

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

Effet de la modalité visuelle sur le jugement de la réverbération d'une salle

Alexis Billon¹, Stéphane Safin², Aurore Defays², Christine Decaestecker³, Nadine Warzée³

¹INTELSIG, Université de Liège, B28, B-4000 Liège, abillon@ulg.ac.be

²LUCID, Université de Liège, B52, B-4000 Liège

³LISA, Université Libre de Bruxelles, Avenue F. Roosevelt, 50, B1050 Bruxelles

Le projet Auralias vise à développer un environnement immersif permettant le travail coopératif entre architectes et acousticiens. Le système développé associe donc un rendu sonore 3D (au moyen d'un dispositif VBAP) et une vue du modèle 3D de la salle considérée. Dans ce cadre, nous nous intéressons à l'impact de l'image projetée sur le jugement subjectif de l'acoustique de la salle. Ici, le critère objectif étudié est le temps de réverbération. Pour cela, des réponses impulsionales synthétiques possédant la même partie précoce (jusqu'à 80 ms) et différentes parties tardives ont été obtenues puis convoluées avec des rythmes de percussions pour obtenir différents échantillons sonores. Après un prétest permettant de définir une série d'images de salles évoquant des réverbérations différentes, un test subjectif en deux parties a été mené sur deux populations différentes. La tâche demandée est à chaque fois un classement des échantillons sonores dans l'ordre de réverbération croissante. Après une situation de contrôle (test réalisé sans l'appui des images), le sujet est soumis à un des trois scénarios suivants : images incitatives (l'échantillon possédant le plus de réverbération est associé à l'image évoquant le plus long temps de réverbération), images distractrices (l'image soutenant l'échantillon sonore est tirée aléatoirement) et de sans image (testant ainsi l'effet d'apprentissage). Les résultats de cette expérience sont détaillés et permettent des premières prescriptions quant aux images à afficher dans l'environnement immersif pour soutenir la tâche de jugement acoustique.

1 Introduction

Le projet Auralias développe un environnement virtuel de travail coopératif entre architecte et acousticien, associant des rendus sonore et visuel tridimensionnels [1]. Une perception directe de l'espace sonore considéré est donc disponible en plus de l'évaluation des critères habituels d'acoustique des salles. Les dimensions perceptuelles permettant de juger et de discriminer l'acoustique d'une salle sont peu nombreuses [2-5] et les trois principales identifiées sont la réverbération [6], l'aspect spatial [4,5] et la force du son [2,5].

La dimension perceptuelle de réverbération renvoie aux critères objectifs de temps de réverbération [5,6] ou de décroissance précoce [2]. Notre étude se limite à l'influence du temps de réverbération Tr30 (temps de décroissance de l'énergie sonore de -5 à -35 dB rapporté à 60 dB), sa plus petite différence perceptible (jnd) étant de 24 ms pour un son à bande étroite, [7] et 5% pour un son complexe ([8] adapté de [9]).

Or, il est maintenant bien connu que les informations perçues par une modalité sensorielle peuvent affecter la perception et le traitement d'une autre modalité sensorielle [10,11]. Ainsi, l'éloignement de la source et le ressenti spatial de l'acoustique d'une salle sont affectés par la ou les modalités stimulées [12]. L'impact de la modalité visuelle sur la réverbération a été aussi étudié mais en termes d'attente par le sujet [13]. Celui-ci devait modifier le rendu sonore d'une salle dont la représentation visuelle lui était soumise jusqu'à ce que les modalités lui semblent cohérentes [13]. Les résultats obtenus mettent en évidence qu'un sujet possède une idée *a priori* de l'ambiance acoustique d'une salle en fonction de son architecture. Cette attente peut donc intervenir comme un biais lors de l'évaluation perceptive d'une salle. L'influence d'une

représentation graphique sur la réverbération est donc ici investiguée, non plus en termes d'attente, mais en termes de discrimination.

Cette article est organisé de la manière suivante : le protocole expérimentale est décrit au sein de la section 2. La section 3 présente les résultats obtenus qui sont discutés en section 4 suivie d'une conclusion.

2 Protocole expérimental

2.1 Plan expérimental

Chaque sujet est amené à effectuer la même tâche de classement des extraits sonores en ordre de réverbération croissante deux fois : une première fois en situation de contrôle (sans image et avec le motif sonore 1) puis une seconde fois avec le motif sonore 2, soit en présence d'images (situations 1 et 2), soit sans image (situation 3), pour juger de l'effet d'apprentissage éventuel (Table 1).

Pour la condition 1 « images évocatrices », le rang de réverbération de l'extrait sonore est corrélé à la réverbération évoquée de l'image affichée. Dans la situation 2 « images distractrices », les extraits sonores et les images sont associés de manière aléatoire. Les expériences se déroulent autour d'un ordinateur portable où des extraits sonores sont diffusés par un casque ouvert Sennheiser HD 600. Une interface Matlab a été développée permettant la lecture des extraits sonores, l'affichage des consignes et des images ainsi que le traitement des résultats. Les consignes affichées sont les suivantes :

- « Bonjour, dans la première partie de cette expérience, vous allez entendre cinq extraits sonores. Nous vous demandons de classer ces différents extraits par ordre de réverbération (1 pour le moins résonnant et 5 pour le plus résonnant). »

- « Maintenant, nous allons vous présenter cinq autres extraits sonores. Ces différents extraits seront accompagnés de l'image de l'endroit où l'extrait sonore a été enregistré. Nous vous demandons également de classer ces extraits, par ordre de réverbération, du moins résonnant (1) au plus résonnant (5). »

Les cinq extraits sonores sont présentés en ordre aléatoire pour chaque sujet et ils peuvent être réécoutés à loisir.

2.2 Extraits sonores

Pour mener à bien ces tests, des échantillons sonores possédant les mêmes caractéristiques sauf la réverbération doivent être obtenus. Pour cela, deux motifs (figure 1) constitués de tambour (grosses structures) et de cymbales (petites structures) ont été réalisés à partir de sons anéchoïques. Ces deux motifs possède une durée de 5 s mais des rythmes différents.

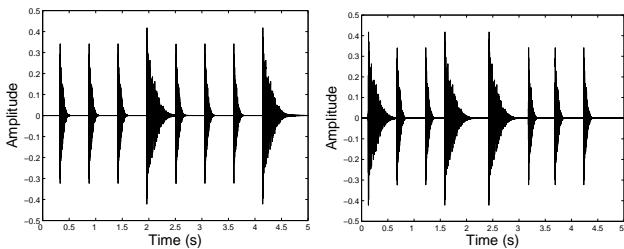


Figure 1 : Evolution temporelle de l'amplitude des motifs sonores anéchoïques : (gauche) motif 1, (droite) motif 2.

Ces échantillons sont ensuite auralisés [14] avec différentes réponses impulsionnelles de salle (RIR) et des fonctions de transfert de la tête (HRTF) pour permettre une écoute au casque réaliste. De tels systèmes d'auralisation ont déjà été utilisés avec succès pour la recherche en psychoacoustique [15,16]. Les RIR sont obtenues à partir d'échogrammes obtenus à l'aide du logiciel de tir de rayons large bande Salrev [17]. Des simulations sont donc réalisées pour une salle ($8 \times 10 \times 3 \text{ m}^3$) et différentes absorptions pour d'obtenir des Tr30 variant par pas de 10% (soit le double de la jnd pour prendre en compte l'inexpérience des sujets) : 1,6 s, 1,8 s, 2,0 s, 2,2 s et 2,4 s. Pour ces simulations, 10^6 rayons ont été émis et le calcul porte sur 200 pas de temps de 10 ms. Le récepteur sphérique de 0,25 m de diamètre est divisé en 26 angles solides [18].

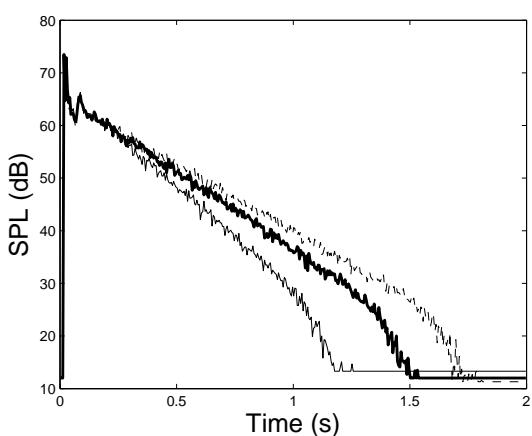


Figure 2 : Echogrammes obtenus pour différentes réverbérations : (—) $\text{Tr30}=1,6 \text{ s}$, (—) $\text{Tr30}=2,0 \text{ s}$, (---) $\text{Tr30}=2,4 \text{ s}$.

Pour obtenir des RIRs ne différant que par leur réverbération, la partie précoce, jusqu'à 80ms, est commune pour toutes les RIRs (celle obtenue pour un $\text{Tr30}=2,0 \text{ s}$) mais les parties tardives sont différentes (figure 2). Les échantillons auralisés ont une durée de 12 s (figure 3) avec une répétition du motif de base (figure 1).

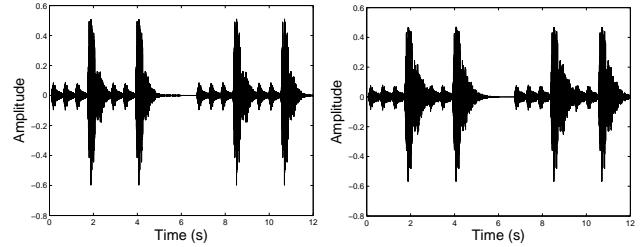


Figure 3 : Evolution de l'amplitude des échantillons sonores obtenus avec le motif 1 : (gauche) $\text{Tr}=1,6 \text{ s}$, (droite) $\text{Tr}=2,4 \text{ s}$.

2.3 Images

Pour obtenir cinq images évoquant une réverbération différente, 13 images ont été prétestées par 14 sujets en leur demandant de classer les images par rang de réverbération croissante. A l'analyse des résultats, un certain ordre se détachait. Pour valider ce résultat, des comparaisons par paires ont été effectuées (test de Wilcoxon) permettant de garder cinq images (figure 4) possédant un niveau de signification suffisant ($p<0,01$).



Figure 4 : Série d'images obtenues par ordre de réverbération croissante de (1) à (5).

2.4 Sujets

Deux types de sujets ont été testés : des étudiants en 5^{ème} année d'architecture (G1, G3 et G5) et des étudiants d'autres filières mais aussi en fin d'étude (G2, G4 et G6). Avant de clôturer cette expérience, les participants devaient remplir un questionnaire. Celui-ci investiguer plusieurs variables

- Données démographiques (age et sexe)
- Pratique de la musique
- Usage de différentes salles (amphithéâtres, salles de concert...). Il s'agit de variables que nous jugeons

- pertinentes pour évaluer l'expertise en qualité acoustique de salles.
- Habitude à lire des plans. Ce critère a pour but de vérifier la distinction entre le groupe architecte et autres.
 - Informations subjectives sur le déroulement de l'expérience.

70 participants (31 hommes et 39 femmes) ont réalisé cette tâche d'écoute et de classement d'extraits sonores. La moyenne d'âge était de 25,1 ans avec un écart type de 5,21 ans.

0	1	2	3
Contrôle, Son uniquement (motif 1)	Images évocatrices (motif 2)	Images distractives (motif 2)	Témoin, son uniquement (motif 2)
G1 : 10 architectes G2 : 17 divers	G1 : 10 architectes G2 : 17 divers		
G3 : 10 architectes G4 : 17 divers		G3 : 10 architectes G4 : 17 divers	
G5 : 8 architectes G6 : 8 divers			G5 : 8 architectes G6 : 8 divers

Table 1 : Plan expérimental.

3 Résultats

L'analyse des résultats porte sur deux variables : le temps d'exécution de la tâche et sa performance. La performance à cette tâche d'écoute est évaluée à l'aide du coefficient de corrélation de Spearman. Celui-ci exprime l'adéquation entre l'ordre théorique (les réponses correctes) et l'ordre observé (les réponses du sujet). Ce coefficient varie de -1 (inverse parfait de l'ordre correct) à +1 (ordre correct). Les différences entre groupes indépendants ont été éprouvées par des tests (non paramétriques) de Mann-Whitney et les différences intra participants (au sein d'un même groupe) par le test des Signes. Tous les tests ont été éprouvés au seuil de 5% en bilatéral.

3.1 Profil professionnel des sujets

Aucune différence significative n'est observée entre les groupes (architecture et autres filières) : ni sur le temps d'exécution, ni sur la performance. La manipulation fréquente de plans ou des vues 3D n'a donc pas d'influence sur la qualité du jugement de l'acoustique d'une salle. Cette variable « usage des plans » n'est pas discriminante. Les deux types de publics peuvent être considérés comme équivalents et les paires de groupes (G1 et G2, G3 et G4 ainsi que G5 et G6) peuvent être fusionnés pour les analyses ultérieures. Cependant cette analyse doit être nuancée car les résultats présentés concernent des sujets relativement jeunes, majoritairement des étudiants. En ciblant des publics de professionnels, des différences pourraient peut être apparaître entre les deux populations testées.

3.2 Profil personnel des sujets

Des données complémentaires sur les participants ont été récoltées pour déterminer si plusieurs profils se dégagent. Les variables prises en compte sont : la pratique d'un instrument de musique, la fréquentation de salles de concert ou théâtres, et enfin sa familiarité avec les concerts de musique non amplifiée.

La seule influence significative identifiée concerne le fait de jouer assez fréquemment un instrument. En effet, les participants sont significativement plus rapides pour classer les extraits lorsqu'ils font fréquemment de la musique par rapport à ceux qui en font peu ou pas ($p=0.002$). En

revanche, aucun impact significatif sur la performance n'a été observé. Les sujets étant habitués à des tâches d'écoute, la phase d'apprentissage n'est plus nécessaire. En outre, il est à noter que la variable « habitude de fréquentation des concerts de musique non amplifiée », sans doute très importante pour ce genre de tâche, ne donne pas de résultats. Cependant, la quasi-totalité des sujets ne sont pas des habitués, ne permettant pas de distinguer deux groupes différents pour ce facteur.

3.3 Effets des images

L'expérience de contrôle (0) ne révèle aucune différence significative entre les différents groupes (figure 5). Les résultats des différents groupes sont donc fusionnés pour cette expérience dans la suite.

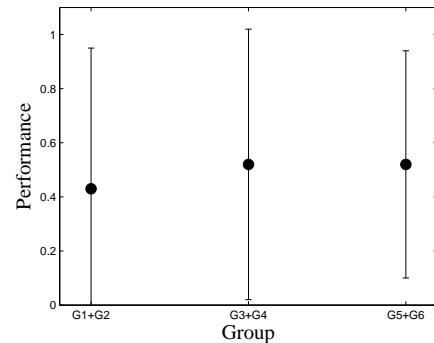


Figure 5 : Moyennes et écart-types des performances par groupe pour l'expérience de contrôle.

La figure 6 présente la seconde partie des tests avec les expériences 1 (G1+G2), 2 (G3+G4) et 3 (G5+G6) ainsi que le résultat de l'expérience 0 pour les six groupes. L'addition d'images évocatrices à l'écoute des échantillons sonores a un impact positif sur la performance des sujets ($p=0.003$). La performance moyenne des groupes G1+G2 évolue 0,52 à 0,75. En revanche, aucune amélioration de la performance n'est observée pour les autres groupes. La présence d'images distractives n'a donc pas d'effet sur la performance ($p=0,7$). Par ailleurs, l'expérience 3 ne met en avant aucun effet d'apprentissage particulier.

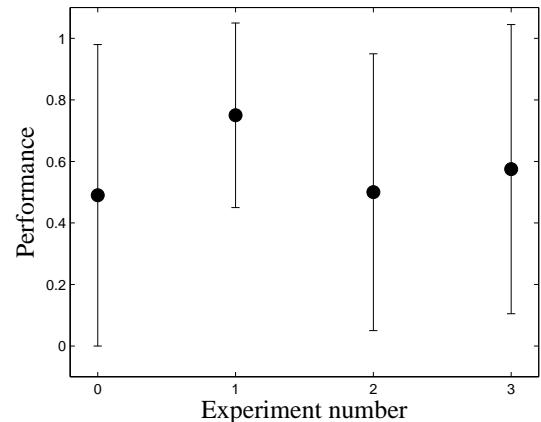


Figure 6 : Moyennes et écart-types des performances par expérience.

En plus d'obtenir de meilleures performances, les groupes G1+G2 ont besoin de moins d'écoutes lorsque les images évocatrices sont présentées pour effectuer leur choix (figure 7). Les résultats des groupes G3+G4 montrent aussi que la présence d'images incite les sujets à choisir plus

rapidement. Les groupes G5+G6 indiquent que ce résultat pas dû à un effet d'apprentissage : le nombre d'écoute reste identique pour les deux parties du test.

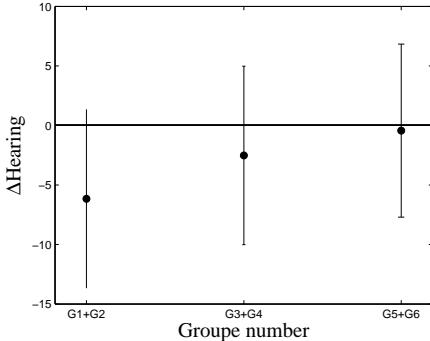


Figure 7 : Moyennes et écart-types de la différence du nombre d'écoutes entre la première partie du test et la seconde par groupe.

La présence d'image a un impact sur le nombre d'écoutes mais aussi sur la durée des écoutes (figure 8) pour les groupes G1+G2 ($p<0,0001$) ainsi que G3+G4 ($p=0,0002$). Lors de la deuxième partie du test, les groupes mis au contact d'images effectuent moins d'écoutes et sur un temps plus bref (de 9,5 s par échantillon sonore à 8,6 s pour G1+G2 et de 10 s à 8,3 s pour G3+G4). Cet effet n'a pas été constaté pour les groupes témoins G5+G6.

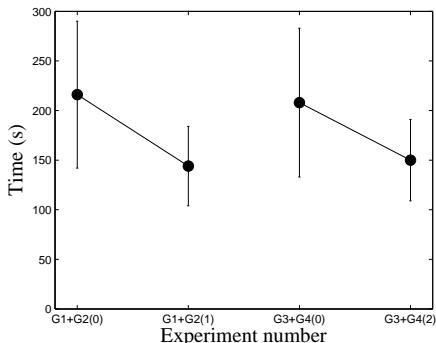


Figure 8 : Moyennes et écart-types de la durée des écoutes par sujets pour les deux parties du test pour les groupes G1+G2 et G3+G4.

La corrélation entre nombre d'écoutes et performance est non significative avec une corrélation maximale de $r<0,31$. Le nombre d'écoutes a peu d'effet sur la performance (figure 9).

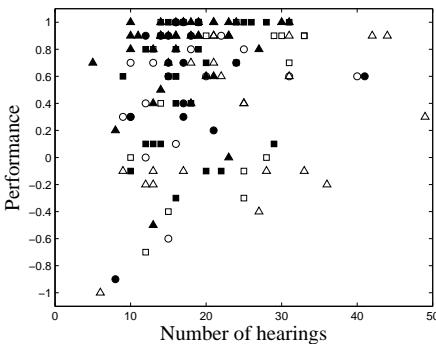


Figure 9 : Performance en fonction du nombre d'écoutes pour la première partie du test (symboles ouverts) et la seconde partie du test (symboles fermés) : (o, ●) G1+G2, (Δ, \blacktriangle) G3+G4, (\square, \blacksquare) G5+G6.

4 Discussion

L'apport d'un soutien visuel cohérent facilite, en performance et en rapidité, le jugement de la réverbération d'une salle. En revanche, la présence d'images incohérentes ne détériore pas la cette tâche d'évaluation. Les modalités auditives et visuelles peuvent donc être qualifiées de redondantes [10,11] : la tâche est réalisable avec la modalité auditive uniquement, mais l'information visuelle permet de désambiguer le message et d'apporter des informations complémentaires.

En revanche, la modalité auditive reste la modalité principale stimulée pour la tâche effectuée. La modalité visuelle peut être outrepassée si elle apporte des informations contradictoires et place le sujet en situation de conflit perceptif. Cette étude suggère donc qu'il est dans tous les cas intéressant d'afficher une image en soutien à l'évaluation auditive : dans le pire des cas, la performance ne sera pas détériorée et dans le meilleur des cas, elle pourra être améliorée.

Notre étude montre aussi qu'il est possible d'utiliser des sujets inexpérimentés pour l'évaluation perceptive de la réverbération à condition d'adapter l'échelle de variations. Ainsi, une variation entre les différents temps de réverbération de 10% au lieu de 5% recommandé [8] a permis un test significatif. Néanmoins, il serait intéressant d'étendre cette étude à une population dont l'oreille serait plus formée à l'acoustique des salles (acousticiens, architectes, musiciens, ingénieurs etc.). L'influence des images pourrait être toute différente pour ces publics.

5 Conclusion

Le développement actuel d'un environnement immersif permettant le travail coopératif entre architectes et acousticiens pose la question de l'impact de la modalité visuelle sur la tâche de l'évaluation perceptive de l'acoustique d'une salle. Une expérience a ici été réalisée pour étudier l'impact d'une image de salle sur le jugement de la réverbération. La réverbération est une des principales dimensions perceptives permettant le jugement et la discrimination de l'acoustique d'une salle. Le critère objectif lié à cette dimension perceptive est le temps de réverbération.

Différents échantillons sonores dont seules la partie tardive varie ont été réalisées. Similairement, une série d'images de salles évoquant des réverbérations différentes a été sélectionnée. Des tâches de classement par réverbération croissante ont ensuite été demandées à des sujets inexpérimentés avec ou sans appui des images.

La présence d'images cohérentes améliore la performance des sujets et accélère cette tâche de classement : les sujets classent mieux et plus vite. L'information visuelle vient en renfort de l'ouïe et permet de désambiguer le choix. La présence d'images incohérentes n'a pas d'effet sur la performance. La modalité auditive reste la modalité principale stimulée. En revanche, les sujets ont tendance à effectuer le classement de manière plus rapide. Une expérience témoin permet de montrer que les phénomènes observés ne sont pas dus à un effet d'apprentissage.

Enfin, cette étude montre que le jugement perceptif de la réverbération, concept facile à appréhender, peut être effectué par des sujets inexpérimentés sans apprentissage particulier en adaptant l'échelle de variations.

Remerciements

Les auteurs remercient la région Wallonne pour son soutien financier (convention n°616416) au projet Auralias (<http://www.auralias.be>).

Références

- [1] Bos L., Embrechts J.-J. "An interactive real-time based auralization system for room acoustics implementing directional impulse response and multiple audio reproduction modules for spatialization (the AURALIAS project)", Proc. of NAG/DAGA 2009, Rotterdam, p. 1459-1462 (2009).
- [2] Beranek L. "Concert and opera halls : How they sound", AIP Press (1996).
- [3] Barron M. "Auditorium acoustics and architectural design", E & FN Spon, London (1993).
- [4] Zahorik P. "Perceptually relevant parameters for virtual listening simulation of small room acoustics", J. Acoust. Soc. Am. 126(2), 776-791 (2009).
- [5] Cerdá S., Giménez A., Romero J., Cibrián R., Miralles J.L. "Room acoustical parameters: A factor analysis approach", Appl. Acoustics 70(1), 97-109 (2009).
- [6] Schroeder M.R., Gottlob D., Siebrasse K.F. "Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters", J. Acoust. Soc. Am. 56(4), 1195-1201 (1974).
- [7] Seraphim, H. P. "Untersuchungen über die Unterschiedsschwelle exponentiellen Abklingens von Rauschbandimpulsen", Acustica 8, 280-284 (1958).
- [8] Vorlander M. "International round robin on room acoustical computer simulations", 15th International Congress on Acoustics, Trondheim (1995).
- [9] Cox T.J., Davies W.J., Lam Y.W. "The sensitivity of listener to early sound field changes in auditoria", Acustica 79(1), 27-41 (1993).
- [10] Stein B.E., Meredith M.A. "The merging of senses", MIT Press, Cambridge MA (1997).
- [11] Delorme A. "L'organisation perceptive", dans A. Delorme et Flückiger (Eds.) "Perception et réalité. Introduction à la psychologie des perceptions", De Boeck Université, Bruxelles. (2003).
- [12] Cabrera D., Nguyen A., Choi Y.J. "Auditory versus visual spatial impression: A study of two auditoria", Proceeding of ICAD, Sydney. (2004).
- [13] Valente D.L., Braasch J. "Subjective expectation adjustements of early-to-late reverberant energy ratio and reverberation time to match visual cues of a musical performance", Acta Acustica united with Acustica 94(6), 840-855 (2008).
- [14] Vorländer M. "Auralization", Springer, Berlin (2007).
- [15] Blauert J., Lehnert H., Sahrhage J., Strauss H. "An interactive virtual-environment generator for psychoacoustic research. I: Architecture and implementation", Acta Acustica united with Acustica 86(1), 94-102 (2000).
- [16] Djelani T., Pörschmann C., Sahrhage J., Blauert J. "An interactive virtual-environment generator for psychoacoustic research. II: Collection of head-related impulse responses and evaluation of auditory localization", Acta Acustica united with Acustica 86(6), 1046-1053 (2000).
- [17] Embrechts J.-J. "Broad spectrum diffusion model for room acoustics ray-tracing algorithms", J. Acoust. Soc. Am. 107(4), 2068-2081 (2000).
- [18] Embrechts J.-J., Werner N., Lesoinne S. "Auralization in room acoustics using directional impulse responses computed by sound ray techniques", 4th Forum Acusticum, Budapest, 2535-2538 (2005).