



IRCAM – UMR 9912
EQUIPE ACOUSTIQUE DES SALLES
RAPPORT D'ACTIVITES - 2004



- Responsable :
Olivier Warusfel
- Thésards :
Etienne Corteel, (Paris 6)
Terence Caulkins, (Paris 6)
Sylvain Busson, (Marseille)
- Collaborations principales:
Brian Katz (LIMSI)
Nicolas Misdariis (IRCAM, équipe Design Sonore)
Isabelle Viaud-Delmon (UMR 7593)
Ludivine Sarlat (UMR 7593)
- Autres collaborations:
FT R&D, SonicEmotion
- Ingénieurs/Chercheur
Olivier Delerue (Projet SHF)
Emmanuel Rio (Projet Opera)

L'activité de recherche et de développement de l'équipe Acoustique des Salles est centrée principalement sur l'analyse, la reproduction et la synthèse de scènes sonores. Il s'agit de proposer des techniques permettant de reproduire ou synthétiser l'organisation spatiale d'une scène sonore, notion qui recouvre la localisation des sources dans l'espace et la signature acoustique de la salle ou de l'environnement. L'activité se répartit entre des études à caractère théorique, basées sur l'acoustique physique et le traitement du signal (reproduction / transformation de champs sonores), des développements informatiques liés à l'exploitation de ces technologies pour les applications de spatialisation, et des études perceptives ayant principalement pour objectif de valider la pertinence perceptive des développements technologiques réalisés. Le travail sur la perception s'est récemment enrichi de problématiques remplaçant la perception auditive spatiale dans un cadre plus général de perception spatiale multi-sensorielle, incluant notamment les interactions avec la vision ou le mouvement dans l'espace.

L'année 2004 a été une année de transition, marquée simultanément par la fin des projets européens LISTEN et CARROUSO qui avaient permis à l'équipe d'initier de nouvelles activités autour des installations sonores et de la Wave Field Synthesis et par le démarrage des projets OPERA et SEMANTIC HiFi.

Sur le plan scientifique, les activités majeures ont concerné la poursuite du travail sur la technique binaurale, la Wave Field Synthesis et le développement d'interfaces pour le contrôle des applications de spatialisation. Les travaux sur la reproduction binaurale ont été réalisés en partenariat avec France Télécom (thèse de Sylvain Busson) et avec le LIMSI et sont centrés sur l'analyse des dépendances inter-individuelles. Le volet Wave Field Synthesis a été mené dans le cadre des thèses d'Etienne Corteel et de Terence Caulkins. Ces travaux ont notamment abouti à une formulation précise des caractéristiques du champ rayonné par un banc haut-parleurs et leurs conséquences sur les procédures d'égalisation ou de compensation de contexte. Par ailleurs, un travail théorique et pratique a été consacré à l'extension de cette méthode de reproduction pour la synthèse de la directivité associée à la source.

En parallèle des activités de recherche, les travaux de développement se sont poursuivis avec notamment l'intégration des résultats de recherche sur la Wave Field Synthesis sous forme de plug-ins étendant la palette des formats audio supportés par le Spatialisateur. Par ailleurs, à l'occasion du projet Semantic HiFi des

travaux de portage du spatialisateur et de MusicSpace ont été entrepris et permettent d'envisager, à terme, des applications intégrant ces différents modules.

Une importante activité de production musicale a été effectuée au cours de l'année afin de valider les travaux de recherche et de développement réalisés pendant ces derniers mois. La constitution de ce répertoire d'esquisses musicales a fourni l'occasion d'un premier retour utilisateur, en situation de production, et permet à l'équipe de disposer de matériaux de démonstrations de ces nouvelles technologies auprès du public et de futurs utilisateurs, compositeurs ou ingénieurs du son.

1.1.1. TECHNOLOGIES DE REPRODUCTION BINAURALES

1.1.1.1. Analyse de la dépendance spatiale du retard interaural

Un problème courant dans le développement et l'amélioration de la restitution binaurale est la résolution des variations interindividuelles. Dans le projet long terme de « l'adaptation individuelle pour le rendu binaural » une étude a été lancée sur les variations de l'indice de base de la localisation, l'ITD, qui ne sont qu'imparfaitement traduites par les modèles classiques. Jusqu'à présent, la plupart des modèles d'ITD sont basés sur une tête sphérique et axisymétrique. Par ailleurs, les variations dépendent non seulement de l'individu, mais la question se pose également de la méthode d'estimation de l'ITD la plus pertinente du point de vue perceptif.

L'étude a mis en évidence les détails de la variation spatiale d'ITD liée aux différents chemins de propagation autour de la tête. Ces résultats seront également utilisés pour la création d'un modèle morphologique d'ITD reposant sur une équation ne dépendant que du paramètre « circonférence » de la tête. Ce modèle issu d'observation statistique à partir de la base de données LISTEN apporte une amélioration significative des performances par rapport aux modèles classiques. Une publication scientifique issue de cette étude est en cours de rédaction.

Collaboration extérieure : B. Katz (LIMSI)

1.1.1.2. Selection des HRTFs non individuelles

Dans le cas où l'auditeur d'une scène binaurale ne dispose pas de ses HRTFs il est nécessaire de lui proposer un jeu aussi proche que possible des siennes. Différents travaux avaient été réalisés dans l'équipe (cf thèse de Véronique Larcher) à partir de l'analyse objective des caractéristiques signal et visant à dégager le jeu minimum de HRTFs, associées à des fonctions spatiales permettant d'engendrer l'ensemble des HRTFs d'une population. Cependant cette étude objective doit être complétée par une approche permettant de dégager les critères de « proximité » perceptive afin de piloter la sélection. Sur la base de données de mesures effectuées à l'IRCAM dans le projet LISTEN, un test perceptif a été mis en place auquel 46 personnes ont pu participer. Chaque sujet a eu pour tâche, d'évaluer subjectivement la qualité de reproduction offerte par les HRTFs des autres individus. Le dépouillement des résultats devrait permettre de dégager une métrique perceptive entre HRTFs. L'analyse des données est en cours, cependant des résultats préliminaires montrent qu'il est possible de composer une base de données des HRTFs réduite, avec 5 ou 6 éléments, qui satisfassent « perceptivement » 95% des participants. Les prochaines étapes, consisteront à établir les liens entre ces éléments de base et les paramètres morphologiques.

Participants : S. Winsberg

Collaboration extérieure : B. Katz (LIMSI)

1.1.1.3. Evaluations des systèmes de mesure pour les HRTFs

Différentes bases de données d'HRTFs existent dans le monde mais reposent sur des configurations expérimentales de mesures différentes, qu'il s'agisse des composants matériels utilisés ou des conditions de mesure (réponse acoustique du lieu, etc.). Ces variations posent la question de l'universalité des HRTFs mesurées et de leur utilisation par les différents laboratoires. En collaboration avec Durand R. Begault du « NASA Ames Research Center », une proposition a été faite à différents laboratoires mondiaux d'effectuer des mesures sur une même tête artificielle (NASA). Cette étude, permettra d'appréhender les différences des HRTFs dues à la variabilité des systèmes de mesures. A la fin de 2005 nous espérons avoir collecté les réponses d'une dizaine de laboratoires réparties en Europe, aux Etats-Unis, et en Extrême-Orient.

Collaboration extérieure : B. Katz (LIMSI)

1.1.1.4. Validation perceptive de modèles de retard interaural

L'analyse des propriétés des filtres binauraux ainsi que leur implémentation reposent généralement sur une séparation de l'information de retard interaural (ITD) et de l'information spectrale consignée dans le filtre à phase minimale associée. Différentes méthodes d'estimation du retard interaural existent sans toutefois conduire à des valeurs identiques. Afin de connaître les performances des estimateurs de l'ITD, du point de vue de leur pertinence subjective, une étude psychoacoustique est actuellement menée pour déterminer dans les valeurs de l'ITD subjectif pour différentes directions de l'espace. La tâche des sujets consiste à apparier une HRIR (Head Related Impulse Response) de référence avec un modèle constitué de la réponse à phase minimale extraite de cette HRIR et d'un ITD ajustable par pas de 10µsec. La valeur moyenne relevée est ensuite comparée à celles livrées par les différents estimateurs.

Les premières écoutes ont mis en évidence la difficulté de faire correspondre les deux sons pour des positions dont l'azimut est supérieur à 60°. Ceci est conforme à la littérature : les articles décrivant l'implémentation ITD+modèle à phase minimale reportent une parfaite similitude sauf pour les positions latérales. Pour ces positions, il semble difficile d'associer un filtre à phase minimum au HRIR correspondantes (réflexions précoces sur les épaules). La phase résiduelle alors négligée serait de nature non linéaire.

Participants : S. Busson (thèse Université de Marseille)

Collaborations extérieures : R. Nicol (FT R&D), P.O. Mattei (LMA)

1.1.1.5. Seuil de résolution auditive du retard interaural

En complément de l'étude consacrée aux estimateurs de l'ITD (cf. 1)1.1.1.4) une expérience a été réalisée afin de déterminer si les variations observées entre estimateurs sont détectables auditivement. L'expérience était également motivée par le souci d'évaluer si les variations objectives de l'ITD sur les cônes de confusion sont pertinentes subjectivement par conséquent si elles doivent être reproduites fidèlement. Si ce n'est pas le cas et que l'utilisation d'un modèle simplifié de l'ITD est suffisante, le coût d'implémentation de la synthèse binaurale sera réduit. Bien que différentes études aient déjà été consacrées à l'évaluation du seuil de perception du retard interaural, les conditions expérimentales apparaissent comme un facteur relativement déterminant, notamment selon le type de stimulus auditif mais également selon les conditions d'écoute (dichotique, champ libre, tête libre ou non, ...). L'étude menée ici se particularise par le modèle d'implémentation utilisé (retard pur associé à un filtre à phase minimale). La méthode adaptative de choix forcé à trois alternatives avec retour des réponses a été utilisée pour mesurer ces seuils.

L'étude menée a mis en évidence des seuils de perception de l'ordre de 40 µsec, supérieurs à ceux des études précédentes réalisées en champs libre ou en écoute au casque avec spatialisation par retard pur. L'interprétation de ces résultats est en cours. Dans l'état actuel il semble donc que le modèle d'ITD basé sur une approximation axisymétrique de la tête soit suffisant au regard des performances de résolution auditive, sauf pour les directions latérales pour lesquelles les variations objectives sont supérieures. L'introduction du paramètre de décalage de l'axe interaural permet de rentrer dans les tolérances.

Participants : S. Busson (thèse Université de Marseille)

Collaborations extérieures : R. Nicol (FT R&D), P.O. Mattei (LMA)

1.1.2. REPRODUCTION DE CHAMPS SONORES – WAVE FIELD SYNTHESIS

La technique Wave Field Synthesis [WFS] a l'ambition de capter, reproduire ou synthétiser un champ sonore préservant les propriétés physiques liées à la position des sources qui la composent. Cette qualité permet notamment aux auditeurs de se déplacer en gardant une perception cohérente de la perspective entre les sources, ce qui contribue à l'impression d'immersion et de présence de la scène virtuelle ainsi produite. L'IRCAM a acquis une expérience dans ce mode de reproduction grâce à sa participation au projet Européen CARROUSO. Un démonstrateur constitué de 48 canaux a été développé et préfigure des systèmes plus

ambitieux visant l'intégration de ce mode de reproduction dans les installations sonores ou les salles de concert.

Les principaux axes de travail scientifique concernent la caractérisation la plus fine possible du champ effectivement produit par ce type de synthèse. Des extensions théoriques ont été proposées pour permettre la synthèse et le contrôle des propriétés de directivité associées aux sources virtuelles. La mise en œuvre de la WFS pose le problème de la compensation du contexte d'écoute à la fois en termes plus complexes mais également plus prometteurs que pour les systèmes conventionnels. Le nombre de transducteurs et leurs caractéristiques individuelles augmentent inévitablement la complexité du travail d'égalisation du dispositif de diffusion, cependant la possibilité de contrôler le champ acoustique rayonné par le banc de haut-parleurs permet d'envisager des méthodes de compensation basée sur l'annulation des contributions réfléchies par les parois de la salle sous certaines conditions.

Parallèlement à ces études scientifiques, le travail d'implantation des techniques de synthèse et d'égalisation se poursuit de sorte à mettre à jour le dispositif expérimental de l'IRCAM. Par ailleurs des développements de plug-in ont été également réalisés et ont été d'ores et déjà validés dans le cadre de réalisation d'exemples de démonstration ou de premières productions musicales (cf § 1.1.7).

1.1.2.1. Travail théorique

1.1.2.1.1. Extension à la synthèse de sources directives

La Wave Field Synthesis ne considère, dans sa formulation initiale, que le cas d'un champ associé à des sources virtuelles ponctuelles omnidirectionnelles en champ libre. Un travail théorique a permis d'étendre ce cadre à la synthèse de sources à directivité complexe en se basant sur la synthèse d'un sous-ensemble d'harmoniques sphériques. Afin de dériver les fonctions d'alimentation des haut-parleurs dans un formalisme simple cohérent avec celui pour les sources omnidirectionnelles, un certain nombre d'approximations sont nécessaires. Leur validité asymptotique en fréquence est assurée mais il s'agirait d'analyser finement leurs conséquences sur le champ synthétisé en basses fréquences.

1.1.2.1.2. Caractéristiques du champ synthétisé

La formulation de la Wave Field Synthesis repose sur plusieurs approximations qui ont des conséquences importantes sur les caractéristiques du champ rayonné (valable uniquement dans une portion du plan horizontal et jusqu'à une fréquence limite). Une étude théorique systématique a été menée dans le cas de la synthèse de sources virtuelles omnidirectionnelles afin d'évaluer les conséquences sur la qualité de la reproduction de chacune de ces approximations. Des descriptions originales des phénomènes à l'œuvre ont été proposées (modification de la loi d'atténuation, diffraction, aliasing) et ont été mis en forme dans la thèse d'Etienne Corteel [Corteel 04a].

1.1.2.1.3. Compensation de l'effet de salle

Un travail systématique théorique sur l'utilisation de bancs de haut-parleurs encastrés dans les parois de la pièce de restitution a été mené afin d'évaluer les performances d'une stratégie de compensation de l'effet par l'émission d'ondes en opposition de phase avec l'onde réfléchi. Une réduction significative du niveau de la réflexion en basses/moyennes fréquence semble possible dans le plan horizontal. En dehors du plan horizontal les performances sont dégradées et sont limitées en termes fréquentiels. Un problème de cohérence de la loi d'atténuation des ondes réfléchies et émises tend aussi à limiter la validité de l'approche. Une évaluation perceptive approfondie de cette technique reste à effectuer ainsi qu'une estimation de ses conséquences sur les caractéristiques du champ diffus.

1.1.2.1.4. Puissance rayonnée par le dispositif

Le rendu d'une impression spatiale réaliste nécessite d'assurer un effet de salle qui soit cohérent avec le son direct des sources sonores virtuelles. C'est un enjeu majeur pour des situations où l'on voudrait faire figurer des sources virtuelles WFS à côté d'instrumentistes réels sur une scène de concert. Une analyse de la puissance effectivement rayonnée par un dispositif WFS idéal synthétisant une source dipolaire en rotation a montré une différence avec ce que l'on attendrait d'une vraie source dipolaire. Ceci implique une

modification du rapport (son direct)/(son réverbéré) à travers la pièce d'écoute pour les sources directives reproduites en WFS. Cette modification peut être compensée dans certains cas de figure par l'utilisation d'un système de réverbération artificiel. Ce travail est décrit dans [Caulkins 04a].

1.1.2.2. Travail pratique

Le travail pratique mené a consisté à continuer le développement de la méthode d'égalisation multicanal permettant la compensation de la réponse fréquentielle des haut-parleurs utilisés dans une large zone d'écoute. Les améliorations proposées se sont nourries du travail théorique mené au cours de l'année 2004. Elles concernent la méthode d'inversion multi-canal qui effectue une première séparation entre la construction du front d'onde et l'égalisation du champ. Cette méthode a été mise en œuvre sur divers dispositifs WFS utilisés dans l'installation permanente dans le studio 4 de l'IRCAM et dans les installations temporaires (Forum Neues Musiktheater, Studio 2 pour les journées Résonances 2004). Une campagne de mesures a permis de vérifier que l'égalisation reste efficace à différentes distances du banc de haut-parleurs.

Participants : E. Corteel (thèse Paris6) , T. Caulkins (thèse Paris6)

Collaborations extérieures : J.D Polack (LAM)

1.1.3. PROJET OPERA

Ce projet concerne la recherche et le développement de méthodes d'optimisation du rendu spatial d'une scène sonore composée d'un grand nombre de sources sonores. De manière à économiser les ressources de calcul, la démarche générale consiste à segmenter et hiérarchiser cette scène sonore en un nombre réduit de sources sonores ou de groupe de sources prémixées soulageant le traitement de spatialisation. Ces travaux ont été initiés à l'INRIA par Nicolas Tsingos, et utilisent une extension des propriétés de masquage auditif au cas de scènes sonores spatialisées. Les modèles se basent sur des métriques perceptives pour décider de l'écrémage des sources en fonction de leur masquage mutuel ou de leur regroupement sous forme d'imposteurs sonores lorsque les performances de localisation auditive ne permettront pas de distinguer leur localisation spécifique. Dans un premier temps ces modèles ont été testés sur des scènes anéchoïques. La mission de l'Ircam est d'étendre le modèle par la prise en compte d'un flux réverbéré lié aux contributions de l'environnement. Différents niveaux de simplifications sont envisagés, chacun d'eux faisant l'objet d'une implémentation et d'une procédure de validation perceptive expérimentale. A cet effet des tests perceptifs ont été réalisés de manière à dégager des règles de regroupement de sources en condition réverbérée.

1.1.3.1. Performances de localisation en présence de réverbération

Il existe dans la littérature différentes études consacrées à la détection et à la localisation des sources en milieu réverbéré. Cependant les conditions expérimentales sont assez spécifiques et ne s'appliquent pas directement aux situations envisagées dans le projet.

Une première expérience a consisté à évaluer, en condition d'écoute binaurale, les performances de localisation de bursts de bruit provenant de différentes directions et avec différents degrés de rapport son direct sur champ réverbéré. Les résultats du test montrent une tendance nette des réponses des sujets à se regrouper autour du plan médian au fur et à mesure que l'intensité du champ réverbéré augmente. La présence de champ réverbéré tend à diminuer la latéralisation des sources.

Cette expérience sera déclinée dans plusieurs versions en variant la nature des stimuli (voix) ou en variant les conditions de réverbération (source unique ou présence de sources adverses).

La suite du travail sera consacrée à une première implantation d'un modèle évaluant estimant les règles de sélection et de regroupement des sources en présence de la réverbération. Ce modèle fera l'objet de tests de validation.

Participants : P. Guiu, J. Ahrens

Collaborations extérieures : N. Tsingos (INRIA)

1.1.4. PROJET SEMANTIC HIFI

La première année consacrée au projet SemanticHIFI a consisté en une étape de recherche et de spécification générale sur les possibilités d'application des outils et techniques de spatialisation développés à l'IRCAM au

cadre d'une utilisation domestique. Ces recherches concernent autant les modules de traitement du signal qui permettent le rendu sonore spatialisé sur un système de restitution domestique (typiquement une écoute de type 5.1, ou par casque) que les interfaces de contrôle qui proposeront à des utilisateurs non experts en mixage et spatialisation un contrôle de la scène sonore pertinent, intuitif et surtout garantissant l'intégrité du résultat sonore. La particularité de ces recherches repose principalement sur l'utilisation de méta-données, d'informations relatives au contenu sonore soit établies au préalable, soit calculées en temps réel, et permettant de rendre plus efficaces les outils de traitement du signal ainsi que d'en permettre l'utilisation systématique pour un corpus musical donné.

Ces investigations ont été exprimées sous la forme d'un ensemble de « cas d'étude » spécifiant des scénarios typiques d'utilisation domestique des possibilités offertes par la spatialisation. Ces cas d'études ont ensuite été décomposés en modules de base au sein d'un document de spécification préliminaire. Finalement, une sélection de ces situations d'utilisation a fait l'objet d'un démonstrateur particulier.

1.1.4.1. MusicSpace

Le projet MusicSpace est une interface de contrôle de la spatialisation du son tirant partie de la programmation par contraintes pour établir et maintenir des relations entre les positions géométriques de sources sonores dans l'espace. Il permet à un utilisateur non spécialiste du mixage sonore ou de la spatialisation d'agir sur la configuration spatiale des sources dans une scène sonore en étant assuré que le résultat auditif sera toujours satisfaisant. Initialement développé au laboratoire Sony CSL, ce projet a été totalement intégré au projet SemanticHIFI. Les méta-données associées à une scène sonore sont donc un jeu de contraintes établissant les relations entre les sources sonores qui doivent être toujours vérifiées. Ce jeu de contraintes n'est pas nécessairement unique pour un morceau de musique donné : au contraire, des jeux de contraintes différents, lorsque la structure des sources sonores l'autorise, permettent différentes « interprétations » du contenu musical.

1.1.4.2. Cocktail Party Browser

L'idée du Cocktail Party Browser est de proposer la navigation dans une base de données de titres musicaux de manière purement auditive. Le principe repose sur la capacité humaine à séparer des flux auditifs simultanés d'autant plus aisément qu'ils sont spatialement disjoints. Ainsi, 3 morceaux de musique sont présentés en continu à l'auditeur, l'un en « avant plan » au centre de l'image sonore, et les deux autres en « arrière plan », de part et d'autre de la source centrale.

A chaque instant, l'auditeur qui perçoit en « demi teinte » les sources sonores secondaires a la possibilité de basculer son attention vers l'une d'entre elle : la source choisie vient alors au cours d'une transition souple se placer au centre de la scène sonore, et les sources secondaires sont remplacées par deux morceaux « voisins » (au sens d'une métrique pré-déterminée) du morceau central vers lesquels, tour à tour, l'auditeur a la possibilité de porter son attention. Les méta-données mises en œuvre dans ce projet résident essentiellement dans la construction des relations de voisinage entre les différents morceaux qui peuvent être établies soit à partir d'informations captées sur le signal (pour séparer par exemple des morceaux incompatibles), soit à partir d'informations sémantiques, permettant par exemple de présélectionner un sous-ensemble de la base de donnée musicale que l'auditeur souhaite explorer (« tous les morceaux de pop-anglaise des années 60 » par exemple...).

1.1.4.3. Interface de contrôle - LSCANNER

Cette recherche porte sur les interfaces de contrôle de la spatialisation et s'inspire des travaux de recherche généraux concernant les interfaces de type « see through ». En effet, la situation de diffusion de son spatialisé est une situation de réalité augmentée par nature: quelque soit le dispositif de restitution utilisé (casque audio, haut-parleurs, réseaux de haut-parleurs,...) le principe consiste toujours à simuler la présence de sources sonores à des emplacements localisés distincts de l'environnement physique dans lequel elles sont diffusées. Il semble ainsi pertinent, afin d'augmenter la facilité de manipulation des scènes sonores, d'imaginer des interfaces de contrôle mettant en relation les mondes physique et virtuel en les superposant dans une même représentation.

Une première version de l'application, dénommée LSCANNER, met en évidence ce concept au moyen d'un ordinateur portable de type « tablet PC » auquel est associé un système de tracking permettant de localiser l'ordinateur en position et orientation en temps réel : l'ordinateur affiche ainsi une représentation de

l'environnement physique (au moyen d'un modèle 3D prédéfini) sur laquelle sont surimposées des icônes décrivant les sources sonores. L'utilisateur peut ainsi en déplaçant l'ordinateur « scanner » l'environnement physique, et, au moyen d'un bouton et de mouvements qu'il imprime à l'ordinateur portable, attraper, relâcher et déplacer les sources sonores dans l'espace. Des simplifications de ce prototype initial, utilisant éventuellement d'autres méthodes de repérage spatial seront nécessaires pour se rapprocher des contraintes de production d'un appareil domestique et grand public.

Les travaux actuels liés au projet Semantic HiFi se concentrent maintenant vers l'étude et la conception d'une architecture globale, permettant d'assembler les différentes briques du système et de proposer un environnement rendant possible des transitions souples entre les différentes applications d'écoute de la musique.

Participants : O. Delerue

1.1.5. COGNITION AUDITIVE SPATIALE

En 2004 ce volant d'activité a consisté essentiellement en un travail d'analyse des expériences réalisées en 2003, nécessitant le cas échéant l'élaboration d'expériences contrôle, et à des travaux de rédaction [Viaud-Delmon 04a&b, 05a]. L'analyse s'est principalement concentrée sur une expérience mettant en jeu un principe de conflit visuo-auditif de la localisation directionnelle. L'introduction d'un conflit entre plusieurs modalités sensorielles est un moyen intéressant pour comprendre les mécanismes d'intégration multi-sensorielle. En effet, la perception spatiale doit être coordonnée à travers plusieurs modalités sensorielles qui encodent chacune l'information spatiale d'une manière propre. Par ailleurs, le cerveau doit continuellement ré-évaluer les différentes données d'entrée pour optimiser la correspondance entre le monde extérieur et sa représentation mentale.

L'étude consistait à soumettre le sujet à un effet ventriloque à savoir présenter des stimuli sonores synchronisés temporellement avec des événements visuels mais avec un biais systématique de localisation en azimuth de 15° vers la gauche (stimulus visuel à gauche à par rapport à l'événement auditif). L'objectif est d'observer si l'exposition du sujet à cette situation conflictuelle « ventriloque » aboutit à une recalibration de l'espace auditif incluant également des positions spatiales non présentées lors de la phase d'immersion, notamment des directions situées de part et d'autre du secteur angulaire soumis au conflit visuo-auditif, ainsi que les directions provenant de l'hemi-espace arrière.

Les principales observations montrent que la recalibration affecte l'ensemble de la sphère auditive avec cependant des hétérogénéités spatiales significatives. Ainsi, la recalibration exhibe une grande symétrie avant-arrière, illustrant le rôle prédominant de l'ITD dans le processus de localisation, tandis que les hémispaces droit et gauche présentent des différences significatives. L'interprétation de cette dissymétrie demande cependant un approfondissement de l'analyse en observant séparément différents groupes de sujets.

Collaboration externe: L. Sarlat, I. Viaud-Delmon (UMR 7593 – Hôpital de la Salpêtrière)

1.1.6. DEVELOPPEMENT

1.1.6.1. Portage Spat~ pour Max sous Windows

Le spatialisateur IRCAM, librairie de composants logiciels de traitement du signal permettant de réaliser – au sein de l'environnement Max MSP – la spatialisation de sources sonores a été entièrement porté sous Windows, plateforme de prédilection pour l'expérimentation et l'échange entre les partenaires du projet SemanticHIFI. Ce portage concerne une centaine d'objets « externes » recompilés pour le système d'exploitation cible, ainsi qu'un ensemble de patches à adapter. Cette version fait désormais partie de la distribution Spat~ au forum IRCAM.

Participants : O. Delerue

1.1.6.2. Adaptation de MusicSpace au protocole de communication OSC

Dans le cadre de son intégration au projet SemanticHIFI, le programme MusicSpace initialement conçu comme une application autonome possédant son propre moteur de rendu sonore (basé sur l'API DirectX de

Microsoft) a été adapté pour s'interconnecter avec les outils de spatialisation développés à l'IRCAM. L'essentiel de ce travail a consisté en la construction dans MusicSpace d'une nouvelle classe d'objets graphiques (méthodes de représentation, méthodes de sauvegarde,...) représentant chacun un canal de spatialisation et interprétant leur position géométrique en un message de commande de spatialisation envoyé au Spat~ sous forme d'un message réseau encodé au format OSC. Cette nouvelle version de MusicSpace peut communiquer avec le Spat~ aussi bien dans sa version pour Mac OSX que dans la nouvelle version pour Windows. Différentes démonstrations ont été réalisées dans le cadre du projet Semantic HiFi, ainsi que des premières expérimentations permettant de piloter le dispositif de diffusion WFS. Des modifications complémentaires, liées à la gestion de priorité des tâches, ont été rendues nécessaires pour autoriser le fonctionnement simultané du moteur de spatialisation et de MusicSpace sur une même machine.

Participants : Olivier Delerue

1.1.6.3. Plug-ins d'édition et de contrôle de Scènes Sonores en WFS

Des plug-ins d'édition et de contrôle de scènes sonores en WFS ont été développés en 2004. Programmés dans l'environnement MaxMSP (Pluggo) ils permettent de contrôler entièrement le déroulement d'une scène sonore en WFS depuis un séquenceur classique comme ProTools, Logic Audio ou Digital Performer. Pour ce faire, les paramètres de position de sources sonores et d'effet de salle sont transformés en paramètres d'automation au sein du séquenceur. L'utilisateur peut ainsi programmer l'évolution temporelle de ces paramètres en utilisant la palette d'outils habituelle du séquenceur. Le développement de ces plug-ins avait été initié par Manuel Poletti lors du projet CARROUSO. Ils ont ensuite été repris par Etienne Corteel et Terence Caulkins en préparation du workshop WFS qui a eu lieu au Forum Neues MusikTheater. Une première utilisation des plug-ins a testée sous forme d'extraits de la pièce « Vertigo Apocalypsis » de Philippe Schoeller transposés pour la WFS. Ces plug-ins ont ensuite été finalisés pendant l'été 2004 en préparation des Résonances de l'IRCAM. Ce travail a fait l'objet d'une première exploitation en production par Gilles Grand à l'occasion de la composition de « Retours » diffusée lors de Résonances en format WFS.

Participants : T. Caulkins, E. Corteel

1.1.7. PRODUCTION MUSICALE

L'année 2004 a vu la réalisation ou la finalisation de productions musicales dédiées, d'une part, à la Wave Field Synthesis et, d'autre part, aux installations sonores basées sur la navigation dans une scène sonore augmentée et spatialisée par technique binaurale.

1.1.7.1. Production pour la Wave Field Synthesis

Différentes productions ont permis d'éprouver les développements technologiques et outils de production développés dans l'équipe en collaboration avec la société SonicEmotion. Ce premier répertoire, auquel il convient d'ajouter des travaux effectués fin 2003 (*Hinano* de **S. Roux** et *Greek Tragedy* de **R. Galichet**), couvre un ensemble de situations de production, tant au niveau des matériaux musicaux (musique concrète, musique instrumentale, musique mixte, dramatique) que des situations de mixage (post-production d'enregistrement de concert, utilisation du paradigme des haut-parleurs virtuels, montage son, ...).

Nicolas Vérin, *11, avenue du Midi*, pièce radiophonique réalisée dans les studios de la Muse en Circuit et de Ligys ; Mixage pour WFS réalisé dans les studios de l'Ircam par Julien Perraudeau (association Hérissons).

Yan Maresz, *Sul Segno*, oeuvre pour guitare, cymbalum, harpe et contrebasse, créée en juin 2004 par Frédérique Cambreling (harpe), Frédéric Stochl (basse), Pablo Marquez (guitare) et Michel Cerutti (cymbalum), avec la collaboration de Manuel Poletti, assistant musical et Sébastien Naves, ingénieur du son.

Gilles Grand, *Retour(s)*, suite radiophonique multipiste, succession d'extraits de l'enregistrement du spectacle *Retour Définitif et Durable de l'Être Aimé* d'Olivier Cadiot, interprété par Valérie Dashwood, Philippe Duquesne, Laurent Poitrenaux, et avec la voix de l'auteur.

Philippe Schoeller : transcription pour WFS de matériaux extraits de *Vertigo Apocalypsis*.

Participants : T. Caulkins, E. Corteel
Compositeurs : G. Grand, N. Vérin, P. Schoeller, Y. Maresz
Collaboration interne : S. Naves
Collaboration externes : J. Perraudeau

1.1.7.2. Production binaurales pour les installations sonores interactives

A l'occasion de Résonances 2004, plusieurs maquettes d'installation sonore interactive ont pu être réalisées et expérimentées par le public. Ces réalisations, dont certaines avaient été initiées fin 2003, sont directement issues des travaux effectués dans le cadre du projet LISTEN réalisé entre 2001 et 2003. Elles mettent en œuvre simultanément la technologie de synthèse binaurale et différents principes d'interactivité basés sur la navigation de l'auditeur dans la scène virtuelle. Les coordonnées spatiales de l'auditeur, captées par un système de suivi de position (caméras infra-rouge et marqueur passif), sont exploitées par l'application ListenSpace, développée au cours du projet LISTEN. Cette application permet à l'auteur de décrire, d'une part, l'organisation spatiale des sources sonores composant la scène virtuelle ou augmentée dans laquelle le sujet déambule et, d'autre part, les lieux et types d'interaction ou de comportement des éléments de la scène. Ces interactions sont généralement basées sur la position et l'orientation relatives des objets sonores (sources ou auditeur(s)) et comprennent également la détection d'entrée ou de sortie de zones virtuelles superposées à la géométrie du lieu de déambulation. L'auteur peut également utiliser des comportements élémentaires tels que l'asservissement de la position d'un objet relativement à l'auditeur (suivi automatique). En phase d'exploitation, les informations sont transmises sous protocole OSC au Spatialisateur pour mettre à jour en temps réel la scène sonore diffusée à l'auditeur par casque sans fil. Plus généralement, ces informations peuvent être exploitées et complétées par des processus comportementaux décrits dans MAX/MSP préalablement à l'étape de spatialisation.

Cécile Le Prado : transcription d'extraits de *Secrète Lisboa*, fiction documentaire résultat d'une commande de l'Atelier de Création Radiophonique de France Culture. Le projet consistait à décliner les matériaux de cette œuvre pour un dispositif interactif. L'une des originalités du travail réside dans l'exploitation conjointe de matériaux sonores monophoniques (voix) et d'ambiances de ville captées en format B (ambisonique) et véhiculant par conséquent leur propre information spatiale. La localisation des sources monophoniques (voix) est synthétisée classiquement en mode binaural. Les propriétés du format ambisonique sont exploitées de manière à appliquer des transformations spatiales congruentes avec les mouvements de l'auditeur. Celui-ci peut donc avoir l'illusion de parcourir le paysage sonore capté préalablement.

Olivier Delerue : *La Volière*. Cette installation sonore interactive est prétexte à présenter une panoplie de situations d'écoute propres à la réalité augmentée sonore et à la restitution binaurale : explorer une scène, écouter des sources statiques, suivre le trajet de sources en mouvements, ou encore percevoir les effets de perspective et de relief produits par les rapports de distance entre chacune des sources et l'auditeur. C'est aussi l'occasion de présenter certaines des possibilités d'interactions accessibles au moyen des outils développés dans l'équipe, ainsi qu'une manière concrète de les mettre en œuvre avec les logiciels ListenSpace, Max MSP et le Spatialisateur. Concrètement, la scène sonore est composée de trois oiseaux, perchés en différents endroits de la salle, immobiles et sifflant chacun un chant qui lui est propre. L'auditeur peut à loisir explorer cet environnement, adopter différents points d'écoute, faire face aux oiseaux, leur tourner le dos, s'en éloigner ou s'en rapprocher, ... Mais dès lors qu'il les approche de trop près, les oiseaux s'envolent et vont se percher ailleurs, pour garder la distance. L'auditeur manipule un bâton sur lequel un oiseau vient parfois se poser, proposant alors une écoute binaurale de proximité au cours de laquelle l'utilisateur contrôle directement la position de la source.

Julien Perraudeau : jeu interactif *Le manoir*. Cette installation sonore basée sur le principe d'un jeu interactif propose un univers composé de différentes sources sonores. Le joueur est tour à tour invité à localiser ou identifier des sources cachées, à mémoriser l'organisation spatiale de la scène ou à reproduire de courtes séquences musicales. La capture des déplacements du joueur est complétée par celle des mouvements de la main pour varier et enrichir l'interaction, notamment pour la reproduction de séquences sonores.

Participants : Olivier Delerue
Compositeurs : C. Le Prado

Collaboration interne : R. Kronenberg

Collaboration externes : J. Perraudou / Association Hérissons

1.1.7.3. Jeu instrumental et directivité

Dans le cadre de l'atelier « Improvisation et Ordinateur » une expérience de contrôle instrumental de la directivité a été réalisée avec le compositeur Georges Bloch. Le contexte initial était la mise en jeu d'une improvisation automatique en réponse à un saxophoniste. Cette improvisation était calculée en temps réel par le logiciel Oracle, développé dans l'équipe Représentations Musicales, et adapté par Georges Bloch pour travailler directement à partir de l'enregistrement live du saxophoniste, segmenté et indexé note à note en temps réel. Afin de contribuer à la *présence* de l'improvisateur virtuel, les différents flux sonores calculés par l'ordinateur étaient diffusés par une Timée jouant « en duo » avec l'instrumentiste. Pour augmenter la lisibilité des différents flux diffusés par la Timée ceux-ci étaient associés à des directivités, spécifiques à chaque flux, et animées de vibratos d'orientation et de directivité asservis respectivement à la hauteur et à l'intensité de la séquence musicale produite.

Participants : O. Warusfel

Collaboration interne : G. Bloch, M. Chemillier, N. Misdariis

Instrumentiste : Philippe Leclerc

1.1.8. PUBLICATIONS ET COMMUNICATIONS

Articles parus dans des revues à comité de lecture

[Viaud-Delmon 05a] Viaud-Delmon I., Warusfel O., Seguelas A., Rio E., Jouvent R., « High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality », *European Psychiatry*, in press.

Actes de congrès avec comité de lecture

[Blum 04a] Blum A., Katz B., Warusfel O. « Eliciting adaptation to non-individual HRTF spectral cues with multi-modal training presence », Proc. 7^{ème} Congrès Français d'Acoustique / 30. Deutsche Jahrestagung für Akustik, Strasbourg (2004)

[Busson 04a] Busson S., Nicol R., Warusfel O. « Influence of the ears canals location on spherical head model for the individualized interaural time difference presence », Proc. 7^{ème} Congrès Français d'Acoustique / 30. Deutsche Jahrestagung für Akustik, Strasbourg (2004).

[Corteel 04c] Corteel E., Caulkins T., Warusfel O., Pellegrini R., Rosenthal M., Kuhn C., « Sound Scene Creation using Wave Field Synthesis », Actes Journées du Design Sonore, Paris (2004).

[Caulkins 04b] Caulkins T., Corteel C., Warusfel O., « Analysis of certain challenges for the use of Wave Field Synthesis in Concert Based Applications », Proc. Digital Audio Effects DAFx'04, Napoli (2004), vol 7, p.250-255.

[Viaud-Delmon 04b] Viaud-Delmon I., Seguelas A., Rio E., Jouvent R., Warusfel O., « 3-D Sound and Virtual Reality: Applications in Clinical Psychopathology », Cybertherapy 2004, San Diego, January 2004.

[Viaud-Delmon 04a] Viaud-Delmon I., Sarlat L., Warusfel O., « Localization of Auditory Sources in Virtual Reality », Proc. 7^{ème} Congrès Français d'Acoustique / 30. Deutsche Jahrestagung für Akustik, Strasbourg 2004.

[Warusfel 04a] Warusfel O., Corteel E., Misdariis N., Caulkins T., « Reproduction of sound source directivity for future audio applications », ICA-International Congress on Acoustics, Kyoto 2004.

[Warusfel 04e] Warusfel O., Misdariis N., Caulkins T., Corteel E., « Radiation control applied to sound synthesis : an attempt for "Spatial Additive Synthesis" », Proc. 75th ASA Meeting, New-York 2004.

Actes de congrès sans comité de lecture

[Warusfel 04b] Warusfel O., Corteel E., Caulkins T., « Perceptual Spatial Exploration of WFS Soundscapes presence », Proc. 7^{ème} Congrès Français d'Acoustique / 30. Deutsche Jahrestagung für Akustik, Strasbourg 2004

[Warusfel 04c] Warusfel O., Viaud-Delmon I., Delerue O., « Binaural rendering assessment in the context of augmented reality », Proc. 7^{ème} Congrès Français d'Acoustique / 30. Deutsche Jahrestagung für Akustik, Strasbourg 2004

[Warusfel 04d] Warusfel O., Misdariis N. « Sound Source Radiation Synthesis : from Stage Performance to Domestic Rendering », Proc. 116th AES Convention, Berlin 2004

Travaux universitaires

[Corteel04a] Corteel E., « Caractérisation et extensions de la Wave Field Synthesis en conditions réelles », Thèse de doctorat, Université Paris 6, soutenue le 9 décembre 2004.

Rapports de recherche

Delerue O. Projet IST-507913 Semantic HiFi Deliverable D1.2.1 Functional specification and refined workplan

Delerue O. Projet IST-507913 Semantic HiFi Deliverable D4.1 First rendering functional demonstrators

Conférences invitées

Warusfel O., Viaud-Delmon I., « La Spazialisazione dei suoni - Prospettive per la musica e la realtà virtuale », Università di Salerno (2004).

Diffusion de connaissances

[Corteel 04b] Corteel E., « Création et manipulation de scènes sonores pour la wave field synthesis », Cahiers Louis-Lumière n°2 (2004).

Colloques et séminaires

Warusfel O. « Acoustique virtuelle et Spatialisation », La Semaine du Son, Paris (2004)

Warusfel O. « LISTEN -Augmenting everyday environments through interactive soundscapes » Workshop VR for Public Consumption - IEEE Virtual Reality – Chicago 2004.

Warusfel O., Collaboration avec G. Bloch et N. Misdariis pour l'utilisation de la Timée dans le cadre du colloque « Improvisation et Ordinateur ».

Emissions radiophoniques et télévisées, entretiens journalistiques, animations

Résonances 2004 – A l'occasion des rencontres Résonances, l'équipe Acoustique des Salles a organisé une série de démonstrations autour des techniques et dispositifs d'écoute immersive pour le concert, les installations sonores et la réalité virtuelle. Ces techniques de création de scènes sonores ont été illustrées par un programme où alternaient démonstrations scientifiques, technologiques et musicales sous forme de remix, d'esquisses ou de propositions interactives (G. Grand, R. Galichet, Y. Maresz, P. Schoeller, N. Verin, C. Le Prado, O. Delerue, J.Perraudeau). 13-22 Octobre 2004

Workshop WFS – En collaboration avec le Forum Neues MusikTheater de Stuttgart et la société SonicEmotion, l'équipe Acoustique des Salles a organisé un atelier de démonstration et de présentation technologique consacré à la Wave Field Synthesis. 30 avril – 2 Mai 2004.