

# Nouvelles techniques de visualisation et d'interaction en composition spatiale pour les interfaces de mixage audio dans le contexte de la spatialisation sonore 3D.

**Laboratoire(s) :** Université Paris-Sud, CNRS & Inria (équipe in|situ|), LIMSI-CNRS, CIRMMT – McGill University.

**Encadrement de la thèse :** Stéphane Huot (Stephane.Huot@lri.fr, LRI), Brian FG Katz (Brian.Katz@limsi.fr, LIMSI/CNRS).

**Collaborations :** Marcelo Wanderley (Marcelo.Wanderley@mcgill.ca, McGill University), Alan Blum (al.blum@ens-louis-lumiere.fr, ENS Louis Lumière).

## Introduction

Ce projet consiste à étudier et concevoir de nouvelles interfaces, 'physiques' et logicielles, pour le mixage audio spatialisé. Les avancées permanentes des méthodes et équipements de diffusion pour le son spatialisé, souvent hétérogènes, entraînent de nouveaux besoins pour lesquels les équipements de mixage actuels (consoles et logiciels) s'avèrent inadaptes ou tout simplement inutilisables :

- placement, manipulation et gestion du mixage de sources audio dans un environnement de restitution en 3 dimensions ;
- interface identique pour la réalisation de bandes son spatialisées cohérentes quel que soit le système d'écoute lors du mixage (binaural par exemple) et de diffusion lors du rendu (22.2 par exemple) ;
- outils pour la compréhension, modification et adaptation de mixages existant pour des systèmes de spatialisation différents (e.g., écoute d'un mixage 7.1 et son 'downmix' stéréo sans perte de sources sonores) ;
- interface mobile cohérente avec l'interface utilisée en studio pour la vérification du mixage et la configuration des systèmes de diffusion spatialisés sur le terrain.

Bien que certains de ces points aient déjà été en partie abordés par la recherche et l'industrie (voir Fig. 1), notamment le premier sur le placement et la manipulation de sources sonores, ces outils n'en sont qu'à leurs prémices, s'adressant pour la plupart à la spatialisation sur un même plan (de type 5.1 ou 7.1 par exemple). Il est pourtant nécessaire d'étudier une approche plus générale, destinée à la spatialisation en 3 dimensions (diffusion sur plusieurs plans sol/murs/plafond, comme par exemple les systèmes de type '22.2', Dolby Atmos, ou Ambisonics).

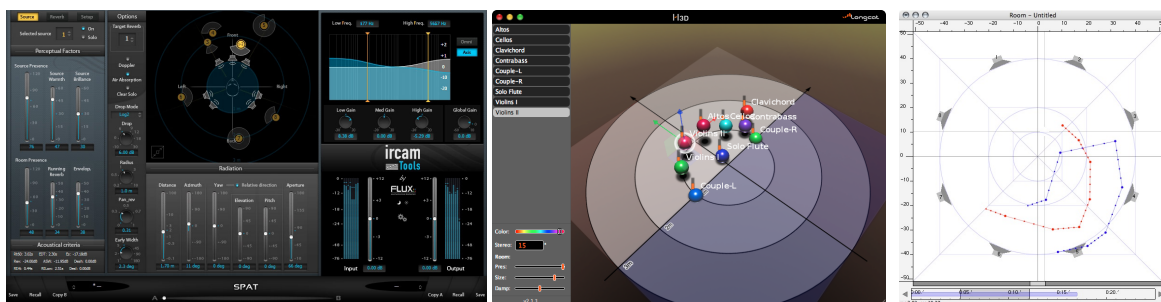


Figure 1 – le logiciel SPAT (IRCAM) [1], le plugin de spatialisation binaurale H3D (Longcat) [2], et le logiciel Holophon [3].

Enfin, étant généralement construits autour d'une solution uniquement logicielle à base de vues en 2D1/2 ou 3D pour représenter et manipuler des sources dans l'espace sonore, ils n'offrent que peu de contrôle sur les paramètres du mixage, en comparaison avec une console physique. De plus, cette différence de représentation entre la vue 3D des sources/sorties et les contrôles de routage/mixage d'une console (bus linéaires) est probablement source d'incohérences qui peuvent rendre inadaptee l'utilisation conjointe de ces approches dans un environnement de mixage professionnel.

### Objectif de la thèse

L'objectif de cette thèse est donc de répondre à cette problématique par une approche centrée sur l'utilisateur, en concevant, réalisant et évaluant de nouvelles techniques de visualisation et d'interaction pour le mixage spatialisé en 3D. Les travaux seront bien sûr basés sur l'existant dans le domaine (des consoles de mixage standard aux logiciels pour la spatialisation sur un même plan [4]), mais surtout sur l'étude des pratiques et méthodes de travail des ingénieurs du son, dans le but de proposer des solutions adaptées aux nouvelles problématiques du mixage 3D. Cette étude servira ainsi de base à l'exploration des possibilités offertes dans ce contexte par de nouveaux dispositifs de visualisation et d'interaction (écrans et tablettes haute-résolution et/ou multi-tactiles, dispositifs à degrés de liberté multiples, interfaces tangibles [5], etc.), voire à la conception de dispositifs originaux appropriés, et à leur association cohérente avec tout ou partie des outils usuels pour le mixage. Ces propositions donneront lieu à la réalisation de prototypes dont l'évaluation en conditions d'usage permettra d'évaluer l'adéquation aux problématiques.

### Méthodologie et déroulement de la thèse

La première partie de la thèse consistera en l'étude de l'état de l'art des techniques de visualisation et d'interaction existantes pour l'édition de sons, et la manipulation et l'édition de mixages audio pour les environnements sonores 3D (domaines de l'informatique musicale, des environnements immersifs, de la visualisation d'information, etc.). En parallèle, le candidat procédera à l'étude des outils actuels et à l'analyse des premiers besoins pour leur utilisation pour le mixage 3D interactif.

Cette étude sera complétée dans un second temps par la création de scénarios d'usages par des séances de conception participative et de prototypage en collaboration avec des utilisateurs cibles de ces outils. Ce travail permettra l'identification des besoins (par exemple, différentes phases de travail nécessaires au mixage 3D, outils existants ou pouvant être adaptés pour une tâche particulière, etc.), ainsi que les premières pistes de solutions issues des compétences et habitudes des utilisateurs.

En se basant sur ces études, le candidat proposera et développera des prototypes de techniques de visualisation et d'interaction qui seront confrontés aux utilisateurs et validés par un processus itératif d'évaluations contrôlées (en laboratoire) et sur le terrain avec des ingénieurs du son professionnels (au LIMSI, à l'ENS Louis Lumière ou au CIRMMT à Montréal). Dans un cadre plus général, ces travaux pourront déboucher sur la proposition de règles et d'un espace de conception pour les techniques avancées de visualisation et de manipulation de sons spatialisés dans le cadre de l'équipe internationale associée Inria MIDWAY (coordonnée par S. Huot et M. Wanderley).

**Prérequis :** cursus en Interaction Homme-Machine, Informatique Musicale, et/ou Visualisation d'Informations (conception, évaluation et développement des systèmes interactifs). Bon niveau en programmation (C/C++, Java, Objective-C). Un Intérêt pour les métiers du son et du cinéma, ainsi que des connaissances en électronique et/ou conception de dispositifs physiques seraient un plus.

## Références

- [1] Plugin SPAT, IRCAM, [http://www.fluxhome.com/products/plugins/ircam\\_spat](http://www.fluxhome.com/products/plugins/ircam_spat).
- [2] Plugin H3d, Longcat, <http://www.longcat.fr/>.
- [3] Holophon, <http://www.gmem.org/>.
- [4] Delerue, O. et Warusfel, O., Authoring of virtual sound scenes in the context of the Listen Project, *Actes de la conférence AES22 (22nd Conference of the Audio Engineering Society)*, Helsinki, Finlande, 2002.
- [5] Jordà, S., Kaltenbrunner, M., Geiger, G. et Bencina, R., "The reacTable\*", *Proceedings of the International Computer Music Conference (ICMC2005)*, Barcelona (Spain).