%) and $IACC_{IR}$ in this theory is illustrated in Figure 8 (a). When the value of α is 1, $IACC_{AC}$ is equal to $IACC_{IR}$, which indicates that the convolution with motif similar to noise do not modify the interaural correlation between left and right impulse responses. When the value of α is 0, $IACC_{AC}$ is always equal to 1, which indicates that a pure tone motif generated in a hall is not affected by the sound field in terms of IACC and the value of IACC is always 1.

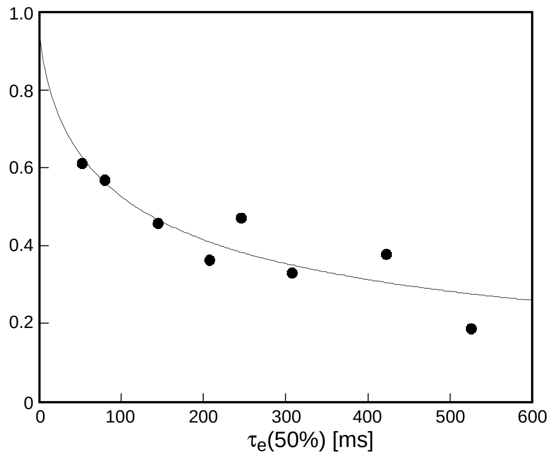

Fig. 7 - – Value of α as a function of τ_e (50%) of musical motif

Using Equation (6) and (7), the estimated value of $IACC_{AC}$ (75%) is calculated by τ_e (50%) of the musical motifs and $IACC_{IR}$. Figure 8 (b) shows the estimated $IACC_{AC}$ as a function of $IACC_{AC}$. The τ_e (50%) of musical motif convoluted with $IACC_{IR}$ explain $IACC_{AC}$ (75%) accurately.

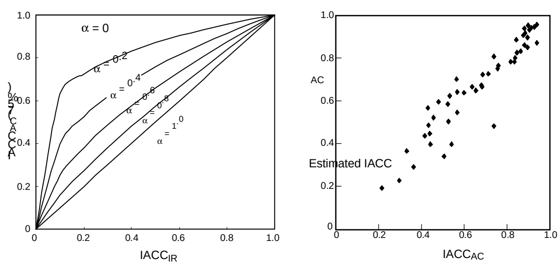

Fig. 8 - Theoretical relationship between $IACC_{AC}$(75%) and $IACC_{IR}$ according to α. (b) Relationship between $IACC_{AC}$ and estimated $IACC_{AC}$ by Equations (6) and (7). The correlation coefficient (R^2) among them is 0.94.

DISCUSSION

Acoustical measurement of concert halls and opera houses has been developed by methodology of obtaining impulse response (e.g. MLS and swept-sine signals) and new implements (e.g. dummy head and omni-directional speaker), and the objective acoustical characteristics of halls can be observed in measured binaural impulse responses. Usually the acoustical measurement is conducted under unoccupied condition. Hidaka and Beranek (2001) compared the acoustical qualities in the two situations with and without audience [3]. According to acoustical measurements, the next assignment is to assume the actual situation in halls where performers play music on stage and audience listens to it in seats.

It is important to recognize that listeners in halls do not enjoy the impulse response but prefer performances like music or songs. Although IACC calculated by BIR is a useful objective acoustical parameter relating to subjective diffuseness, the acoustical design of halls should be carried out assuming the performances that would be presented in the halls.

To evaluate concert halls or opera houses acoustically, Farina (2001) utilized IACC calculated with all passed impulse responses ($IACC_E$) [4], while Hidaka and Beranek (2000) utilized the average of IACCs calculated with 0.5, 1, and 2 kHz band passed impulse responses ($IACC_{E3}$) [5, 6]. As a result, Farina did not find a correlation between $IACC_E$ and subjective evaluation of halls, while Hidaka et al. described that $IACC_{E3}$ had a high correlation with a rank order of halls' acoustical qualities. Although many reasons might explain this contradiction (e.g. different methodologies of evaluation or different subjects), it is a fundamental problem that the values of IACC treated by them are calculated by BIR of different frequency ranges. Hidaka and Beranek used the $IACC_{E3}$ considering about the blending of sound field and performance, because spectral energy of symphony distributes mainly around 0.5-2 kHz. It is important to evaluate the sound quality in halls based on the music performed in them.

It was found that IACC calculated by virtual sound after convolution with a BIR exceeds the IACC of the BIR itself, and the degree of excess is dependent on the value of τ_e that the autocorrelation function of anechoic musical motif presents. Although the IACC after convolution or effective value of IACC for subjective spatial impression can be evaluated by the frequency components of motif [7, 8], the autocorrelation function of anechoic musical motif also presents important factors to realize the blending of sound field and performance. When τ_e of musical performances in a theatre is short (e.g. trumpet), architectural plans in order to lowering $IACC_{IR}$ of the theatre are effective to make the practical $IACC_{AC}$ lower. On the other hand, τ_e of musical performance in a theatre is long (e.g. organ), the efforts to make $IACC_{IR}$ lower are less meaningful. For example, when a concert of organ is assumed to be performed, the values of $IACC_{AC}$ at most theatres exceed 0.7 independently from $IACC_{IR}$ of the theatres. In this case, acoustical planers should give priority to improve the other acoustical parameters for subjective preference.

CONCLUSIONS AND FUTURE WORK

This preliminary paper reported on the steps taken for setting up a recording/measurement and reproduction/simulation system capable of recreating a realistic reconstruction of the three- dimensional soundfield inside an existing concert hall.

The method can be applied either to multichannel "realtime" recordings, or to synthetically simulated sound samples obtained by convolution of anechoic music with measured impulse responses. These signals are replayed inside a special listening room, equipped with two integrated reproduction chains: a dual Stereo Dipole for "transaural" presentation of binaural signals, and an advanced Ambisonics decoder for periphonic (3D) presentation of B-format signals.

The two systems can be operated separately or simultaneously, provided that, in the latter case, a proper correction for the gain and for the processing delay of the two systems is applied.

Even if this paper is specially focused on the experimental equipment, the measured impulse responses have been utilised starting from the theoretical principles of Ambisonics and Stereo Dipole, implementing a method for the reproduction of sound characteristics of the auditoria. The reproduction of the sound samples employed for the listening tests was driven by specially-written software, which also enabled for the automatic collection of questionnaires.

The hybrid Ambiophonics system resulted in very natural and convincing listening experience, and consequently this opens the possibility to comparatively assess minor acoustical differences between halls very far each other, particularly with reference to the

spatial perception (envelopment, source imaging, depth, etc.) and to the temporal factors (Initial Time Delay Gap, difference between EDT and the subsequent reverberation time, etc.).
The encouraging results obtained by the first comparative experiments allows for the continuation of the research, which will move to the execution of several listening tests, aimed principally to defining the optimal listening conditions in terms of spatial attributes of the sound field and of system's frequency response, which are actually the less-explored perceptual aspects for a concert hall

REFERENCES

[1] Ando Y., "Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners", New York: Springer (1998).

[2] Morimoto M., Iida K., "A practical evaluation method of auditory source width in concert halls" J. Acoust. Soc. Jpn., Vol. 16, pp. 59-69 (1995).

[3] Hidaka T., Beranek L. L., "Relation of acoustical parameters with and without audience in concert halls and a simple method for simulating the occupied state" J. Acoust. Soc. Am., Vol. 109, pp. 1028-1042 (2001).

[4] Farina A., "Acoustical quality of theatres: correlations between experimental measurements and subjective evaluations" Applied Acoustics, Vol. 62, pp. 889-916 (2001).

[5] Hidaka T., Beranek L. L., "Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas" J. Acoust. Soc. Am., Vol. 107, pp. 368-383 (2000).

[6] ISO/DIS 3382, Acoustics: Measurement of the Reverberation Time with Reference to Other Acoustical Parameters. International Organization for Standardization (1997).

[7] Nakajima T., Yoshida J., Ando Y., "A simple method of calculating the interaural cross-correlation function for a sound field" J. Acoust. Soc. Am., Vol. 93, pp. 885-891 (1993).

[8] Hidaka T., Beranek L. L., Okano T., "Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls", J. Acoust. Soc. Am., Vol. 98, pp. 988-1007 (1995).