A Psycho-Acoustical Experiment Using a Stereo Dipole for Spatial Impression of Music Signals

Benedetto Nastasi – Delft University of Technology, The Netherlands – benedetto.nastasi@outlook.com
Massimiliano Manfren – University of Southampton, United Kingdom – m.manfren@soton.ac.uk
Francesca Merli – University of Bologna, Italy – francesca.merli8@unibo.it

Abstract

Acoustic performance of concert halls and opera houses is usually assessed by measuring the BIRs (Binaural Impulse Responses). Anechoic music convoluted with BIRs constitutes the virtual sound in the way it is played in the sound field, i.e. the room. From BIRs, the IACC (Inter-Aural Cross Correlation) can be computed. This parameter makes it possible to evaluate the spaciousness of the hall. However, the calculation of the IACC value is affected by the convolution technique used as well as the kind of musical motif. For example, in the same concert hall, the BIR provides three different IACC values in the case of three different motifs played in it. This study has conducted a psycho-acoustic experiment by using a virtual sound field representation produced by the stereo dipole technique in a listening room. In the experimental set-up there were two or four loudspeakers, corresponding to the single stereo-dipole or the dual stereo-dipole, respectively. By cancelling the cross-talk pathways (i.e. from left loudspeaker to right ear), the parallel sound presentation creates a 3D sound field for listeners sitting in the target point. The invert Kirkeby method was adopted to determine the inverse filters. Finally, the auralization technique with measured BIRs in theatres was utilized and the virtual sound field was generated in the Arlecchino listening room (Bologna, Italy), a low reverberation room equipped with an Ambisonic system. In the virtual sound field, the BIR was recorded again by the same dummy head used during the measurement in the theatres. The similarity between real and virtual sound fields was evaluated by comparing some acoustic parameters. The stereo-dipole technique demonstrates a good degree of accuracy of the sound field appearance. Moreover, the accuracy of the sound field appearance was analysed using two musical motifs and three musical instruments, comparing the values of the IACC calculated by echoic music with the virtual echoic music.

1. Introduction

In general, acoustical qualities of concert halls and opera houses are evaluated on the basis of BIRs (Binaural Impulse Responses) measured in them. The anechoic music convoluted with BIRs realizes the virtual sound as it was played in the sound field (Shimokura et al., 2011; Tronchin, 2012; Tronchin et al., 2020). The IACC (Interaural Cross Correlation) calculated from the BIRs represents one of the parameters to evaluate the spaciousness of halls. However, the value of the IACC is changed by the convolution technique according to the kind of musical motif used (Farina and Tronchin, 2000; Tronchin, 2013). For example, one BIR measured in the concert hall in Tsuyama (Japan) presents an IACC value of 0.16 whilst the IACC value of the symphony ”Royal Pavane” (Orlando Gibbons) convoluted with that BIRs is 0.39 and the IACC value of the symphony “Symphonietta n14 the fourth movement” (Malcolm Arnold) convoluted with that BIR is 0.32. A proper evaluation of acoustic quality is also important for other purposes, such as retrofitting design (Caniato et al., 2019; Fabbri et al., 2014; Fabbri and Tronchin, 2015; Mancini et al., 2017; Tronchin et al., 2014, 2016 and 2018; Tronchin and Fabbri, 2017). The aim of this study is to conduct the psycho-acoustical experiment by using a virtual sound field representation by a stereo dipole technique in a listening room. The stereo-dipole technique is realized using two or four loudspeakers (corresponding to single stereo-dipole or dual stereo-dipole). The calculation of proper inverse filters, by means of the Kirkeby method, makes it possible to reproduce the virtual sound field by means of the (dual) stereo dipole method, avoiding cross-talk paths. In this study,

Part of
Pernigotto, G., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). (2020). Building simulation applications BSA 2019. bu,press. https://doi.org/10.13124/9788860461766

the psychoacoustic experiments were conducted using the stereo-dipole technique, considering measured BIRs in theatres, virtually reproduced in the Arlecchino listening room located at the University of Bologna, Italy. This listening room was previously redecorated and equipped with the Ambisonic reproducing system. After having calculated the inverse filters, the virtual sound field of the rooms was obtained and the BIRs were recorded again by the same dummy head used during the measurement in the theatres. The similarity between the real and the virtual sound fields was evaluated by comparing some acoustical parameters (SPL, EDT, IACC etc) calculated using real and virtual BIRs. These acoustical parameters were compared and the results suggest that the stereo-dipole has a good degree of accuracy of the sound field appearance (Farina and Tronchin, 2005 and 2013; Tronchin and Coli, 2015). In this further study, we examine the accuracy of the sound field appearance using some musical motifs, comparing the values of the IACC calculated by “echoic music” and “virtual echoic music”. The echoic music indicates the anechoic musical signal convoluted with a BIR measured in a hall, while the virtual echoic music indicates the recorded echoic musical signal by means of the single stereo-dipole. Using MIDI, the anechoic musical signals are composed by considering two different melodies and three kinds of musical instruments. The IACC of a long continuous music motif was calculated by sliding the fixed integration interval along time (Tronchin and Coli, 2015).

2. Materials and Methods

2.1 IACC (Interaural Cross Correlation)

When sound is propagated from a sound source, the signals received at the 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 of 1ms. The IACC is one maximum value in the IACF. The IACC can be expressed by

$$IACC = \left| \frac{\int_{-T}^{T} p_l(t) p_r(t-\tau) dt}{\sqrt{\int_{-T}^{T} p_l^2(t) dt \int_{-T}^{T} p_r^2(t) dt}} \right|_{max} \quad (1)$$

where 2T is the integral interval, τ is the time delay, and $p_l(t)$ and $p_r(t)$ are signals obtained at left and right ears. In the case of evaluating a sound field, the $p_l(t)$ and $p_r(t)$ are corresponding to the impulse responses recorded at the left and right ear positions of a dummy head.

2.2 Acoustical Parameters Based on the IACC

In some research, the IACC has been modified based on the auditory nerve process or the acoustical characteristics of the musical performances in a hall.

Ando (1998) proposed that the IACF should be calculated with $p_{l'}(t)$ and $p_{r'}(t)$ which are obtained after passing through the A-weighting filter, which corresponds approximately to the sensitivity of the human ear (Tronchin, 2013). These calculative steps are based on the auditory-brain model for subjective responses. Unlike the spectral filtering, the $IACC_E$ is calculated in the limited integration time in 80 ms (Shimokura et al., 2011). The early part of the signal, such as EDT (Early Decay Time), is often evaluated to be important because most symphonic compositions include successive notes changing rapidly. The $IACC_{E3}$ is taken into account both for the spectral and temporal limitation. A signal is divided into one-octave spectral bands, and the values of the IACC are led from each band-passed signal limited to an integration time of 80 ms. The $IACC_{E3}$ is the IACC averaged with the results of the bands whose center frequencies are 0.5, 1, and 2 kHz because the spectral energy of a symphony distributes mainly around 0.5-2 kHz. To evaluate concert halls or musical instruments and opera houses acoustically, Farina utilized the $IACC_{E5}$, while Hidaka and Ando utilized the $IACC_{E3}$ (Ando, 1998; Farina, 2001; Farina et al., 1998). As a result, a correlation between $IACC_E$ and subjective evaluation of the halls was found by Farina. In another research, the $IACC_{E3}$ was found with a high correlation with a rank order of the halls' acoustical qualities by Hidaka et al. This controversial results can be explained by several reasons (e.g. diverse assessments or subjects). However, it is noteworthy that the IACC values chosen by them were computed referring to the BIRs of different frequency ranges.

2.3 Acoustical Characteristic of a Signal

Ando proposed τ_1 and τ_e to determine temporal acoustical characteristics of musical performances and adopted them for virtual sound reconstructions (Ando, 1998; Tronchin and Knight, 2016). A normalized autocorrelation function (ACF) was used to extract τ_1 and τ_e as follows:

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

Where

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

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_1 is a delay time of the first positive peak, and τ_e is an effective duration of the ACF, defined by the delay time where the envelope of the normalized ACF becomes and, then, remains smaller than 0.1 as depicted in Fig. 1. The value of τ_1 indicates the pitch of the signal, and the value of τ_e represents repetitive features, which corresponds to different kinds of musical instruments, tempo of the motif and the pattern of playing, such as legato or staccato. Generally, a fast tempo or a snap playing makes the τ_e shorter.

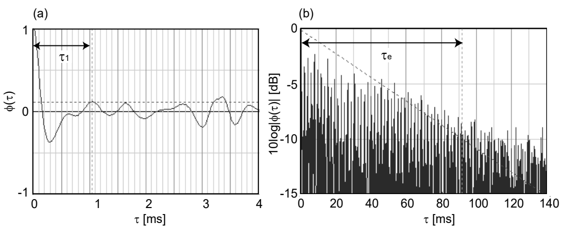


Fig. 1 – (a) Definition of τ 1 in normalized ACF; (b) Definition of τ e in normalized ACF in logarithm scale

During a performance of music, the acoustic characteristics (e.g. pitch and tempo) varies as a function of time. For observing the fluctuation of acoustical characteristics, it is necessary to run the ACF. The running ACF is defined by:

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

Where

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

After passing, the normalized ACF of p′(t) was calculated in the range of integral interval 2T once passed through the A-weighting filter. 2T slides along the duration of the motif. The structure of the running ACF is reported in Fig. 2.

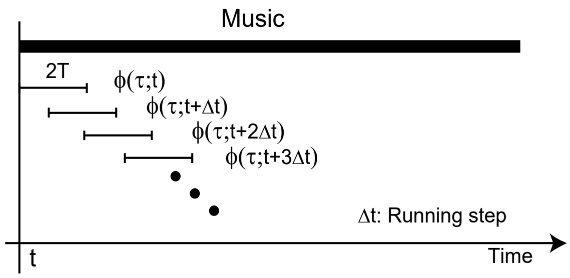


Fig. 2 – Running ACF of long musical motif

2.4 3D Sound Representation by Stereo-Dipole

2.4.1 BIR of the theatre

In this study, we utilize two kinds of BIR ("BIRn1" and "BIRn2") measured in the traditional Italian opera house, Teatro Nuovo in Spoleto (Italy). In the acoustical measurement, the sound source and the receivers are an omnidirectional speaker (LookLine dodecahedral configuration) and a dummy head (Sennheiser), respectively. The loudspeaker was located in the two positions of the stage; one near (BIRn1) and another one far (BIRn2) from the frontal edge of the stage, and the dummy head was located in one position in the middle of the stalls. The values of the IACC of all-passed BIRn1 and BIRn2 are 0.39 and 0.26, and $IACC_{E3}$ of the BIRs are 0.32 and 0.24, respectively.

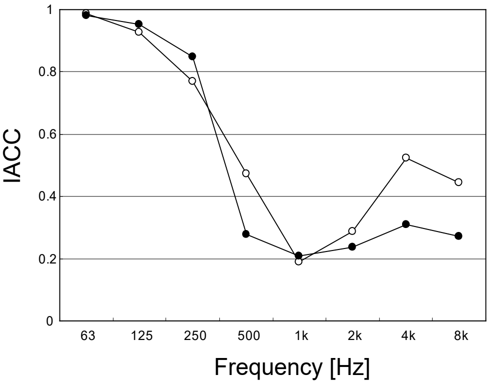


Fig. 3 – IACC of BIRn1 (○) and BIRn2 (●) as a function of frequency band

Fig. 3 shows the spectral characteristics of the IACC in these BIRs.

2.4.2 Anechoic musical motifs

Three kinds of an anechoic musical motifs were used: "Melody A by trumpet", "Melody A by piano", and "Melody B by organ". The scores of Melody A and Melody B are shown in Fig. 4. These anechoic musical motifs were generated by MIDI. The duration of the musical motifs is 30 s.

Fig. 4 – (a) Scores of Melody A (b) Scores of Melody B

To observe the acoustical characteristics of these anechoic musical signals, the running ACF calculation (see Equations (4) and (5)) was carried out along the signal duration. Fig. 5 shows the changes of τ_1 and τ_e in the early 5 s. The integral interval (2*T*) and the running step are 1 s and 0.1 s, respectively. Although the trends of "*Melody B by piano*" were included in Fig. 5 for comparison reasons, this musical motif is not employed in this stereo-dipole examination.

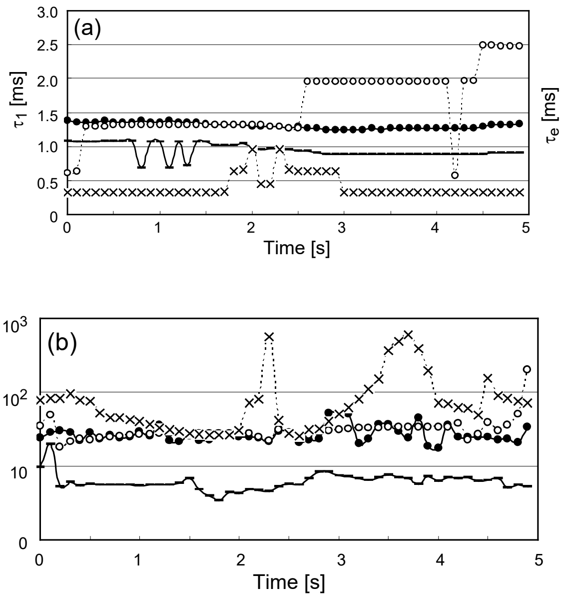


Fig. 5 – (Different symbols indicate different musical motifs: (-): *Melody A* by trumpet; (●): *Melody A* by piano; (×): *Melody B* by organ; (○): *Melody B* by piano. (**a**) Acoustical parameters for 5 s. τ_1 (**b**) Acoustical parameters for 5 s. τ_e

It has to be noticed that τ_1 is affected by the difference in musical instruments (trumpet, organ or piano) and τ_e is mainly affected by the difference in melody (*Melody A* or *B*).

Since both τ_1 and τ_e changes dynamically over time as shown in Fig. 5, it is not easy to determine a unique representative value to express distinctly the differences between these musical motifs. Particularly, the values of τ_e increase to a high value, so that the mean value of τ_e is meaningless. In this study, the 300 values obtained by the running ACF in a rate of 0.1 s along the duration of 30s are converted into the histogram, and the representative values are determined by the 50 % probability of cumulative frequency. These values are termed "τ_1 (50%)" and "τ_e (50%)", and they are listed in Table 1.

Table 1 – Anechoic musical motifs and their τ1 (50%) [ms] and τe (50%) [ms]

Musical motif	τ_1 (50%) [ms]	τ_e (50%) [ms]
Melody A by piano	1.33	246.5
Melody A by trumpet	0.88	54.9
Melody B by organ	0.46	526.7
Melody B by piano	1.94	308.8

2.5 Procedure of Dual Stereo-Dipole

2.5.1 Measurement in Arlecchino listening room

The single stereo-dipole representations were carried out in an Arlecchino listening room at University of Bologna (Italy). Two loudspeakers (Montarbo W400A) were located in front of a dummy head (Sennheiser) as reported in Fig. 6, whereas the other two loudspeakers (Montarbo W400A) were located to the rear of it. To obtain the BIRs in the listening room, a log swept-sine signal was generated by Adobe Audition and was presented by the two loudspeakers alternately.

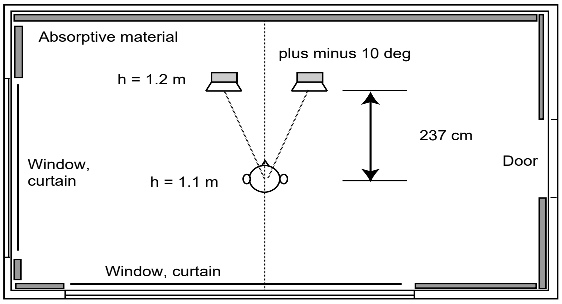


Fig. 6 – Arrangement of two loudspeakers and a dummy head in the Arlecchino listening room

After deconvolution of the signals recorded by the dummy head, the impulse response of the listening room can be obtained respectively for the left and right loudspeakers. The envelopes of impulse responses are smoothed in order to remove extra reflections and to leave only the direct sound.

2.5.2 Generation of cross-talk canceling filter

The smoothed impulse response was converted into cross-talk cancelling filter by using the plug-in of "Invert Kirkeby"1 in Adobe Audition. Table 2 shows the calculation conditions of the Invert Kirkeby plug-in.

Table 2 – Properties of Invert Kirkeby plug-in

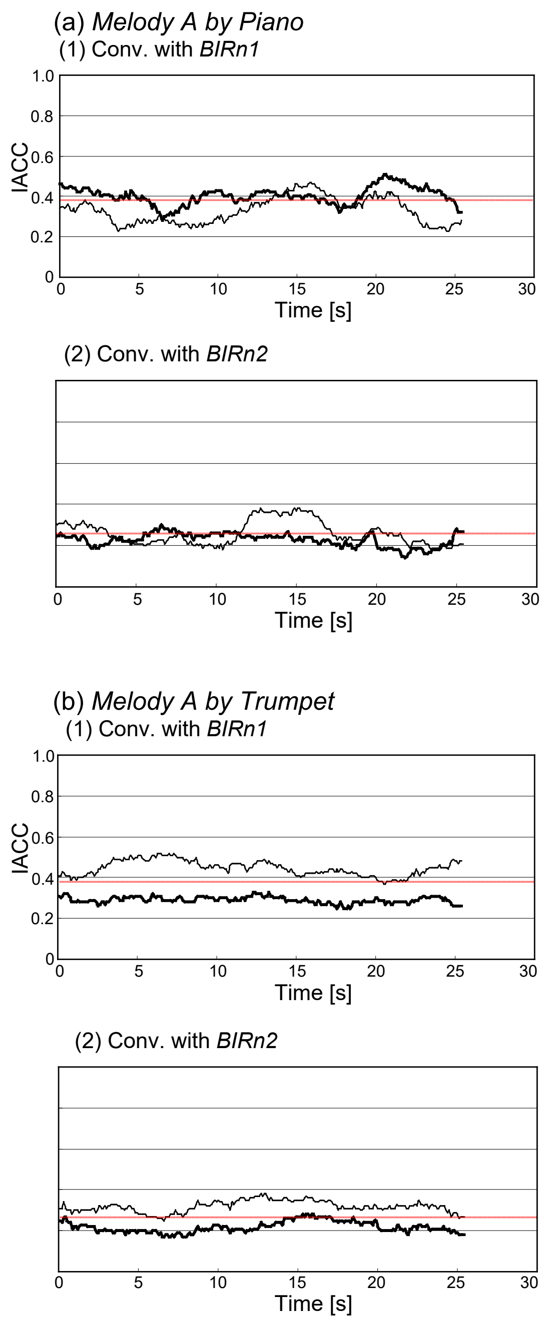
Filter length [sample]	2048
IN-band parameter	1
OUT-band parameter	10
Lower cut freq. [Hz]	80
High cut freq. [Hz]	16000
Width	0.33

2.5.3 Presentation

The "anechoic music" was convoluted with the impulse responses of the theatres. Conversely, the "echoic music" was convoluted again by the cross-talk cancelling filters based on the impulse response of the listening room. The resulted signals were presented by the two loudspeakers at the same time, and the sounds were recorded by the dummy head under almost the same conditions as when the impulse response of the Arlecchino listening room was measured. The recorded musical motifs are defined by "virtual echoic music".

3. Results

During a musical performance, the IACC values vary as a function of time. The observation of the IACC fluctuation is helped by running IACF like the running ACF (Farina and Tronchin, 2013; Tronchin, 2013). Then, the IACF was computed in the range of integral interval 2T (1 s) that is sliding (step: 0.1 s) along the duration of the motif (30 s) after passing through the A-weighting filter. Fig. 7 compares the temporal fluctuation of the IACC produced by the running IACF in the cases of the echoic music (thick line) and the virtual echoic music (thin line). For *Melody A* by piano, the values of IACC are similar among the echoic and virtual echoic music, although it is difficult to observe the synchronous change.



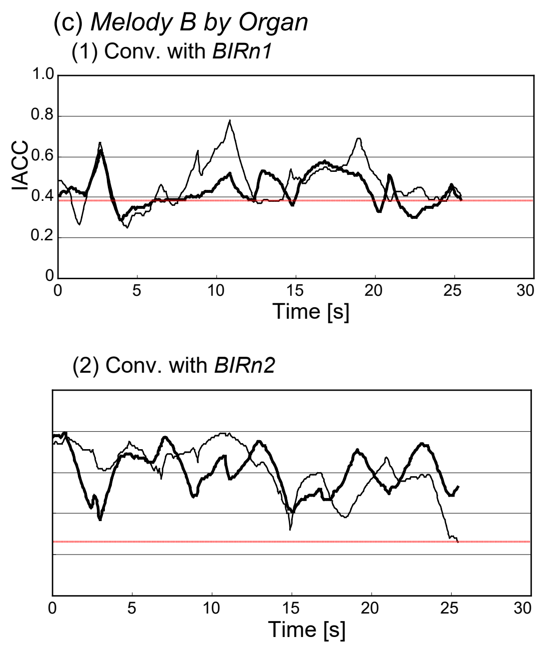


Fig. 7 – Running IACF as a function of time. The thick and thin lines indicate echoic and virtual echoic music, respectively. The red dotted line indicates the values of IACC calculated from the all-passed BIRs

For *Melody A* by trumpet, the differences between them are evident. For *Melody B* played by organ, the IACCs at some moments are quite different from each other; however, some IACCs simultaneously fluctuate between the echoic and virtual echoic music. In a further step, we compared the different distribution of IACC between the echoic and virtual echoic music. It is unlikely that listeners follow the dynamical change of IACC; they are more likely to judge the spaciousness inclusively during musical performances. The running IACC arranged on a long time is converted into a histogram, and the cumulative frequency is rearranged along the IACC. Fig. 8 shows the results. The distribution of IACC is close when the sound source is *Melody B* by organ. On the other hand, in the case of *Melody A* by trumpet, the distributions of IACC are more different from each other. Moreover, the results show that the case of BIRn2 can represent the vertical echoic music more accurately than the case of BIRn1.

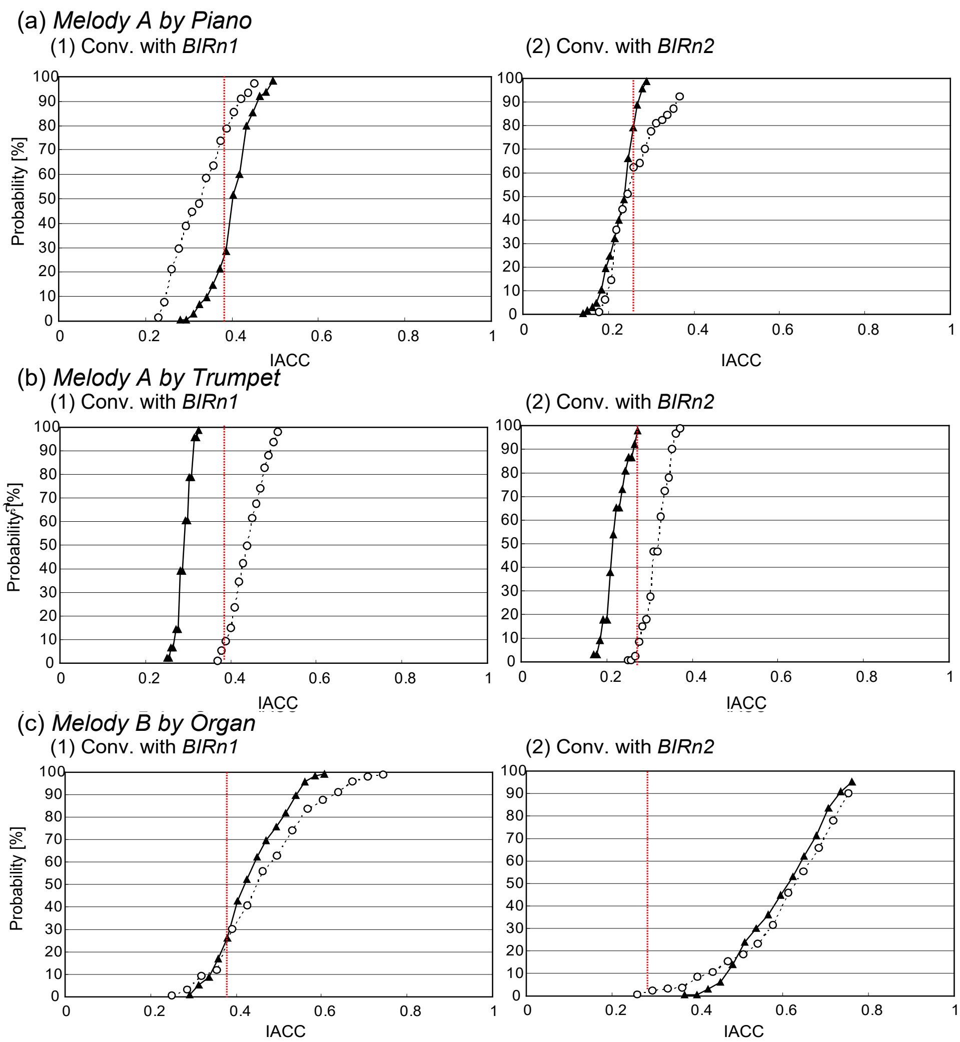


Fig. 7 Cumulative frequencies as a function of IACC of the echoic music (▲) and the virtual echoic music (○). The red dotted line indicates the values of IACC calculated from the all-passed BIRs

In this study, the gap of IACC between the echoic and virtual echoic music was evaluated with

$$Error = \int_{1}^{100} |IACC_{echoic}(x) - IACC_{vecho} \quad (6)$$

where $IACC_{echoic}(x)$ and $IACC_{vechoic}(x)$ are the values of the IACC calculated from the echoic music and the virtual echoic music in the probability x %. The errors can be seen in Table 3. It is important that the accuracy of the stereo-dipole is dependent not only on the kinds of BIRs, but also on the kinds of musical motifs. Although the kind of melody is the same, the errors in *Melody A* by piano and *Melody A* by trumpet are different. Although the kinds of motif are not enough to support the statistical significance, the error values have a good correlation with τ_e (50%) extracted from anechoic musical motifs.

Table 3 – Errors of IACC arranged in terms of BIR and musical motif

Musical motif	**BIRn1**	**BIRn2**
Melody A by piano	0.07	0.03
Melody A by trumpet	0.16	0.10
Melody B by organ	0.04	0.03

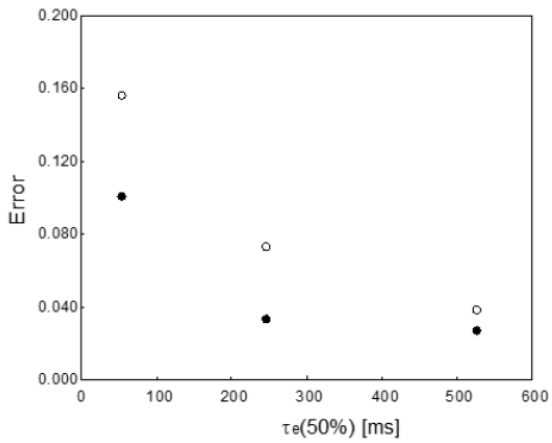


Fig. 8 – Errors of IACC as a function of τ_e(50%) of anechoic musical motif

4. Conclusion

Ando developed a running ACF and IACF calculation of long continuous signal and proposed the effective duration, τ_e, extracted from the ACF of an anechoic musical signal to quantify the acoustical characteristics of it (Ando, 1998). These studies commonly emphasize the usefulness of τ_e in blending musical motif and sound field. In this study, three kinds of anechoic musical signals were employed to examine the accuracy of sound field representation by the stereo-dipole with a view to conducting the subjective experiment judging spatial impression of echoic musical motifs. The error of the IACC ranges from 0.03 to 0.16; this result seems to suggest that the stereo-dipole systemized in the listening room can reproduce the virtual sound field of the opera house, Teatro Nuovo di Spoleto, with a high correlation. The accuracy of results is dependent both on the kinds of BIR and on the kinds of the musical motif. It is interesting that the anechoic musical signal with longer τ_e improves the accuracy of stereo-dipole representation.

Acknowledgement

The authors wish to thank Ryota Shimokura and Lamberto Tronchin, for their collaboration during the experimental measurements

References

Ando, Y. 1998. *Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners*. AIP Press/Springer-Verlag, New York.

Caniato, M., and A Gasparella. 2019. “Discriminating People’s Attitude towards Building Physical Features in Sustainable and Conventional Buildings.” *Energies* 12(8):1429. doi:10.3390/en12081429

Fabbri, K., L. Tronchin and V. Tarabusi. 2014 “Energy retrofit and economic evaluation priorities applied at an Italian case study.” *Energy Procedia* 45: 379-384. doi:10.1016/j.egypro.2014.01.041

Fabbri, K., and L. Tronchin. 2015. “Indoor environmental quality in low energy buildings” *Energy Procedia* 78: 2778–2783. doi: 10.1016/j.egypro.2015.11.625

Farina, A., A. Langhoff, and L. Tronchin. 1998. "Acoustic Characterisation of “virtual” Musical Instruments: Using MLS Technique on Ancient Violins." *Journal of New Music Research* 27(4): 359-379. doi:10.1080/09298219808570753

Farina, A., and L. Tronchin. 2000. "On the "Virtual" Reconstruction of Sound Quality of Trumpets." *Acustica* 86(4): 737-745.

Farina, A. 2001. “Acoustical quality of theatres: correlations between experimental measurements and subjective evaluations.” *Applied Acoustics* 6: 889-916.

Farina, A., and L. Tronchin 2005. “Measurements and reproduction of spatial sound characteristics of auditoria.” *Acoustical Science and Technology* 26(2): 193-199. doi.org/10.1250/ast.26.193

Farina, A., and L. Tronchin. 2013. "3D Sound Characterisation in Theatres Employing Microphone Arrays." *Acta Acustica United with Acustica* 99(1): 118-125. doi:10.3813/AAA.918595

Mancini, F., C. Clemente, E. Carbonara, and S. Fraioli. 2017. "Energy and Environmental Retrofitting of the University Building of Orthopaedic." *Energy Procedia* 126: 195–202. doi: 10.1016/j.egypro.2017.08.140

Shimokura, R., L. Tronchin, A. Cocchi, and Y. Soeta. 2011. "Subjective Diffuseness of Music Signals Convolved with Binaural Impulse

Responses." *Journal of Sound and Vibration* 330(14): 3526-3537. doi:10.1016/j.jsv.2011.02.014

Tronchin, L. 2012. "The Emulation of Nonlinear Time-Invariant Audio Systems with Memory by Means of Volterra Series." *AES: Journal of the Audio Engineering Society* 60(12): 984-886.

Tronchin, L. 2013a. "On the Acoustic Efficiency of Road Barriers: The Reflection Index." *International Journal of Mechanics* 7(3): 318-326.

Tronchin, L. 2013b. "Francesco Milizia (1725-1798) and the Acoustics of His Teatro Ideale (1773)." *Acta Acustica United with Acustica* 99(1): 91-97. doi:10.3813/AAA.918592

Tronchin, L., M.C. Tommasino and K. Fabbri. 2014. "On the cost-optimal levels of energy-performance requirements for buildings: A case study with economic evaluation in Italy." *International Journal of Sustainable Energy Planning and Management* 3: 49-62. doi:10.5278/ijsepm.2014.3.5

Tronchin, L., and V. L. Coli. 2015. "Further Investigations in the Emulation of Nonlinear Systems with Volterra Series." *AES: Journal of the Audio Engineering Society* 63(9): 671-683. doi:10.17743/jaes.2015.0065

Tronchin, L., and D. J. Knight. 2016. "Revisiting Historic Buildings through the Senses Visualising Aural and Obscured Aspects of San Vitale, Ravenna." *International Journal of Historical Archaeology* 20(1): 127-145.

Tronchin, L., M. Manfren and L C. Tagliabue. 2016. "Optimization of building energy performance by means of multi-scale analysis – Lessons learned from case studies" *Sustainable Cities and Society 27*:296-306. doi: 10.1016/j.scs.2015.11.003

Tronchin, L., and K. Fabbri. 2017. "Energy and Microclimate Simulation in a Heritage Building: Further Studies on the Malatestiana Library." *Energies* 10(10). doi:10.3390/en10101621

Tronchin, L., M. Manfren and P. A. James. 2018. "Linking design and operation performance analysis through model calibration: Parametric assessment on a Passive House building." *Energy* 165(A): 26-40. doi:10.1016/j.energy.2018.09.037

Tronchin, L., M. Manfren, V. Vodola. 2020a. "The carabattola - vibroacoustical analysis and intensity of acoustic radiation (IAR)." *Applied Sciences* 10(2), 641.

Tronchin, L., M. Manfren, V. Vodola. 2020b. "Sound characterization through intensity of acoustic radiation measurement: A study of persian musical instruments." *Applied Sciences* 10(2), 633.

Tronchin, L., F. Merli, M. Manfren. B. Nastasi. 2020c. "The sound diffusion in Italian Opera Houses: Some examples." *Building Acoustics,* in press. doi:10.1177/1351010X20929216

Tronchin, L., F. Merli, M. Manfren. B. Nastasi. 2020d. "Validation and application of three-dimensional auralisation during concert hall renovation." *Building Acoustics*, in press. doi:10.1177/1351010X20926791