SUBJECTIVE ASSESSMENT OF AURALIZATIONS CREATED BY MEANS OF GEOMETRICAL ACOUSTICS AND FINITE ELEMENT NUMERICAL METHODS

Luis A. Tafur Jiménez, Santiago Ruiz Sánchez, Jonathan Ochoa Villegas and Libardo A. Londoño Ciro.

University of San Buenaventura, Faculty of Engineering, Medellín, Antioquia, Colombia
email: decano.ingenieria@usbmed.edu.co

This paper presents a subjective assessment of auralizations created by means of numerical methods such as Geometrical Acoustics (GA) and Finite Elements Method (FEM). A subjective test was implemented in order to evaluate the virtual sound environments created for a meeting room at the Institute of Sound and Vibration Research (ISVR), in the University of Southampton, UK. The auralizations were reproduced in a binaural 3D system based on Optimal Source Distribution (OPSODIS). The perceptual parameters assessed in the subjective test were source localization, reverberation or spatiality, warmth and brightness. Three sound sources were used to create the auralizations: a human voice (speech), a percussion instrument and a wind instrument, all of them recorded in a recording studio. Three groups of auralizations were created, the first two using the following numerical approaches: GA and the combination of GA and FEM. The third group of auralizations was created by means of Binaural Impulse Responses (BIR) measurements. A comparison between the auralizations created using the numerical approaches was carried out having as a reference the auralizations obtained by means of BIR measurements. The reproduction system was located in a controlled environment and the evaluation scale was established according to the recommendations of the ITU-R BS.562-3,1990. The statistics results indicated that the auralizations created using a hybrid approach presented more similarity for all the perceptual parameters in comparison with the auralizations created with only GA.
Keywords: auralizations subjective assessment, hybrid approach.

1. Introduction

This study was part of a PhD research [1], where the sound wave propagation was analysed in two rooms with different conditions (size, shape and purpose) by means of subjective assessment. This paper presents the results obtained for a subjective assessment of three groups of auralizations for the room number 2011 located on the second floor in the TIZARD building of University of Southampton. A virtual sound environment or Auralization is the process of audibly rendering the sound field created by a source in a space, to reproduce the binaural hearing experience at a given position. Vorländer [2] defined auralization as a technique consisting of three stages: sound generation, sound transmission and sound reproduction. The first stage describes the procedures to characterize the sources to be auralized. Sound transmission involves the methods used to estimate the BIR and the last phase includes the reproduction system and the signal processing required to convolve the output of the first two stages.

The virtual sound environments were created using Geometrical Acoustics (GA) simulations, a combination of GA and Finite Element Method (FEM) and, as a reference, Binaural Impulse Responses (BIR) measurements. The reproduction was done by means of the OPSODIS reproduction system. The subjective test was designed with the purpose of evaluating perceptual parameters such as: localization of the source, reverberation or sense of space, warmth and brightness. In the statistical results of the subjective test, the hybrid approach (FEM-GA) presented better results than those of GA simulation. The localization parameter was evaluated as the most similar in the simulations with respect to the measured reference auralizations.

2. Creation of auralizations

This section presents the procedure to create three groups of auralizations used in this research. The numerical approaches compared in this study considered auralizations created using GA and a hybrid combination of FEM-GA. The third group of auralizations were created by means of BIR measurements. The last were used as a reference to evaluate the performance of each numerical approach.

2.1 Acoustic measurements

Binaural Impulse Response (BIR) measurements in the selected room were carried out. The aim was to obtain the room's transfer path for the source-receiver combinations and to create the reference auralizations. The measurements were carried out in room number 2011 located on the second floor in the TIZARD building of University of Southampton. The enclosure has a rectangular shape of 7.3 m long, 3.1 m width and 3 m height, with an approximate volume of 67 m³. In Fig. 1 the shape and dimensions of the room can be seen.

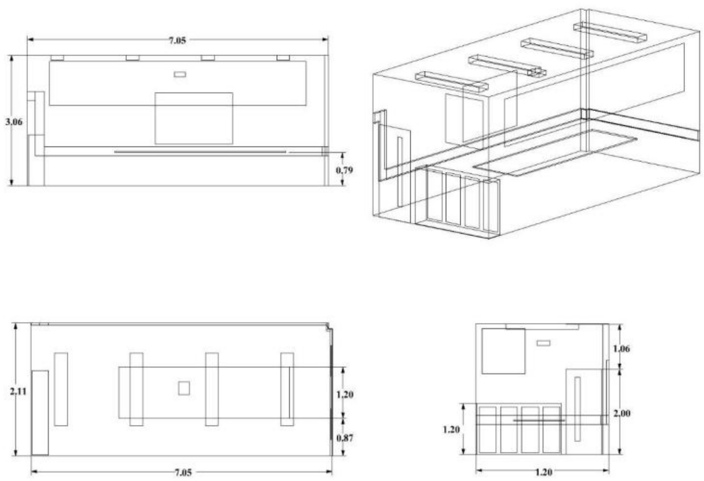

Figure 1: Top, frontal, lateral and isometric views.

The loudspeaker was placed on the table which has a height of 0.81m and the dummy head height was 1.2m. Figure 2 shows the positions used in the measurements.

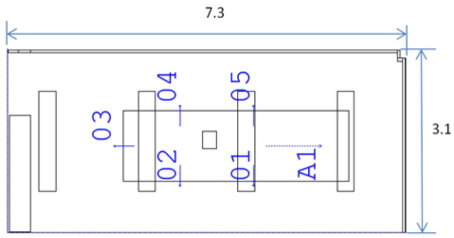

Figure 2: Source (A1) and binaural receiver (01 – 05) positions.

The equipment used in the measurements consisted of a two-way active loudspeaker MACKIE SRM305v2, a omni-directional microphone B&K 4159 with a 505mV/Pa sensitivity and the excitation signal was a Log Sine Sweep from 20Hz to 20kHz.

2.2 Geometrical Acoustic simulations

The software CATT-Acoustic version 9 was used to implement the GA methods. With the purpose of creating virtual sound environments, the hybrid technique “Randomized Tail-corrected Cone-tracing” (RTC), which includes Image Source Method (ISM) for low-order reflections and Ray Tracing (RT) method for high-order reflections, has been applied.

The 3D model was constructed in a CAD software first to import it into each software (GA and FE) to guarantee similar geometric conditions. In CATT-Acoustic, after importation process a *GEO-file* was generated. The final 3D model created in the CATT software is shown in Fig. 3.

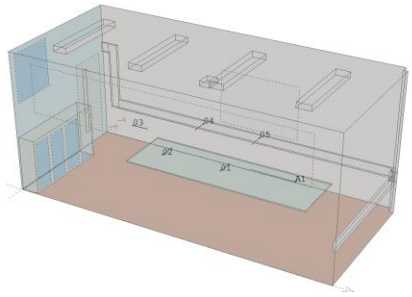

Figure 3: GA model created in CATT-Acoustics after importing 3D model in CAD language. A1 denotes the acoustic source location and the consecutive numbers from 01 to 05, the receiver positions simulated.

The model required each planar face assigned with its material acoustic properties of absorption and scattering. Hence, the absorption coefficients for each surface was taken from the library provided by the software. The scattering coefficients for flat surfaces were set at a minimum of 10%. For standalone objects, such as table, furniture, projector and lights, a frequency dependent scattering coefficient was applied using the *Automatic edge diffusion* option provided by the software. This function applies significant diffusion to a surface if its size is small compared to the sound wavelength [3]. To insert a sound source in CATT-Acoustic the following information is needed: acoustic centre location, source orientation, sound level pressures defined at 1 m distance for all octave band frequencies and directivity data. For the first two, the Cartesian coordinates were introduced matching the location and source orientation applied in the acoustic measurements. The directivity information provided by the manufacturer was introduced using the source directivity module available in the software.

2.3 Hybrid approach (FEM-GA)

For the FEM, a time-harmonic simulation up to 700 Hz with 1Hz frequency steps, was carried out in the software COMSOL 4.3. The source-receiver combinations were the same as in GA simulations and acoustic measurements. To be able to combine the FE simulation results with GA results, the parameters used in the latter were also applied in the FE model.

The above-mentioned CAD model was used to generate the geometry in COMSOL. Figure 4 shows the 3D model created in COMSOL.

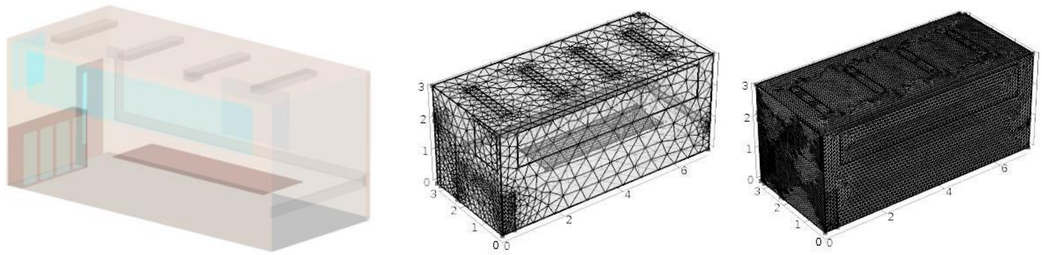

Figure 4: FE model created in COMSOL after importing the 3D model in CAD language (left). The coarsest and the finest mesh resolutions implemented in COMSOL (right).

In order to estimate up to a frequency of 700Hz, a system with approximately 1,000,000 degrees of freedom had to be solved. Therefore, different mesh resolutions for groups of frequencies were used (see Fig. 4), changing the maximum element size according to the maximum frequency estimated, as shown in Table 1. In this model the MUMPS solver was applied for solving the algebraic linear system, for dealing with symmetric and non-symmetric matrices.

Table 1: Maximum frequency estimated, maximum element size, DOF, average time estimation per frequency and approximate number of points per wavelength for each simulation ran in COMSOL.

<table>
<tr><th>Max. Frequency Estimated (Hz)</th><th>Max. Element Size (m)</th><th>DOF</th><th>Average time estimation per frequency (min)</th><th>Number of points per wavelength</th></tr>
<tr><td>50</td><td>0.686</td><td>122.728</td><td>0.3</td><td>10.0</td></tr>
<tr><td>100</td><td>0.343</td><td>207.869</td><td>2</td><td>10.0</td></tr>
<tr><td>150</td><td>0.229</td><td>326.760</td><td>4</td><td>10.0</td></tr>
<tr><td>200</td><td>0.172</td><td>486.905</td><td>9</td><td>10.0</td></tr>
<tr><td>250</td><td>0.140</td><td>558.951</td><td>12</td><td>9.8</td></tr>
<tr><td>300</td><td>0.122</td><td>802.695</td><td>18</td><td>9.4</td></tr>
<tr><td>350</td><td rowspan="15">0.111</td><td rowspan="15">922.648</td><td rowspan="15">26</td><td>8.8</td></tr>
<tr><td>375</td><td>8.2</td></tr>
<tr><td>400</td><td>7.7</td></tr>
<tr><td>425</td><td>7.2</td></tr>
<tr><td>450</td><td>6.8</td></tr>
<tr><td>475</td><td>6.5</td></tr>
<tr><td>500</td><td>6.2</td></tr>
<tr><td>525</td><td>5.9</td></tr>
<tr><td>550</td><td>5.6</td></tr>
<tr><td>575</td><td>5.4</td></tr>
<tr><td>600</td><td>5.1</td></tr>
<tr><td>625</td><td>4.9</td></tr>
<tr><td>650</td><td>4.7</td></tr>
<tr><td>675</td><td>4.6</td></tr>
<tr><td>700</td><td>4.4</td></tr>
</table>

According to Aretz [4], to specify the impedance in FE simulations for extended reaction materials, an approach consisting in defining a real valued frequency dependent impedance, corresponding to the absorption coefficients can be considered. Therefore, the impedance boundary conditions were defined using the field incidence absorption coefficient applied in the GA model to find the resistance part of impedance, which is associated with energy loss by either dissipation or transmission [4]. The receiver positions (see Fig. 2) and acoustic conditions in the GA model were considered in the FE model. The source was characterised as a monopole point radiating uniformly 1Pa of acoustic pressure at 1m distance, in a frequency independent spherical propagation. For obtaining the BIR, a simple approach using a cube and two receiver points at corresponding ear positions simulated the HRTF (see Fig. 5). The cube size was defined according to the average physical dimensions of a head established in the IEC/TS 60318-7-2011 standard [5].

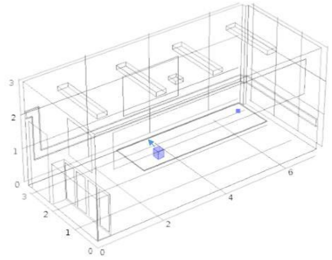

Figure 5: Binaural receiver model in FE simulations.

A high-pass filter for GA results and band-pass filter for FE simulations were implemented, with the purpose of combining both numerical results. The phase responses of the filter applied to the FE signals, high pass filter applied to GA signals and the combination given by FE and GA results are shown in Fig. 6. In the FE method, frequency domain results are obtained with an apparent sampling frequency which depend on the frequency steps defined. In order to match the sampling frequency of the GA simulation results it is therefore necessary, padding with zeros outside the frequency interval estimated with a number according to the GA BIR length. An Inverse Discrete Fourier Transform (IDFT) follows this procedure to obtain the FE BIR. Finally, both filtered signals are summed to obtain the FEM-GA auralizations.

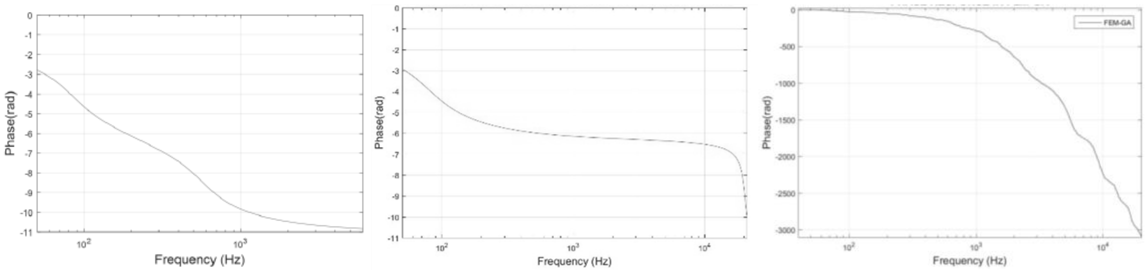

Figure 6: Phase responses of band-pass filter applied to the FE signals (left), high pass filter applied to GA signals(centre) and an example of a FEM-GA signal (right).

3. Subjective evaluation procedure

To evaluate the virtual sound environments created with the numerical methods with respect to the auralizations created by means of BIR measurements, a subjective test was implemented. Three groups of auralizations (BIR measurements, GA and FEM-GA) were created using three audio samples: a human voice (speech), a bass drum and a sax horn. Three different receiver positions were used in the auralizations in order to have information for spatial evaluation. The parameters evaluated were: localization of the source, reverberation or sense of space, warmth and brightness. The auralizations were reproduced in the Optimal Source Distribution (OPSODIS) reproduction system that enables a lossless approach of inverse filters [6]. The system was developed as a 3D sound reproduction technology applying a conceptual pair of monopoles transducers, where the angle span is changing continuously as a function of frequency. The OPSODIS system is given by the ninety-degree phase modification on the crosstalk path in the inverse filter matrix, which guarantees that synthetized sound is always reproduced by constructive interference with no loss of dynamic range. Moreover, the radiation pattern becomes constant over frequency and does not emanate excessive sound to the surrounding environment, which allows multiple listeners and robustness against spurious reflections.

The test was conducted in the recording studio A of the University of San Buenaventura, which has favourable acoustical conditions to carry out subjective tests (see Fig. 7). This room allowed presenting the auralizations and the subjective tests, since it has a much lower background noise level and shorter reverberation time than the classroom investigated. The methodology of the test consisted of a pairwise comparison of samples A and B, where the first was the reference auralization created by BIR measurements and the second was given by the virtual sound environment generated numerically, either with GA or by the combination of FEM-GA. The participants had to rate the parameters mentioned above of sample B, with respect to reference sample A. The listener was free to play, stop or repeat samples using a tablet with an application designed for the experiment in Pure Data language. The scale used in the test is shown in Table 2, following the recommendations by the ITU Radio Communication Assembly in the document “Subjective assessment of sound quality” [7].

Table 2: Subjective assessment scale.

RATING	ASSESSMENT
Not different	5,0
Slightly not different	4,0
Slightly different	3,0
Rather different	2,0
Completely different	1,0

Figure 7: Picture of the place where the subjective tests were carried out

The definitions of the parameters evaluated are as follows:

- Localization: attribute associated to a subjective perception of the direction indicating the origin of sound and the relative position of the source.
- Sense of space: Similar to reverberation, this parameter refers to a subjective permanence of reflected sound in the enclosure. In other words, it indicates a subjective size impression of the room in acoustic terms.
- Warmth: attribute denoting a subjective perception of loudness at low frequencies of the corresponding source.
- Brightness: parameter indicating a subjective perception of loudness at high frequencies of the corresponding source.

The population determined for this research consisted of students from the sixth semester onwards of the Sound Engineering programme at the University of San Buenaventura, in Medellín Colombia. According to university data, there were 103 students enrolled in these semesters at the moment the test was applied. The sample size consisted of a simple random sampling of 40 students taking into account three aspects: the estimated confidence level, the permissible error and the finite character of the population (less than 100,000 people). Box-and-Whisker Plots were used to represent the statistical data obtained, with the purpose of summarizing statistic measures such as a minimum, a lower quartile, a median, an upper quartile and a maximum, and indicate the presence of outliers.

4. Results

Figure 8 shows the spatially averaged reverberation time (T20) of the room obtained from measurements and numerical approaches. Figure 9 shows the statistical results for the bass drum of the four parameters assessed for each numerical simulation (localization=LOC, sense of space=REV, warmth=WRM and brightness=BRI) at three specific positions in the room (see Fig. 2). Figures 10 to 12 show the spatially averaged results of the four parameters for the speech voice, the bass drum and the sax horn on each numerical technique. The point symbols represent the outliers which values exceed 1.5 times the interquartile range from the edge of the box. Any point which value exceeds more than three times the interquartile range is called far outside point, and is indicated with a plus sign.

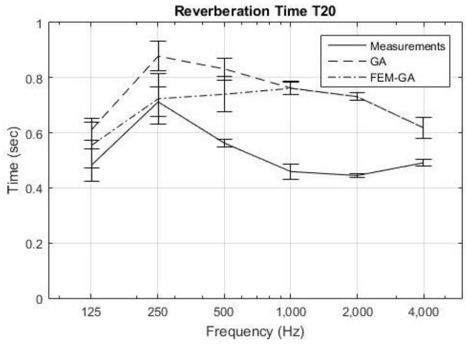

Figure 8: Spatially averaged $T20$ results of the room calculated from measurements, GA simulations and the numerical approach combination of FEM-GA.

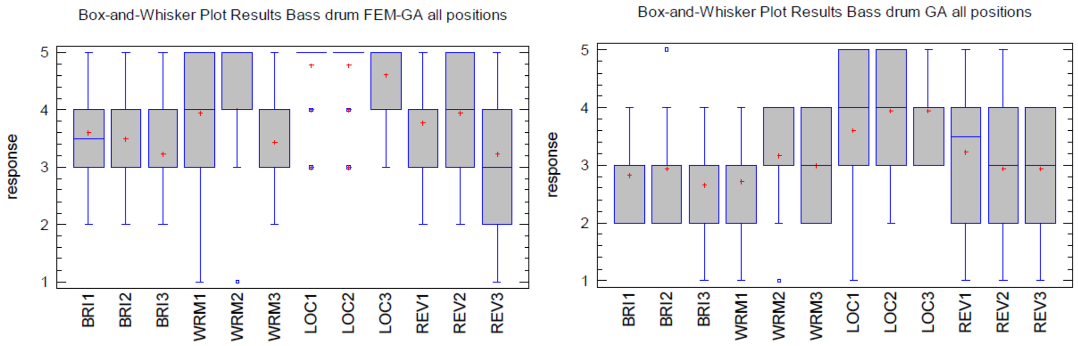

Figure 9: Estimates at each position of the four parameters evaluated for the bass drum on each numerical technique (on left the hybrid approach FEM-GA and on right, GA).

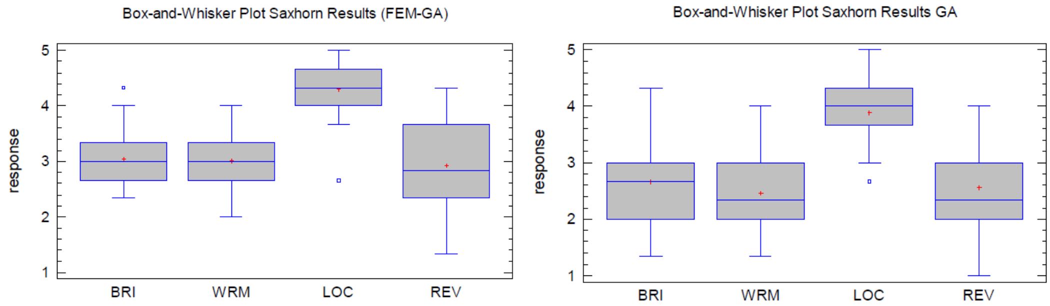

Figure 10: Spatially averaged estimates of the four parameters evaluated for the saxhorn on each numerical technique (on left the hybrid approach FEM-GA and on right, GA).

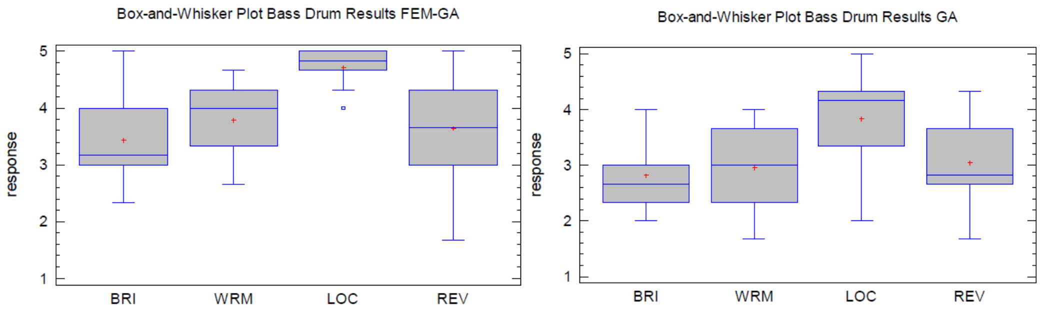

Figure 11: Spatially averaged estimates of the four parameters evaluated for the bass drum on each numerical technique (on left the hybrid approach FEM-GA and on right, GA).

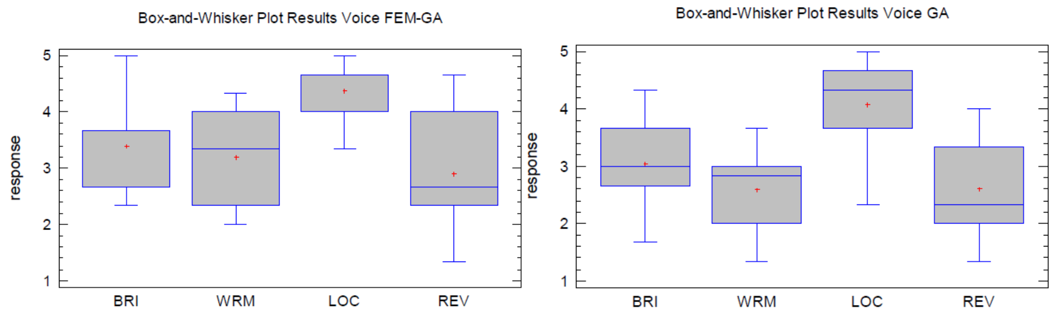

Figure 12: Spatially averaged estimates of the four parameters evaluated for the male voice on each numerical technique (on left the hybrid approach FEM-GA and on right, GA).

5. Discussion

The bass drum results indicate that the auralizations created by means of the hybrid approach were assessed as being more similar with respect to reference auralizations than GA simulations, for all the parameters and receiver positions (see Fig. 9). The tendency results for the hybrid approach sorted from the highest to the lowest as follows: localization (4.8), warmth (4.0), reverberation (3.8) and brightness (3.5), the first two in position 2 and the last two in receiver location 1. This numerical technique has presented the lowest rate for the parameter of reverberation (3.2) at position 3. The GA subjective results exhibited the best tendencies at position 2 for the parameters of localization (4.0), warmth (3.1) and brightness (3.0), and location 1 for reverberation (3.3). In this case, the parameter exposing the lowest score was the brightness (2.7) at position 3. Figure 10 to Figure 12 illustrate how every single source obtained higher scores in all the parameters evaluated, when the auralizations simulated by means of the hybrid approach (FEM - GA) were listened. It is observed from the figures how the localization of the source was the parameter subjects judged to be the most similar with respect to measured reference auralizations. In the evaluation of this parameter, the bass drum is the acoustic source with the best scores meaning that it was very difficult to distinguish between measurement and simulation. It is important to note for that instrument how the hybrid approach exhibited the best responses in all the parameters but brightness, which had comparable results in the male voice auralizations.

6. Conclusions

A hybrid numerical approach combining FEM and GA simulations was applied to estimate the sound propagation in all the audibly frequency range, in order to assess the advantages and drawbacks of the hybrid model to create auralizations of non-built spaces. The use of material parameter databases was considered with the purpose of defining the boundary conditions in the numerical modelling, in this sense the GA absorption coefficient databases provides reasonable input data to construct the GA models. Nevertheless, FE boundary conditions data is limited and more difficult to access than GA information.

The implementation of the hybrid numerical approach to create auralizations of the meeting room improved the subjective perception of localization, warmth and reverberance in comparison with GA methods. The subjective test results indicated that for the three sources of study (saxhorn, bass drum and male voice), all the subjective parameters evaluated obtained higher scores when the auralizations simulated by means of the hybrid approach were assessed. It was observed that localization was judged as the most similar with respect to measured reference auralizations, which indicates the importance of the transient responses for this psychoacoustic parameter. Moreover, the bass drum was the source with the best localization assessment that points out how the hybrid approach improved the transient response given by the low frequency range.

REFERENCES

1 Tafur J., L. A., *Assessment of a hybrid numerical approach to estimate sound wave propagation in an enclosure and application of auralizations to evaluate acoustical conditions of a classroom to establish the impact of acoustic variables on cognitive processes*, PhD dissertation, Institute of Sound and Vibration Research – University of Southampton, Southampton, UK (2016).

2 Vorländer, M., *Auralization*. Springer, Berlin, (2008).

3 CATT, *CATT-Acoustic v8.0 User's Manual*. Gothenburg, (2007).

4 Aretz, M., *Specification of realistic boundary conditions for the FE simulation of low frequency sound fields in recording studios*. Acta Acustica united with Acustica, August, **95**(5), 874 – 882, (2009)

5 International Electrotechnical Commission, *Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index*, Geneva: BSI, (2011).

6 T. Takeuchi and P. Nelson, "Extension of the Optimal Source Distribution for Binaural Sound Reproduction," ACTA ACUSTICA UNITED WITH ACUSTICA, p. 981 – 987, 2008.

7 ITU Radiocommunication Assembly, *Recommendation ITU-R BS.562-3. Subjective Assessment of Sound Quality*, Geneva: Radiocommunication, (1990).