On the perception of reflections from directive sources in binaural simulations

Frank Melchior[1,2], Diemer de Vries[1]
[1] TU Delft, Faculty of Applied Physics, The Netherlands
[2] Fraunhofer IDMT, Ilmenau, Germany

Introduction

The directivity of a sound source in an anechoic environment or in a direct sound only reproduction system can not be perceived. By moving the listener can explore the different emission of a source depending on his listening position. In case of a non moving listener one will only perceive a difference in the source timbre itself without effecting the impression of different directivity or a rotating source. To make the directivity of a source perceivable even for a static listener room simulation is required. In the present study the importance of correct auralisation of reflections generated by a mirror image source model (MISM) is investigated.

Room model

To generate the examples for the listening test a MISM was used. The implementation is analog to [1]. As an extension directive sources with idealized directivity are implemented. The directivity is described with:

$$g(\phi,\theta) = \beta + (1-\beta)cos(\phi)cos(\theta) \quad (1)$$

With $\beta \in [0, 1]$. The gain factor $g(\phi, \theta)$ is used to express the directivity depending on the elevation angle ϕ and the azimuth θ using β to vary the directivity between figure-of-eight ($\beta = 0$) and omnidirectional characteristic ($\beta = 1$). For the used idealized characteristic $g(\phi, \theta)$ is set to zero for all $\pi > \theta > 2\pi$. A binaural impulse response ($BRIR$) from an image source model can be calculated using:

$$BRIR(\omega) = \sum_{n=1}^{N} HRTF(\omega, n) f(\omega, n) d_{air}(\omega, n) g(n) \quad (2)$$

This equation includes the sum over N reflections with a frequency response $f(\omega, n)$, an air damping factor $d_{air}(\omega, n)$, a direction factor $g(n)$ and the head related transfer function (HRTF) corresponding to the direction $HRTF(\omega, n)$.

Listening experiment

To minimize the influence of the HRTF individualized HRTFs are used. For this purpose a pretest was performed to select the best fitting HRTF for each subject. The HRTFs were taken from the CIPIC database [2] and diffuse field equalized. The subjects have the possibility to listen to 45 different HRTFs convolved with a test signal. The test signal is a 30Hz modulated white noise. 3 Cycles of these bursts were convolved with one HRTF corresponding to a horizontal direction before switching to the next direction. By that a sound which should move two times around the listeners head is generated. Subjects were advised to choose the HRTF with no elevation and a smooth frequency response during the movement as well as the best externalisation. For the main experiment the test items were generated with the individual HRTF for each subject. During a training phase listeners had the possibility to listen to a rotating sound source with the same characteristics as in the later test. The source was rotated in 45 degree steps around its z-axis. For each source signal two versions were presented and marked with *Good* and *Bad*. *Good* corresponds to the reference of the later test while the example marked with *Bad* was calculated with a omnidirectional characteristic for the MISM (corresponding to item 3 of the later test). The hypothesis of the listening experiment was: *To achieve a good quality in perception of source directivity and direction it is sufficient to generate only the first order reflections in a correct way. Reflections of higher order can be calculated with omnidirectional source directivity.* To proof this hypothesis a multiple stimulus with hidden reference and anchor (MUSHRA) quality grading experiment was conducted. The subject's task was to rate the quality of the presented item compared to a open reference. The criterias are the perceived source direction and differences in the room impression. The MISM was used to generate a BRIR of a rectangular room which is shown in figure 1. The reflection coeffi-

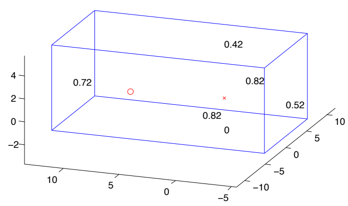

Figure 1: Simulated room with reflection coefficients for each wall (source marked with o receiver marked with x).

cients are frequency independent and presented at the corresponding walls. The aim was not to generate a realistic environment but to try to analyse the worst case. It was expected that in a MISM without any diffraction and diffuse tail the differences between the different reflection patterns are most audible. In other words, if in this situation a given order of correct reflections is sufficient for the perception of directivity than this order is also sufficient for more realistic environments. Two different source signals with different directivities were used while the room conditions have been kept constant. As source signal one sentence of male speech was used (Speech). As a second source signal a mixed sequence of drums and a accompanying electric bass was used (Drums). For each signal the directions A=0° ,B=90°,C=135°,D=292°) were graded. Direction 90° points to the right wall from the listener point of view. Beside this direction an anchor with the same source characteristic pointing in the opposite direction as the reference was always included. For the speech signal a directivity with $\beta = 0.3$ was used. For the drum example the simulation was divided into 2 frequency ranges. Up to 400Hz a onmidirectional directivity was calculated and from 400Hz to 20kHz the same directivity as for the speech sample was applied. For each of the four directions 5 different items were evaluated (hidden reference, anchor, item 1 item 2 and item 3). The reference corresponds to a correct simulation of all reflections. For Item 1 only the direct sound, first order and second order reflections were calculated corresponding to the source direcivity. For item 2 only direct sound and first order reflections were correct and for item 3 only the direct sound was correct. The energy for the reflections have been kept constant for the directional source and the unidirectional simulation. The task for the subject was to grade the quality of different version of one source direction on a 5 point scale as indicated in figure 2. The grading was performed in direct comparison to a open reference. Each grading was performed two times, while the order of the items is randomised.

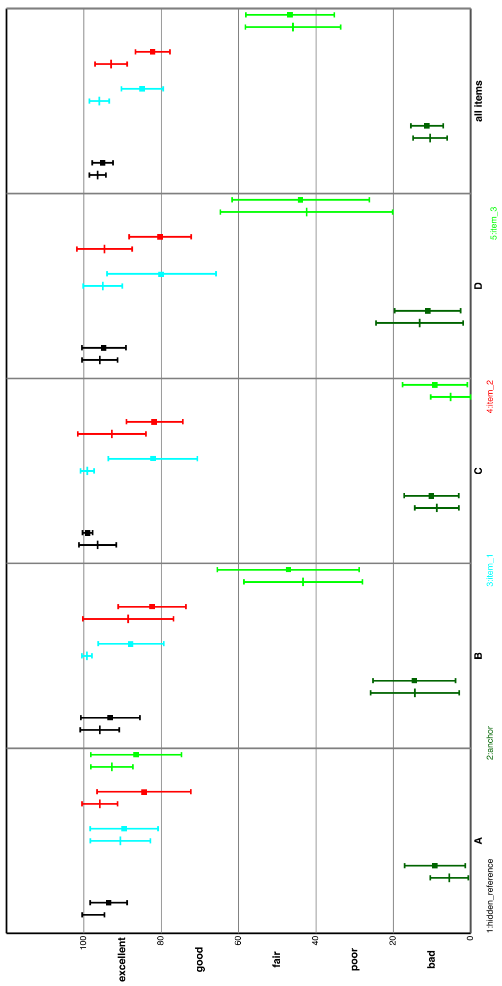

Figure 2: Average results and 95% confidence interval for drum example (square) and speech (bar) example graded by 9 subjects.

Results

The subjects were trained listeners but unexperienced with this kind of simulations. For this reason the first grading passes of the subjects were omitted. The results for the different signals were analysed separately. Figure 2 presents the results. The results for the speech example are marked with a bar for the mean value while the results for the drum example are marked with a square. In the over all rating for the speech example, significant results can be obtained by the hidden reference, item 3 and the anchor. Item 2 and item 3 show significant difference in comparison to the hidden reference but both are in the excellent range. The difference in the rating between item 1 and item 2 is not significant. From this experiment we can conclude that for the correct perception of source direction under the given conditions it is sufficient to model only the first order reflections correctly and render all higher order reflections under the assumption of an omnidirectional source. In comparision of the different source directions it is obvious that subjects could not distinguish the difference between the examples for direction A. Only if the source is rotated a difference is perceived and a grading becomes possible. The overall tendency of the results for the drum examples are comparabel to the speech signal. Remarkable are the results for direction B and C. In this case there is a tendency to grade item 1 and item 2 differently. This is due to the fact that the drum example contains more transients which makes it possible to identify flutter echoes in the room more easily, especially in the given simple image source model. If second order and higher order reflections are not modeled properly these flutter echos are not generated and a clear difference to the reference can be perceived.

Conclusions

For each signal the difference in the grading between different items depends on the source direction. The anchor is always identified as well as the hidden reference. To get an excellent or good grading for the source directivity and room quality the correct simulations of the first order and second order reflections are sufficient. This holds for a speech signal and a broadband directivity simulation as well as for a drum example with a directivity simulation above 400Hz. The simulation of the correct direct sound only results in a bad grading of the examples. It can be concluded that for the simulation of source directivity the first order and second order reflections are most important. The presented work is part of the project EDcine, supported by the IST 6th framework program of the European Commission.

References

[1] J. B. Allen, D. A. Berkley: Image method for efficiently simulating small-room acoustics. J. Acoustic Soc. Am. 66 (1979), 943-950

[2] CIPIC database, URL: `http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm`