

Environnements Virtuels Sonorisés

Nicolas Tsingos
REVES-INRIA Sophia Antipolis
Nicolas.Tsingos@sophia.inria.fr
<http://www-sop.inria.fr/reves>

Si la visualisation 3D bénéficie d'un développement privilégié pour les applications de réalité virtuelle, le son en est un autre composant incontournable. Les techniques de spatialisation du son permettent aujourd'hui de simuler des sources sonores virtuelles placées arbitrairement dans l'espace autour de l'auditeur. Sur casque ou enceintes, elles permettent une immersion décuplée dans les environnements de synthèse tout en restant naturelles et peu intrusives pour l'utilisateur.

Les systèmes de réalité virtuelle interactifs combinent des représentations visuelle, sonore et haptique, afin de simuler de manière immersive l'exploration d'un monde tridimensionnel représenté depuis le point de vue d'un observateur contrôlé en temps réel par l'utilisateur. La plupart des travaux effectués dans ce domaine ont historiquement porté sur les aspects visuels (par exemple des méthodes d'affichage interactif de modèles 3D complexes ou de simulation réaliste et efficace de l'éclairage) et relativement peu de travaux ont été consacrés à la simulation de sources sonores virtuelles ou *auralisation*. Il est pourtant certain que la simulation de sources sonores virtuelles est un facteur clé dans la production d'environnements de synthèse, la perception sonore s'ajoutant à la perception visuelle pour produire une interaction plus naturelle.

En particulier, les effets sonores spatialisés, dont la direction de provenance est fidèlement reproduite aux oreilles de l'auditeur, sont particulièrement importants pour localiser les objets, séparer de multiples signaux sonores simultanés et donner des indices sur les caractéristiques spatiales de l'environnement (taille, matériaux, etc.). La plupart des systèmes de réalité virtuelle immersifs, des simulateurs les plus complexes aux jeux vidéo destinés au grand public mettent aujourd'hui en œuvre des algorithmes de synthèse et spatialisation des sons qui permettent d'améliorer la navigation et d'accroître le réalisme et la sensation de présence de l'utilisateur dans l'environnement de synthèse.

Comme la synthèse d'image dont elle est l'équivalent auditif, l'auralisation, appelée aussi *rendu sonore*, est un vaste sujet à la croisée de multiples disciplines : informatique, acoustique et électro-acoustique, traitement du signal, musique, calcul géométrique mais également psycho-acoustique et perception audio-visuelle. Il regroupe trois problématiques principales : synthèse et contrôle interactif de sons, simulation des effets de propagation du son dans l'environnement et enfin, perception et restitution spatiale aux oreilles de l'auditeur.

Historiquement, ces trois problématiques émergent de travaux en acoustique architecturale, acoustique musicale et psycho-acoustique. Toutefois une différence fondamentale entre rendu sonore pour la réalité virtuelle et acoustique réside dans l'interaction son-image et dans l'efficacité des algorithmes devant être mis en œuvre pour des applications interactives. Ces aspects importants contribuent à en faire un domaine à part qui prend une importance croissante, tant dans le milieu de l'acoustique que dans celui de la synthèse d'image/réalité virtuelle.

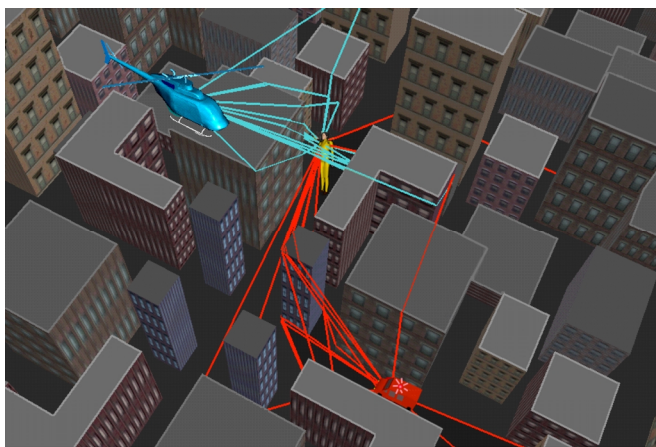


Figure 1 – Simulation de chemins de propagation du son d'un hélicoptère et d'un camion dans un environnement urbain. La réflexion et diffraction du son sur les obstacles est évaluée par une technique de lancer de faisceaux.

La première composante d'un système de rendu sonore pour la réalité virtuelle est la synthèse de sources : la production et la manipulation interactive de sons et la modélisation des caractéristiques physiques des sources sonores. L'application principale en est la production de sons résultant de l'interaction de l'utilisateur avec l'environnement comme les bruits d'impact, de contact entre objets, etc. Suivant les effets voulus, on pourra travailler avec des sons enregistrés ou purement synthétiques mais, en général, les sons utilisés doivent être les plus « bruts » possibles (sans réverbération par exemple). Cette première problématique est très proche de la synthèse musicale par ordinateur bien que les sons considérés ne soient en général pas musicaux. Toutefois, elle comprend également un aspect de contrôle basé sur l'interaction 3D de l'utilisateur avec l'environnement qui lui est propre.

Une fois synthétisés ou enregistrés, les signaux sonores sont ensuite traités de manière à reproduire les effets de propagation du son dans l'environnement de synthèse (occlusions par les obstacles, réflexions sur les parois et réverbération, effet Doppler pour les sources en mouvement). Là encore, ces effets dépendent fortement de la géométrie de l'environnement, de la position des sources et de l'auditeur. Les traitements effectués sont donc contrôlés par des calculs géométriques (Figure 1) très proches de ceux effectués pour la synthèse d'image (du lancer de rayon par exemple) mais ont également leurs particularités comme le traitement de la diffraction du son par les obstacles qui est généralement négligeable pour la lumière. Pour les applications où le « réalisme » plus que la précision physique de la simulation est suffisant, des modèles perceptifs peuvent être également utilisés pour décrire les traitements à appliquer aux sons sans recourir à des simulations géométriques trop coûteuses en temps de calcul.

Enfin, la dernière partie du processus est la restitution du son aux oreilles de l'auditeur de manière la plus fidèle possible. Au delà des notions de *fidélité* généralement admises en audio (précision des calculs et de la représentation numérique des signaux, qualité de l'électronique et des hauts-parleurs), l'acoustique virtuelle s'intéresse tout particulièrement à la fidélité de la restitution spatiale du son, la *spatialisation*. La spatialisation du son est un traitement spécifique qui vise à reproduire aux oreilles de l'auditeur la sensation correcte de la direction de provenance des sons virtuels. C'est l'un des composants les plus importants dans la production de son pour la réalité virtuelle. On peut le comparer à la production d'images en relief pour le canal visuel. Depuis la mise au point des premiers systèmes de restitution stéréophonique et multi-canal dans les années 1930-1940, la reproduction spatiale du son a été un enjeu majeur de la recherche en acoustique et électro-acoustique. Longtemps cantonné aux laboratoires de recherche ou aux amateurs de musique expérimentale, le son spatial a récemment effectué une percée fulgurante chez le grand public au travers du son « 3D » ou *surround* (limité dans le plan horizontal) pour les jeux vidéos et le « home-cinéma ». De nombreux systèmes de restitution spatiale du son ont été développés et étudiés durant les 30 dernières années. La référence reste la stéréo *binaurale* qui permet de reproduire à l'aide d'un simple casque stéréo et d'un filtrage spécifique des signaux, la direction de provenance d'un son synthétisé ou enregistré. Ce système permet de reproduire aussi bien des sons provenant de l'avant que de l'arrière, du dessus ou du dessous de l'auditeur. Néanmoins, la qualité du résultat est très dépendante de l'auditeur puisque les filtres à appliquer sont fortement dépendants de la morphologie de la personne. D'autres systèmes comme l'*holophonie*, équivalent acoustique de l'holographie, multiplient le nombre d'enceintes utilisées pour reproduire le plus fidèlement possible les ondes sonores physiquement correctes dans le local d'écoute. Enfin des systèmes, dits *transauraux*, combinant filtrages spécifiques et restitution sur enceintes permettent d'obtenir un bon compromis entre qualité des résultats et étendue de la zone d'écoute optimale.

De nombreuses optimisations visant à réaliser des algorithmes rapides permettant de traiter des scènes très complexes, comprenant des centaines voire milliers de sources sonores, sont également au cœur du problème. Parmi les plus récentes avancées, on peut mentionner les techniques de synthèse et spatialisation sonore perceptives. Ces techniques, au croisement de la psycho-acoustique, de la compression audio perceptive (proches du standard *mp3* par des aspects d'évaluation de masquages acoustiques) et de l'analyse de scène sonore, introduisent des concepts de rendu à différents niveaux de détail, similaires à ceux développés pour le rendu graphique (Figure 2). Elles vont très certainement permettre de réaliser des simulations sonores de très haute qualité pour des scènes complexes dans un futur très proche. Elles permettront également de traiter des environnements distribués dans le cadre de « chat-rooms » ou de jeux en ligne massivement multi-joueurs.

Au cours des quelques dernières années, s'est également développée une problématique nouvelle dans le cadre du rendu audio-visuel pour les environnements virtuels : l'interaction multi-modale son/image. Cette problématique vise à étudier comment les aspects auditifs et visuels s'influencent mutuellement dans notre perception de l'expérience offerte à l'utilisateur. Au final, de meilleurs algorithmes de rendu intégrant de manière plus unifiée les aspects visuels et sonores pourraient voir le jour.

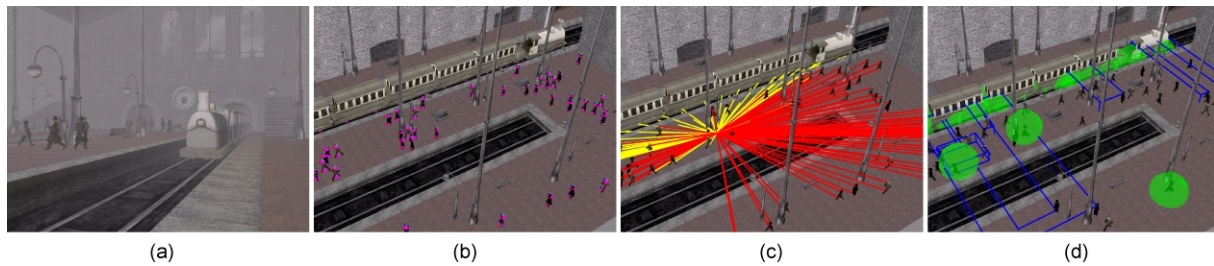


Figure 2 – (a) Une application d'un système de rendu audio perceptif à un environnement complexe, ici une gare. (b) Chaque piéton dans la gare est modélisé par deux sources sonores (sa voix et ses bruits de pas). Chaque roue du train est aussi modélisée comme une source sonore pour restituer l'impression d'une source spatialement étendue. Au final, 160 sources sonores doivent être simulées en temps réel (en magenta). (c) Les lignes colorées représentent les chemins de propagation directs entre les sources et l'auditeur (au centre). Les lignes en rouges représentent les sources sonores masquées par les limitations de notre système auditif. Celles en jaunes sont audibles. Dans cette figure, on remarque que le bruit du train, beaucoup plus fort, masque les conversations et les bruits de pas des piétons. (d) Des groupes sont automatiquement construits pour spatialiser les sources sonores audibles. Plusieurs sources (dans les boîtes bleues) sont ainsi remplacées par une source sonore perceptivement équivalente (en vert).

Pour en savoir plus

Ouvrages

- [1] D.R. Begault. *3D Sound for Virtual Reality and Multimedia*. Academic Press Professional, 1994.
- [2] J. Blauert. *Spatial Hearing : The Psychophysics of Human Sound Localization*. M.I.T. Press, Cambridge, MA, 1983.
- [3] A.S. Bregman. *Auditory Scene Analysis, The perceptual organization of sound*. M.I.T Press, Cambridge, MA, 1990.
- [4] M. Kahrs et K. Brandenburg Ed., *Applications of Digital Signal Processing to Audio and Acoustics*, Kluwer Academic Publishers, 1998.
- [5] Perry Cook. *Real Sound Synthesis for Interactive Applications*, AK Peters Ltd., 2002.
- [6] David Luebke, Martin Reddy, Jonathan D. Cohen, Amitabh Varshney, Benjamin Watson, et Robert Huebner. *Level of Detail for 3D Graphics*. Morgan Kaufmann Publishing, 2002.

Articles scientifiques

- [7] M. Kleiner, B.I. Dalenbäck et P. Svensson, *Auralization - An Overview*, Journal of the Audio Engineering Society 41(11), Nov. 1993.
- [8] L. Savioja, J. Huopaniemi, T. Lokki et R. Väänänen. *Creating interactive virtual acoustic environments*. Journal of the Audio Engineering Society, 47(9), Sept. 1999.
- [9] T. Funkhouser, J.M. Jot et N. Tsingos, *Sounds good to me ! Computational sound for graphics, VR, and interactive systems*, SIGGRAPH 2002 course #45, 2002.
- [10] N.Tsingos, E. Gallo et G. Drettakis. *Perceptual audio rendering of complex virtual environments*, à paraître à SIGGRAPH 2004, Los Angeles, Août 2004.
- [11] K. Van Den Doel, K. Kry et D.K. Pai. *FOLEYAUTOMATIC: Physically-based sound effects for interactive simulation and animation*. SIGGRAPH 2001.
- [12] J.F. O'Brien, P. Cook et G. Essl. *Synthesizing sounds from physically-based motion*. SIGGRAPH 2001.
- [13] K. Van Den Doel, D. Knott and D.K. Pai, *Interactive Simulation of Complex Audio-Visual Scenes*, Presence: Teleoperators and Virtual Environments 13(1), 2004.
- [14] Y. Dobashi, T. Yamamoto et T. Nishita, *Real-time rendering of Aerodynamic Sound using Sound Textures based on Computational Fluid Dynamics*, ACM Transactions on Graphics 22(3), Proceedings of SIGGRAPH 2003, Août 2003.

Programmation Audio

- [15] Soundblaster, © Creative Labs. <http://www.soundblaster.com>
- [16] Direct X homepage, © Microsoft. <http://www.microsoft.com/windows/directx/default.asp>
- [17] Environmental audio extensions: EAX © Creative Labs. <http://www.soundblaster.com/eaudio>, <http://developer.creative.com>
- [18] EAGLE © Creative Labs. <http://developer.creative.com>
- [19] ZoomFX, MacroFX, ©Sensaura. <http://www.sensaura.co.uk>