



IRCAM – UMR 9912
EQUIPE ACOUSTIQUE DES SALLES
RAPPORT D'ACTIVITES - 2003



•Responsable :

Olivier Warusfel

•Ingénieurs/Chercheur

Olivier Delerue (Projet Listen)

Cedric Peigné (Projet Listen)

Emmanuel Rio (Projet Listen)

Sébastien Roux (Projets Listen & Carrouso)

Riitta Väänänen (projet Carrouso)

Guillaume Vandernoot (projet Listen)

•Thésards :

Alexis Baskind, (Paris 6)

Etienne Corteel, (Paris 6)

Terence Caulkins, (Paris 6)

Sylvain Busson, (Marseille)

•Stages :

Alan Blum, Clemens Kuhn,

Mathieu Nogues, Charles Verron

• Collaborations principales:

Alain de Cheveigné (IRCAM équipe PCM)

Brian Katz (LIMSI)

Nicolas Misdariis (IRCAM, équipe Design Sonore)

Isabelle Viaud-Delmon (UMR 7593)

Ludivine Sarlat, Angeline Seguelas (UMR 7593)

•Autres collaborations:

FT R&D, Université de Delft (NL), KunstMuseum Bonn, Fraunhofer IMK (Bonn), Fraunhofer AEMT (Ilmenau), Ecole Polytechnique Fédérale de Lausanne (EPFL), AKG, Université de Vienne (AUT), Université d'Erlangen (DE), Studer.

Responsable Olivier Warusfel

L'année 2003 représente pour l'équipe l'aboutissement d'un cycle de recherches et de développements suscités par les projets européens Carrouso et Listen. Des avancées notables ont été recueillies dans les domaines de la reproduction de champs sonores holophoniques et de la réalité augmentée interactive. A l'issue de ces 3 années, ces projets ont atteint une première phase de maturité permettant leur exploitation dans le cadre du concert ou des installations sonores interactives.

Une part importante du travail a été consacrée au développement de la technique de rendu holophonique, en mettant l'accent sur les procédés d'égalisation multicanal qui permettent simultanément le contrôle du champ direct et celui de l'effet de salle associé. Par ailleurs, une méthode originale de synthèse de la directivité des sources a été développée et validée objectivement dans le cadre de la reproduction holophonique. Le contrôle de l'effet de salle est envisagé parallèlement sous l'angle de la compensation du contexte d'écoute lorsque la restitution exige sa neutralité, ou au contraire sur la mise en cohérence de l'effet de salle

synthétisé et de la salle d'écoute afin de contribuer à la sensation de la présence des sources virtuelles.

Les travaux sur la reproduction binaurale ont été consacrés principalement aux techniques d'adaptation individuelle, envisagées successivement à travers l'optimisation et la validation de procédés d'encodage multicanal, l'adaptation du paramètre de retard inter-aural et sa modélisation à partir de paramètres morphologiques ou à partir de simulations numériques basées sur une modélisation physique de la tête. Enfin, une étude originale a été consacrée à la plasticité auditive spatiale en observant comment l'adaptation du système auditif à des HRTFs non-individuelles peut être stimulée grâce à une phase d'apprentissage mettant en jeu un principe perception-action.

Sur le plan de la recherche amont, les études ont concerné, d'une part, l'analyse automatique de scène et, d'autre part, la cognition spatiale auditive. L'année 2003 a vu l'aboutissement du travail de thèse d'Alexis Baskind avec des résultats très prometteurs sur l'extraction automatique de paramètres de description de l'organisation spatiale de la scène. Les études sur la cognition spatiale auditive élargissent le champ de recherche consacré à la perception spatiale en y intégrant d'autres modalités sensorielles telles que les dimensions visuelles ou idiothétiques. La motivation de ces recherches est liée à l'expansion des installations sonores interactives et au rapprochement avec les domaines de la réalité virtuelle ou augmentée pour lesquels différentes voies sensorielles sont sollicitées conjointement. Ce nouveau volet de recherche a fait l'objet d'une série d'expériences exploratoires permettant de baliser le terrain de recherche pour les prochaines années.

1.1.1. Outils auteurs pour l'édition et le contrôle de scènes sonores

L'année 2003 a été consacrée à la poursuite des efforts visant la création d'un outil de construction et de visualisation de scènes sonores virtuelles dans le cadre d'applications de la réalité audio augmentée. Cet objectif original s'est progressivement élargi à d'autres types d'applications s'articulant essentiellement autour de besoins en termes de contrôle de la spatialisation. L'outil a ainsi permis d'intégrer des développements effectués dans le cadre du projet CARROUSO pour la création et la transmission de scènes au format MPEG-4. Le travail a pu être évalué au cours de mises en situation dans le domaine de la création ou de l'expérimentation scientifique sur la cognition spatiale (cf. § 1.2.3 et § 1.2.6.1).

1.1.1.1 ListenSpace

Réalités audio augmentées

La construction de scènes sonores interactives, en particulier dans le cadre de réalités audio augmentées, nécessite un outil auteur permettant la création, la représentation et la manipulation des éléments sonores virtuels qui les composent. Ainsi, ListenSpace représente sous forme de projection bidimensionnelle les éléments physiques et virtuels d'un univers tridimensionnel simplifié. Les éléments de la scène sont représentés de façon schématique et proposent une interface permettant de contrôler l'ensemble des paramètres qui les caractérisent (paramètres perceptifs d'une source sonore, par exemple).

Les pré-requis pour un tel environnement impliquent également une flexibilité suffisante pour faire face aux différentes situations d'interaction désirées. Les développements relatifs à

ListenSpace ont été guidés par des objectifs liés à la simplicité et à la connectivité : ListenSpace est avant tout un moyen de représentation géométrique de la scène sonore, capable de transmettre aux applications connectées différents paramètres géométriques de la scène nécessaires à l'interactivité (position absolue des éléments, positions relatives entre des éléments, interactions entre éléments telles que les entrées et sorties de zone, ...).

Nous avons mis en évidence, particulièrement dans le cadre de réalités audio augmentées, la nécessité d'agir sur la scène sonore au moment même de son exploration. ListenSpace a été ainsi adapté pour les ordinateurs ultra-portables de type « tablet PC » permettant par une communication sans fil, de contrôler la scène représentée dans un repère absolu ou centré sur l'utilisateur et rafraîchie en temps réel en fonction de ses déplacements.

Différents développements complémentaires ont été réalisés tels que l'application «StarApp» qui implémente une redirection « en étoile » des messages réseau transmis par ou vers ListenSpace, ainsi que le module « Reactor » qui constitue un cadre de programmation permettant le développement rapide de comportements pour les éléments de la scène. Quelques exemples de «comportements» ont été développés : l'un d'entre eux consiste à forcer une source sonore à suivre un auditeur avec un déplacement dont il est possible de contrôler la vitesse et l'inertie. Deux autres applications ont été conçues par ailleurs pour enregistrer et rejouer les déplacements de l'auditeur.

Le contrôle de la spatialisation et la transmission de scènes sonores

Très proche des objectifs initiaux, le contrôle de la spatialisation est l'une des applications envisagées pour ListenSpace. ListenSpace apporte une solution indépendante des moteurs de rendu du signal sonore : son utilisation permet en outre la répartition des calculs en parallèle sur plusieurs machines effectuant le rendu sonore, ainsi que la communication avec des moteurs de rendu de natures différentes. L'aspect distribué sous-jacent propose un avancement significatif puisqu'il permet de détacher l'interface de contrôle du choix du moteur de rendu sonore et de piloter la spatialisation directement depuis la zone d'écoute. ListenSpace est ainsi régulièrement utilisé pour le contrôle de scènes sonores dans le cadre de la diffusion en mode Wave Field Synthesis basé sur l'utilisation d'un ou plusieurs réseaux de haut-parleurs et d'une puissance de calcul répartie.

Les mises en situation

L'évaluation de ListenSpace s'effectue au sein d'un ensemble de mises en situation : chacune permettant de mettre en évidence les réels besoins en termes de fonctionnalités d'édition et d'interactivité.

- Expériences sur la cognition spatiale auditive menées en collaboration avec le CNRS-UMR7593 – Hôpital de la Salpêtrière (cf. § 1.2.3). ListenSpace a été largement mis à contribution dans ces expériences cognitives dont la mise en œuvre rejoignait de près une situation de réalité audio augmentée. A cette occasion, ont été conçues des applications complémentaires permettant l'enregistrement et la reproduction des mouvements de l'auditeur.
- Réalisation d'un jeu sonore interactif par Julien Perraudou / Association Hérisson.

- Esquisse autour de « Secrète Lisboa » de Cécile Le Prado (cf. § 1.2.6.1) pour laquelle l'objectif est d'adapter le matériau musical originellement construit pour une diffusion en 5.1 en visant une écoute interactive basée sur la navigation de l'auditeur dans une partition spatiale.
- Expérimentation avec la danse (Espace de projection, août 2003) : dans cette expérience, le processus de spatialisation est renversé et l'espace est utilisé comme paramètre d'entrée du système. Un dispositif de tracking délégué à un(e) danseur(se) contrôle la lecture, la synthèse ou la transformation de flux sonores. En cette occasion, ont été conçus les modules de type «reactor» permettant, par exemple, d'asservir la position d'objets sonores par rapport au danseur, par rapport à un point d'écoute prédéfini, ou selon un angle dépendant des mouvements. L'objectif de cette expérience était d'observer dans quelle mesure l'utilisation de la spatialisation du son combinée à un système de suivi de position permettait de jouer sur la cohésion entre le contenu sonore et la chorégraphie.

Participants : O. Delerue, E. Rio,

Collaborations extérieures : C. le Prado, J. Perraudeau, I. Viaud-Delmon, J. Mason (chorégraphe).

1.1.1.2 Outil auteur pour le format MPEG-4

Les développements menés dans le cadre du projet CARROUSO et permettant la définition et la transmission de scènes sonores au format MPEG-4 ont été finalisés et définitivement intégrés à ListenSpace : la scène sonore construite dans ListenSpace est convertie dans un format textuel et transmise à un encodeur MPEG-4 qui se charge du multiplexage avec le contenu sonore. A ce stade, ListenSpace est donc utilisé comme outil auteur dédié à la description de la scène sonore à transmettre. Au delà de la description de la position des différentes sources sonores et de l'auditeur, l'outil permet d'importer des descripteurs de l'effet de salle issus, par exemple, d'une analyse de réponse impulsionnelle mesurée in-situ. Des fonctions d'édition et d'automatisation de trajectoires ont été également implantées.

Le résultat est transmis soit sous forme de flux, soit sous forme de fichier à un décodeur MPEG-4. L'une des originalités de la démarche est de transmettre, simultanément aux données de description de la scène, une représentation graphique interactive de ses composantes grâce aux primitives graphiques du format MPEG-4. Ainsi, au moment de la réception et du décodage, la représentation graphique de la scène sonore est partiellement reconstituée, semblable à celle définie dans ListenSpace, et fournit automatiquement à l'auditeur une interface graphique lui donnant accès aux paramètres de contrôle. Cependant, de manière à préserver l'intégrité de la scène transmise, l'auteur a le choix de définir les composantes de l'interface qui seront accessibles à l'interaction de l'utilisateur final. L'auteur décide de la représentation graphique de la scène et l'éditeur de source lui permet d'indiquer les options de visualisation et d'interaction de chaque paramètre marqué comme « editable », « visible » ou « caché ». Cette propriété garantit le niveau de propriété du résultat sonore.

Participants : R. Väinänen, O. Delerue

Collaborations extérieures : G. Zoia (EPFL)

1.1.2. Description automatique de scène

Les futures générations de contenu audio feront appel à des techniques d'incrustation et de mixage automatique permettant par exemple le mixage entre des formats audio divers ou l'inclusion d'un objet sonore dans une scène pré-enregistrée ou pré-mixée. Pour faciliter le travail de l'ingénieur du son ou du musicien il s'avèrera important de guider sinon automatiser le travail de mixage pour assurer la transparence et la cohérence du résultat final. Ces techniques s'appuieront notamment sur des modèles de description automatique de scènes sonores permettant de dégager une description symbolique de la scène à partir des signaux audio qui la composent. La disponibilité de cette description ouvre la voie à des transformations ultérieures ou à l'inclusion d'un nouvel élément dont la spatialisation sera synthétisée en fonction de la description de scène obtenue (placement automatique des sources dans l'espace pour minimiser les masquages mutuels et respect de l'effet de salle). Le travail de thèse présenté ci-dessous constitue une première étape dans une voie que nous désirons poursuivre dans les prochaines années.

1.1.2.1 Extraction automatique de descripteurs spatiaux

L'objet de cette recherche est de proposer des méthodes automatiques de description objective des aspects spatiaux d'une scène sonore enregistrée, sans connaissance préalable ni des sources qui la composent, ni du message sonore diffusé, ni des caractéristiques du lieu d'enregistrement. L'étude se concentre plus spécifiquement sur l'estimation de la direction de la source (supposée unique et stable dans un premier temps) et sur une caractérisation de l'enveloppe de réverbération. Un cadre théorique, inspiré des modèles d'audition spatiale, permet de développer un ensemble homogène de méthodes de détection et d'estimation, basées sur des statistiques non stationnaires d'ordre 2 relatives aux relations entre les voies de l'enregistrement. Les informations obtenues dans chaque bande de fréquences sont ensuite regroupées et interprétées au moyen de descripteurs de plus haut niveau. La pertinence de cette approche est étudiée sur des enregistrements binauraux synthétiques et réels.

L'essentiel du travail mené en 2003 a consisté en une synthèse des différentes études menées jusqu'ici, dans une optique de rédaction du mémoire [Baskind03b]. Cela a été notamment l'occasion de développer et de consolider le modèle de détection et de localisation dit "par égalisation et annulation", inspiré par des modèles psychoacoustiques d'audition spatiale, et reposant en pratique sur des algorithmes inspirés des techniques monophoniques de détection de sons harmoniques développées par Alain de Cheveigné de l'équipe Perception et Cognition Musicales à travers le programme 'Yin' [DeCheveigne02k]. La récente implémentation en temps réel de Yin par Norbert Schnell de l'équipe "Applications temps réel" permet de souligner la pertinence de cette approche, et d'envisager le développement d'applications basées sur les méthodes de détection et de description proposées dans ce travail. Cette thèse a été présentée et soutenue publiquement en décembre 2003.

Toujours dans le cadre de cette collaboration ont été présentés, à la 24^e conférence spécialisée de l'AES, des travaux annexes basés sur l'estimation de fréquence fondamentale en milieu réverbérant [Baskind03a].

Participants : A. Baskind

Collaborations internes : A. de Cheveigné (équipe PCM)

1.1.3. Cognition spatiale auditive

Dans le domaine de la reproduction sonore ou de la communication les technologies audio futures tenteront de privilégier les sensations *d'immersion* et de *présence*. Ces notions sont intimement liées aux dimensions spatiales d'une scène multimedia et en particulier sonore. Bien que la modalité auditive présente des performances de localisation moins précises que la vision, l'audition accède en permanence à l'ensemble de la sphère spatiale perceptive et contribue fortement à la sensation d'immersion. La *présence* est notamment renforcée dans les situations impliquant une action de la part de l'auditeur, soit à travers sa navigation dans la scène, soit à travers une interaction gestuelle avec les objets qui la composent. Dans ces conditions, rendues possibles par les technologies holophoniques ou binaurales (à condition, pour ces dernières, d'être asservies à un système de suivi de position) la congruence et le rafraîchissement en temps réel des indices auditifs spatiaux en fonction des mouvements ou actions de l'auditeur ont un impact majeur sur la sensation de présence. Ce nouveau contexte a motivé le lancement d'une série d'expériences consacrées à la cognition spatiale auditive, notamment à travers l'étude de processus d'intégration multi-sensorielle, en privilégiant les modalités auditives et idiothétiques (indices induits par le mouvement du sujet et incluant l'équilibre et la proprioception). La méthodologie expérimentale fait appel à des expériences comportementales liées à l'observation des performances de localisation ou de navigation des sujets soumis à différents contextes exploratoires. Ce domaine d'étude est mené conjointement avec le CNRS UMR 7593, Hôpital de la Salpêtrière, qui nous apporte ses compétences dans le domaine expérimental et méthodologique sur l'étude de la perception multisensorielle et sur la cognition spatiale.

1.1.3.1 Environnement expérimental

Les différentes expériences décrites dans les paragraphes suivants ont nécessité la mise au point préalable d'un environnement expérimental de réalité virtuelle permettant de gérer les modalités visuelles, auditives en interaction avec la capture de mouvement ou de contrôles gestuels du sujet.

Afin de minimiser les développements nécessaires et rester compatible avec le Spatialisateur, nous avons opté pour une architecture ouverte et modulaire basée sur la mise en réseau de différents composants disponibles en interne ou accessibles aisément sur le marché. Les principaux éléments sont le logiciel Virtools, permettant l'importation de scènes virtuelles, et l'édition de scripts comportementaux à l'aide d'un langage graphique. Ces scripts ont ici le rôle de description et d'exploitation du protocole de tests. Pour la composante audio3D, un module de communication UDP utilisant le protocole OSC a été développé afin d'établir la communication avec notre outil d'édition de scène sonore ListenSpace et/ou notre moteur de rendu 3D, le Spat~. Une version ultérieure pourra utiliser la version C du Spat~ développée par FranceTelecom et supportée par Virtools.

A l'occasion de cette collaboration, différents modules ont été développés dans ListenSpace et dans le Spat~. Dans ListenSpace, il s'agit d'un ensemble de plugins périphériques permettant par exemple l'enregistrement et la lecture des données de navigation du sujet.

Des efforts spécifiques ont été entrepris afin de rendre plausible la scène sonore présentée aux oreilles de l'auditeur, ainsi que sa cohérence avec la scène visuelle (navigation dans un environnement urbain). Parmi les différentes solutions envisagées, une composition en deux niveaux de restitution a été retenue. Le premier niveau est constitué de scènes sonores au

format Ambisonics, actives sur tout l'espace de navigation et entre lesquelles l'auditeur peut se déplacer de manière fluide. Les scènes sonores retenues sont issues d'enregistrements effectués en situation urbaine et ont été choisies pour leur caractère immersif et réaliste. Le deuxième niveau, qui se superpose au premier, est constitué de sources sonores ponctuelles spatialisées en mode binaural et associées à des éléments identifiables de l'environnement visuel.

Participants : O. Delerue, E. Rio,

Collaborations extérieures : I. Viaud-Delmon, Y. Ouallouche (UMR7593)

1.1.3.2 Conflit visuo-auditif de localisation directionnelle.

La perception spatiale doit être coordonnée à travers plusieurs modalités sensorielles. Cette tâche représente un problème non trivial dans la mesure où chaque modalité encode 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. Pour étudier comment nos différents sens contribuent à l'émergence d'un percept unique, la réalité virtuelle offre un cadre approprié permettant, par exemple, de créer des situations de conflit sensoriel. Dans le cas présent, l'idée était de présenter des stimuli sonores synchronisés temporellement avec des événements visuels mais en introduisant 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. Le protocole utilisé est basé classiquement sur une étude en trois phases. Dans une phase initiale, on mesure les performances de localisation auditive du sujet en situation auditive pure (sujet masqué). Dans un second temps, on expose le sujet au conflit sensoriel pendant une durée contrôlée à l'issue de laquelle on reitère la phase de mesure de localisation en modalité auditive pure. L'influence du conflit est évaluée à travers l'observation des éventuelles différences de performances intervenues entre les deux situations de mesure de la localisation auditive. Différentes hypothèses peuvent être émises sur l'influence de la phase de re-calibration. Si dans l'espace frontal on peut s'attendre à une rotation des réponses de localisation, la question se pose pour les stimuli auditifs arrière qui n'ont pu par évidence être directement soumis à une situation de conflit. On peut émettre trois hypothèses concurrentes: (1) la localisation subjective ne présente pas d'altération ; (2) la localisation auditive fait apparaître une rotation des événements perçus par rapport aux stimuli ; (3) la localisation auditive fait apparaître une translation (latéralisation) des événements. A l'issue de l'expérience on a pu montrer que la phase de conflit entraînait bien une recalibration auditive de l'ensemble de la sphère auditive, affectant principalement le demi-espace droit, mais incluant par conséquent les positions avant et arrière. Cependant, les difficultés expérimentales liées à l'usage de fonctions de transfert de tête (HRTFs) non individuelles rendent délicate l'interprétation sur les héli-champs avant et arrière, dans la mesure où de nombreuses inversions avant-arrière subsistaient. Cette étude devrait être poursuivie en améliorant différents aspects du protocole, notamment la phase préalable de sélection des HRTFs pour les sujets, les modalités de report de la localisation et de la phase de conflit (modalité visuelle virtuelle ou réelle incluant ou non la représentation de la main du sujet) [Viaud-Delmon04b].

Participants : G. Vandernoot,

Collaborations extérieures : L. Sarlat, I. Viaud-Delmon (UMR7593)

1.1.3.3 Comparaison des performances de navigation en réalité virtuelle visuelle ou audio-visuelle.

Le but de l'étude est de fournir des données sur l'importance de la modalité auditive dans les environnements de réalité virtuelle en particulier au regard des tâches de navigation. L'expérience consiste à mesurer les performances de navigation spatiale en comparant différents environnements virtuels mettant en œuvre une synthèse purement visuelle (Vis) ou audio-visuelle (Avis). Un aspect conjoint, lié au domaine de recherche de nos partenaires du laboratoire de la Salpêtrière, est de mesurer l'impact pour les applications de réalité virtuelle à usage clinique. Différents groupes sujets sont ainsi conviés à l'expérience afin de comparer l'importance de la composante auditive pour les sujets normaux ou agoraphobiques.

Les sujets étaient équipés d'un casque de vision 3D et immergés dans un environnement urbain virtuel. Les sujets pouvaient avancer à l'aide d'un bouton de souris, tandis que les changements de direction étaient induits par la rotation du corps captée par un dispositif de suivi électromagnétique. La tâche consistait à repérer un ensemble de balises et à se familiariser avec la topographie du lieu. Dans la condition audiovisuelle les sujets étaient équipés d'un casque d'écoute délivrant un paysage sonore rafraîchi en temps réel en fonction de leur déplacement. Le paysage était constitué d'une série d'ambiance tri-dimensionnelle captées en situation urbaine à l'aide du format ambisonic permettant un traitement spatial ultérieur afin de retranscrire les déplacements relatifs du sujet dans la scène. Par ailleurs différentes balises auditives, spatialisées en mode binaural, étaient également disposées le long du parcours. Les résultats principaux de l'étude montrent une amélioration des performances de navigation dans la condition audio-visuelle comparativement à une situation visuelle pure. Cette amélioration est quantifiée par le temps passé à trouver les différentes balises dans la ville, ou à replacer celles-ci lors de la phase de report graphique après l'expérience. Par ailleurs l'analyse des questionnaires établis à l'issue de l'expérience révèle une sensation de présence accrue dans la condition audiovisuelle. La comparaison entre les deux groupes sujets montre une sensibilité supérieure des sujets agoraphobiques aux échelles évaluant différents aspects d'inconfort (cybersickness) pouvant révéler un degré moindre d'intégration multi-sensorielle. Différents aspects de congruence inter-modalitaire, technologiques (latence) ou cognitifs, peuvent en être l'origine et doivent être évalués [Viaud-Delmon04a].

Participants : E. Rio,

Collaborations extérieures : I. Viaud-Delmon, A. Seguelas, Y. Ouallouche (UMR7593)

1.1.3.4 Auditory Watermaze

Les performances de l'audition spatiale sont généralement considérées pauvres par rapport à la vision. La résolution angulaire est plus faible et les tests de localisation auditive pure, menés avec des HRTFs, montrent d'importantes déviations ou confusions sauf pour des sujets entraînés. Cependant l'audition est la seule modalité perceptive qui donne accès instantanément à l'ensemble de la sphère spatiale. Afin d'étudier comment une scène spatiale peut être mémorisée sur la seule base d'indices auditifs et idiothétiques (sensation induite par le mouvement propre du sujet), nous avons reproduit l'expérience du Watermaze de Morris, classiquement utilisée en neurosciences en l'adaptant au contexte auditif. Les sujets masqués sont équipés d'un casque d'écoute sans fil qui délivre un paysage sonore rafraîchi en temps réel en fonction de leur déplacement. Un système de suivi de position utilisant des marqueurs passifs repérés par des caméras infra-rouge envoie la position et

l'orientation de l'auditeur au système de spatialisation qui effectue en temps réel la synthèse de la scène sonore en mode binaural. Le développement technologique est directement issu du projet Listen. La tâche du sujet consiste à se familiariser avec l'organisation spatiale de la scène et à y repérer le plus rapidement un objet sonore caché qui n'est déclenché que lorsque le sujet se rend à une position bien précise. Une fois la phase d'apprentissage achevée et la scène acquise, on peut également introduire des modifications topographiques des balises auditives ou des lois acoustiques utilisées pour le rendu de la scène sonore de manière à évaluer l'impact sur les facultés d'orientation du sujet. Ces expériences ont montré que l'acquisition d'une représentation mentale de l'organisation spatiale de la scène était accessible sur la seule base d'une situation auditive et idiothétique, y compris dans le cas de situations non réalistes sur le plan sémantique ou acoustique. En revanche, la congruence entre la modalité auditive et idiothétique est nécessaire. Par ailleurs les sujets ont unanimement noté un haut degré de présence et d'immersion lors de la tâche qui leur était proposée et qui a été jugée très captivante. Au delà des résultats scientifiques cette expérience a ainsi révélé un potentiel intéressant pour certaines installations sonores qui pourraient utiliser avantageusement le principe d'une privation totale ou partielle de la modalité visuelle au profit d'une modalité auditive assortie de moyens de navigation [Warusfel04c].

Participants : O.Delerue

Collaborations extérieures : I. Viaud-Delmon (UMR7593)

1.1.4. Technologies de reproduction binaurales

La restitution réaliste de sources ponctuelles sur casque d'écoute implique l'usage d'une spatialisation binaurale. Ce type de spatialisation est basé sur le filtrage dynamique de la source sonore par les fonctions de transfert de la tête de l'auditeur (HRTF) et pose deux problèmes majeurs. D'une part, les fonctions de transfert utilisées doivent respecter certaines caractéristiques individuelles qui diffèrent d'une tête à l'autre. D'autre part, une implémentation exacte pour chaque source peut se révéler prohibitive en termes de coût de calcul lorsque le nombre de sources s'accroît. Dans le cadre du projet Listen, l'équipe a poursuivi son travail sur l'optimisation de l'implémentation par utilisation d'un format intermédiaire d'encodage multicanal et la validation perceptive des solutions développées.

La spatialisation sur casque d'une scène sonore, réactive en temps-réel aux mouvements de l'auditeur, est un des points d'étude du projet Listen. Différentes stratégies peuvent être adoptées en fonction des matériaux sonores, en particulier de sa nature (contenu ou format disponible) et de sa fonction dans la scène sonore. Suivant les critères de composition, on peut hybrider différentes techniques de restitution, depuis l'emploi d'un son monophonique (intra-crânien) jusqu'à la spatialisation binaurale de sources ponctuelles.

1.1.4.1 Optimisation et validation perceptive du format d'encodage multicanal pour la restitution binaurale

La constitution d'un format intermédiaire multicanal destiné à l'écoute en binaural est une solution à ces deux problèmes. Il consiste en une représentation de la scène sonore sous la forme d'un flux multicanal. Pour être intégrée à la scène sonore, chaque source doit d'abord être distribuée sur les différents canaux selon des coefficients qui dépendent de sa position dans l'espace (fonctions spatiales d'encodage). L'utilisation de gains variables permet ainsi de réduire le coût de calcul nécessaire à chaque source additionnelle. La scène sonore ainsi

constituée est restituée aux oreilles de l'auditeur après une étape globale de filtrage dont le coût ne dépend pas de la complexité de la scène sonore. Les filtres utilisés pour ce décodage (filtres de reconstruction) peuvent s'adapter aux caractéristiques individuelles de chaque tête.

Les fonctions spatiales et les filtres de reconstruction associés peuvent être obtenus par analyse statistique (PCA) des HRTF. Une nouvelle méthode a été proposée afin d'optimiser ce processus d'analyse statistique et d'équilibrer la qualité de restitution selon les directions de provenance du son et la gamme de fréquence. Il s'agit d'adapter la restitution aux caractéristiques de la localisation auditive (amélioration de la restitution pour les directions et les fréquences clefs) ou aux circonstances de l'application cible (scène sonore limitée au plan horizontal par exemple). Ces optimisations ont fait l'objet d'une série de tests de validation perceptive [Rio03a].

Participants : E. Rio, G. Vandernoot

1.1.4.2 Format d'encodage multicanal pour la restitution binaurale

La palette des stratégies de spatialisation ne se restreint pas au cas des sources ponctuelles. Afin de restituer des scènes complexes éventuellement issues d'enregistrements, l'utilisation du format Ambisonics peut s'avérer nécessaire. Ce format peut être interprété comme une représentation multi-canal des HRTF, pour laquelle les fonctions spatiales d'encodage seraient définies mathématiquement par les harmoniques sphériques. Le décodage d'un tel format a donné lieu à des optimisations similaires à la méthode par analyse statistique, notamment en faisant varier le nombre et la position des directions pour lesquelles la reconstruction est optimisée (haut-parleurs virtuels).

Implémentations :

- Décodage paramétrable de scènes Ambisonics sur casque audio, et objets de manipulation de scènes au format Ambisonics (bibliothèque Max/MSP).
- Nouvelles versions de l'implémentation du binaural dans le Spat~ (simplification et optimisation du modèle existant, implémentation par convolution).
- Développement d'une plate-forme de tests perceptifs basée sur l'exploitation des données du suivi de position (Max/MSP) en relation avec une interface de réponse dans l'environnement Virtools.

Participants : E. Rio

1.1.4.3 Adaptation individuelle

Un problème critique lié à la restitution binaurale est le degré d'adaptation des auditeurs à une synthèse binaurale non basée sur des fonctions de transfert (HRTFs) individuelles qui peut conduire à des artefacts de localisation tels que l'inversion avant-arrière, des excès ou défauts de latéralisation ou encore la localisation intra-crânienne. Ces différences ont pour origine l'influence de la morphologie des individus sur la signature acoustique de la localisation caractérisée par les HRTFs. Une série d'étude est menée pour tenter de fournir des solutions d'adaptation de la synthèse binaurale à la morphologie du sujet.

Adaptation de l'ITD

Le but est de trouver une combinaison linéaire de paramètres morphologique pouvant servir de prédicteur de la dépendance spatiale de l'ITD (inter aural time delay) en fonction du sujet.

La procédure a consisté à soumettre la base de données des variations inter-individuelles d'ITD à une analyse en composantes principales. Cette décomposition permet de caractériser la dépendance spatiale de l'ITD en paramétrant l'influence du sujet. L'étude a prouvé que la première composante délivrée par l'analyse permet de traduire un fort pourcentage des variations spatiales et inter-individuelle et améliore notablement les résultats précédemment obtenus par l'équipe d'Algazi. Une bonne approximation de cette fonction peut être obtenue à partir de la mesure de la circonférence de la tête, paramètre auquel on pourrait facilement accéder dans un cadre pratique d'accueil de public (préalablement au choix de HRTFs adaptées dans le cadre d'une installation sonore basée sur des casques).

Modélisation physique de l'ITD

D'une manière générale la mesure de fonctions de transferts d'oreille (HRTFs) est coûteuse du point de vue de l'infrastructure et du temps nécessaire. Pour éviter cette étape on peut également envisager une modélisation physique permettant de prédire ces HRTFs à partir de la fourniture d'un modèle morphologique (obtenu par un scan 3D plus ou moins simplifié). Une première approche peut consister à utiliser un modèle de tête sphérique à partir duquel on peut estimer différents paramètres tels que l'information d'ITD. Si l'on considère un modèle sphérique avec conduits auditifs centrés l'analyse de l'ITD montre des déviations importantes (de 10 à 20%) par rapport à ce que l'on mesure sur une tête humaine.

Une étude a été menée de manière à introduire un décalage des conduits auditifs sur la sphère et mesurer son influence sur la prédiction de l'ITD. Grâce à cette modification on peut retraduire plus exactement la dépendance spatiale de l'ITD. Une expérience psycho-acoustique est en cours de traitement pour évaluer l'importance de ces modifications au regard des performances de localisation auditive [Busson 04].

Participants : S. Busson, G. Vandernoot

Collaboration extérieure : B. Katz (LIMSI), Rozenn Nicol (FT R&D), P.O. Mattei (LMA)

1.1.4.4 Etude de la plasticité auditive spatiale

Contrairement aux études précédentes qui tentent d'adapter la synthèse aux caractéristiques individuelles cette étude vise à mesurer les capacités d'adaptation des auditeurs à des HRTFs non individuelles. Différentes études ont démontré les possibilités de plasticité de l'audition spatiale, cependant astreinte à une phase d'apprentissage lente. L'idée de la présente étude est de susciter un apprentissage rapide grâce à un protocole mettant en jeu un principe de rétroaction auditif et gestuel.

Le protocole du test consistait à mesurer les performances de localisation en contexte purement auditif (sujets masqués) avant et après une phase d'apprentissage. Deux groupes de sujets participaient à l'expérience, l'un utilisant ses propres fonctions des transferts, l'autre utilisant des fonctions de transferts hybrides pour lesquelles l'information individuelle d'ITD était préservée mais en utilisant la partie à phase minimale non individuelle, choisie dans la base de données parmi les HRTFs les moins bien adaptées au sujet.

La phase d'adaptation était présentée sous forme d'un jeu pendant lequel le sujet était amené à manipuler un « faisceau » sonore (bruit rose intermittents) pour chercher des cibles autour de la tête. Le faisceau était spatialisé en fonction des mouvements de la main du sujet captés par un dispositif de suivi de position électromagnétique. Selon l'appartenance du sujet au groupe contrôle ou au groupe test, la spatialisation binaurale était effectuée avec HRTFs individuelles ou non. Dans ce dernier cas, l'adaptation est « espérée » par un processus d'interaction auditive et proprioceptive.

L'expérience a prouvé les capacités d'adaptation des sujets utilisant des fonctions de transfert non-individuelles. Avant la séance d'apprentissage, les performances de localisation des sujets du groupe test sont moindres que celle du groupe contrôle, notamment au regard de la localisation en élévation et des confusions avant-arrière. Après la séance d'entraînement, les performances du groupe test rejoignent celles du groupe contrôle qui elles n'ont pas varié. Au delà de cette observation, il est intéressant de constater que la recalibration auditive a pu être obtenue sans recours à la modalité visuelle mais par interaction auditive et proprioceptive [Blum03] [Blum04].

Participants : A. Blum,

Collaboration extérieure : B. Katz (LIMSI)

1.1.5. Reproduction de champ sonore

1.1.5.1 Compensation de contexte dans le cadre de la reproduction holophonique

La technique de restitution Wave Field Synthesis (WFS) permet de reproduire les propriétés spatio-temporelle d'un champ sonore sur une audience étendue. Parmi les différentes techniques de réseau de transducteurs envisageables, l'Ircam s'est intéressé à la technologie MAP (multi-actuator panel) qui offre l'avantage d'un dispositif intégré sous la forme de panneaux éventuellement utilisables comme surface de projection dans le cadre d'applications multi-media. Cependant, la constitution d'une couronne de MAP entourant la zone d'auditeurs pourrait créer des réflexions qui peuvent s'avérer gênantes auditivement. Une technique de compensation de l'effet de salle a été étudiée sur la base de procédés d'inversion multicanal. Après analyse de l'agencement géométrique des panneaux, l'idée consiste à séparer l'égalisation champ libre du son direct des sources virtuelles que l'on veut créer et la partie d'égalisation dédiée à la compensation des réflexions indésirables. L'analyse géométrique préalable permet de sélectionner les réflexions qui pourront être objectivement réduites sans provoquer d'artefacts. Lorsque la géométrie s'y prête il est ainsi possible de concentrer la compensation sur les réflexions d'ordre 1, l'atténuation des réflexions d'ordre supérieur étant automatiquement obtenue.

L'étude a fait apparaître la nécessité d'une analyse approfondie des propriétés du champ acoustique que l'on peut produire grâce au système WFS, en particulier sur les lois d'atténuation associées aux sources acoustiques virtuelles qui ne peuvent en toute rigueur égaler les lois naturelles compte tenu des différentes approximations pratiques inhérentes à la technique.

Cependant les simulations effectuées en collaboration avec l'université de Delft montrent la possibilité d'une atténuation des réflexions effective sur l'ensemble de la zone d'écoute [Van Zon04].

Participants : E. Corteel

Collaboration extérieure : R. Van Zon, D. de Vries (University of Delft)

1.1.5.2 Contrôle de la directivité des sources

Les techniques de reproduction holophoniques visent la reconstruction du champ sonore associé aux sources virtuelles. L'application de cette technique pour la synthèse de sources

virtuelles présentes dans un espace clos à l'intérieur duquel les auditeurs peuvent évoluer peut engendrer des effets saisissants et intéressants dans le cadre des applications de réalité augmentée. Cependant la prise en compte de l'interaction du champ acoustique virtuel généré avec la salle d'écoute doit être prise en compte si l'on veut optimiser cette sensation de présence. En particulier, au moment de la synthèse de la source sonore, une importance doit être accordée à la restitution des caractéristiques de directivité de la source et de son interaction avec la salle. Ces caractéristiques sont reconnues comme jouant un rôle dans la perception de la signature acoustique d'un instrument et de la salle. D'un côté ces caractéristiques sont facilement repérables lorsque la source évolue autour de l'auditeur ou réciproquement. Par ailleurs, la directivité se répercute sur la réponse de la salle. Pour un effet de présence optimal il est donc important de veiller à la mise en cohérence de la salle d'écoute avec les propriétés de directivité de la source que l'on cherche à simuler. Cependant la reproduction du champ sonore direct associé à la source virtuelle n'est qu'une approximation de la réalité, liée aux propriétés des bancs linéaires de haut-parleurs utilisés pour la reproduction. Ainsi, l'effet de salle engendré par excitation de la salle par le champ sonore virtuel ne peut être conforme à la réalité. Une étude a été réalisée afin simultanément d'opérer la synthèse d'une fonction de directivité associée aux sources virtuelles, modifiable en temps réel le cas échéant, et de dresser le cadre pour la création d'un effet de salle cohérent. Ces études ont fait l'objet de simulations, de premières réalisations et de tests perceptifs informels [Caulkins 03a] [Caulkins 03b] [Warusfel 04a].

Participants : T. Caulkins, E. Corteel

1.1.5.3 Contrôle de l'effet de distance dans le cadre de la reproduction holophonique

Un test d'écoute a été organisé dans le cadre d'une simulation d'une situation de production de contenu audio utilisant le principe de restitution par Wave Field Synthesis. Le but du test était d'évaluer les possibilités de création d'une perspective sonore en concentrant l'étude sur le contrôle de la distance apparente des sources.

A cet effet, la particularité de l'approche holophonique est de fournir deux moyens d'accès à cette dimension auditive. D'une part, grâce au contrôle de la courbure du front d'onde associé à la source, qui permet de piloter directement la position apparente de l'événement sonore. D'autre part, en ayant classiquement recours au contrôle du rapport énergétique entre le son précoce et le son tardif (effet de salle), dimension utilisée couramment par les ingénieurs du son.

La situation classique consiste à ajuster le niveau du son direct en gardant le niveau de réverbération pour jouer sur la distance apparente de la source. Les études psychoacoustiques menées dans différents laboratoires ont affiné cette proposition et certains logiciels comme le Spatialisateur, propose l'utilisation de règles psychoacoustiques pour optimiser le contrôle de la « présence » de la source. En plus de ce critère, la WFS permet de jouer sur la courbure du front d'onde associé à la source. Même si cet effet de courbure n'est pas directement accessible à l'audition sauf en situation de très grande proximité, cette propriété devient déterminante dès lors que l'auditeur est autorisé à se déplacer dans la scène sonore. C'est l'une des grandes spécificités de cette technique de reproduction. Elle permet notamment la gestion de la profondeur des différentes sources sonores par la sensation des effets de parallaxe lorsqu'on se déplace ou de sensation de variations de niveau naturelles lorsqu'on s'approche ou s'éloigne des sources. Par opposition à la sensation de « présence » obtenue par

contrôle du rapport énergétique entre le champ direct et le champ réverbéré, nous dénommons ce paramètre de contrôle, « distance holophonique ».

L'idée du test était par conséquent d'étudier « le mérite » respectif de ces deux paramètres de contrôle, en vérifiant tout d'abord leur maniabilité et leur perception respective dans le cadre d'une production sonore utilisant la WFS. Dans un second temps, l'objectif était d'étudier la nécessité du maintien de leur contrôle indépendant ou au contraire l'intérêt d'une manipulation conjointe.

Les différentes phases du test ont confirmé, d'une part, la pertinence perceptive de la distance holophonique, dans le cadre d'installations sonores autorisant la navigation de l'auditeur. Par ailleurs, les réponses des sujets ont clairement montré la nécessité d'une distinction entre la manipulation de la « présence » et de la « distance holophonique ». Ces résultats sont importants pour la conception d'outils auteurs dédiés à la création de contenu pour les installations sonores interactives.

Participants : M. Noguès, E. Corteel

1.1.6. Projet Listen

L'année 2003 a vu l'achèvement du projet Listen. Outre les études décrites dans les paragraphes précédents, la dernière année du projet a été consacrée à la mise au point des démonstrateurs qui ont fait l'objet d'expositions publiques au KunstMuseum de Bonn. Deux expositions ont permis d'illustrer les potentialités du système dans les cadres respectifs d'une installation artistique multimédia et dans le cadre d'une application à caractère pédagogique illustrant une nouvelle génération de dispositif d'audio-guide.

L'installation "Raumfaltung", réalisation du plasticien Beat Zöderer et du compositeur Ramon G. Arroyo avec l'assistance de Gerhard Eckel, a permis de valider le concept LISTEN dans le domaine artistique pour en explorant ses capacités à produire un contenu multi-sensoriel. L'exposition « Macke Labor » a permis d'étudier de nouvelles formes de présentation interactive d'un contenu à caractère didactique. Ces expositions ont permis d'accueillir respectivement 4000 et 1600 visiteurs pendant plusieurs semaines.

En conclusion du projet, Listen est apparu comme fournissant un nouveau médium reliant de manière originale un contenu sonore interactif, le corps en mouvement et l'espace. Grâce au dispositif d'écoute et de suivi sans fil l'auditeur-visiteur est immergé dans une scène sonore qui augmente le monde exploré par la navigation.

Différentes voies d'interaction spatiale ont pu être explorées, sous forme de déclenchement de sources, de flux ou de transformations sonores, ou encore sous des formes plus évoluées mettant en œuvre une interaction spatio-temporelle, reliant une partition spatiale ouverte et l'historique de la navigation de l'auditeur. Au cours des études à caractère scientifique il est également apparu la force « d'immersion » inhérente à ces situations interactives et leur acceptation spontanée par les sujets visiteurs. Cette propriété n'est pas nécessairement liée à la construction de mondes réalistes, mais se prête également à des propositions de nature symbolique tant du point de vue du contenu sémantique que de la mise en œuvre des propriétés acoustiques associées au monde sonore synthétisé.

1.1.6.1 Esquisse musicales

Diffusée dans le cadre de l'Atelier de Création Radiophonique (France Culture) le 30 Novembre 2003, *Secrète Lisboa* intègre dès sa conception la possibilité de dispositifs d'écoute multiples : diffusion stéréo sur le réseau hertzien avec possibilité de décodage Dolby Surround, ainsi qu'une installation sous la forme d'un dispositif 5.1.

Les enregistrements de cette fiction documentaire, réalisés dans la ville de Lisbonne, ont été effectués sous deux formats différents : stéréo conventionnelle et prise de son Ambisonics (microphone Soundfield). Cette dernière technique permet de capter une scène sonore selon l'ensemble des directions de l'espace. Sa mise en place a nécessité l'étude d'une solution portable permettant l'enregistrement des 4 canaux que fournit le microphone Soundfield. Lors de la phase de production, différents outils de manipulation et de décodage du format Ambisonics (issus du projet Listen) ont été réalisés sous forme de plug-ins (Pluggo) pour une utilisation en studio sous environnement Protools.

La collaboration avec Cécile Le Prado a été poursuivie afin de décliner le matériau sonore de *Secrète Lisboa*, et l'adapter au dispositif d'écoute du projet Listen. Cette étude a permis de dégager des règles d'écriture propres au contexte de la réalité augmentée. De premières esquisses ont été réalisées et ont donné lieu à plusieurs démonstrations.

Commande de l'Atelier de Création Radiophonique (France Culture) en collaboration avec l'Ircam

Compositeur : C.Le Prado

Participants : R. Kronenberg, M. Poletti, E. Rio

Collaboration extérieure : G. Mardirossian (France Culture)

1.1.7. Projet Carrouso

L'année 2003 a vu l'achèvement du projet Carrouso. Au cours de ce projet, l'ensemble de la chaîne de traitement d'une scène sonore véhiculant ses dimensions spatiales a été élaboré et validé. Les aspects aussi divers que la captation par un réseau de microphones, l'encodage au format MPEG4, la transmission, le décodage de la scène sonore et sa restitution en mode holophonique ont été étudiés et ont fait l'objet du développement d'un démonstrateur complet.

Le projet a fourni l'occasion à l'Ircam de se doter d'un système de restitution holophonique à la conception duquel il a participé. Ont été notamment développés à l'Ircam, un logiciel de création et de manipulation de scènes au format MPEG4, des procédés d'égalisation multi-canal utilisé pour la restitution en mode WFS (Wave Field Synthesis) de la localisation des sources virtuelles, de leur directivité et de leur effet de salle associé.

Un premier prototype constitué d'un réseau modulaire de 56 canaux a été développé et démontré dans plusieurs manifestations publiques et professionnelles (convention de l'AES à Amsterdam en mars 2003, Résonances à l'Ircam en octobre 2003 et Nicéphore days en novembre 2003). Ces démonstrations ont donné l'occasion de premières esquisses de

production audio, sous forme de remixage de pièces existantes (Extraits d'œuvres de Yann Maresh et Nicolas Verrin) ou de création de pièces originales (Rémy Gallichet).

Par ailleurs, le projet a permis d'amorcer une évolution du logiciel Spatialisateur pour son adaptation à la technique de rendu holophonique et pour les contextes de production répartie, utilisant massivement les communications réseau. Cette mise en réseau est utile pour déléguer certaines parties du traitement audio à des unités dédiées (modules de diffusion holophoniques, plugin de spatialisation sous séquenceurs, ...) ou encore pour autoriser le monitoring depuis un logiciel auteur embarqué sur une unité portable sans fil, préfigurant ainsi de nouvelles situations de production.

1.1.8. Publications et communications

Actes de congrès avec comité de lecture

[Baskind 03a] Baskind A., de Cheveigné A., « Pitch tracking of reverberant sounds, application to the spatial description of sound scenes », Proc. of 24th Intern. Audio Engineering Society Conference "Multichannel Audio – the new Reality", June 2003.

[Baskind 03b] Baskind A., « Modèles et méthodes de description spatiale de scènes sonores : application aux enregistrements binauraux », Thèse en Science Mécanique, Acoustique et Électronique. Paris : Université Paris 6 (Pierre et Marie Curie), 2003. [Thèse non publiée]

[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).

[Caulkins 03a] Caulkins T., Corteel E., Warusfel O. « Wave field synthesis interaction with the listening environment : improvements in the reproduction of virtual sources situated inside the listening room », Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003

[Corteel03] Corteel E., Nicol R., Listening Room Compensation for Wave Field Synthesis. What can be done? Proceedings on the 23rd International Conference of the Audio Engineering Society, Copenhagen, Denmark, May 2003.

[Noguès 03a] Noguès M., Corteel E., Warusfel O. « Monitoring Distance Effect With Wave Field Synthesis », Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003.

[Rio 2003a] Rio E., Vandernoot G., Warusfel O., « Perceptual Evaluation of Weighted Multi-Channel Binaural Format » Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003.

[Viaud-Delmon 2004a] 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 2004b] 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.

[Vos 03] Vos H.J., Warusfel O., Misdariis N., de Vries D., « Analysis and reproduction of the frequency spectrum and directivity of a violin », Dutch Acoustical Society Conference, Septembre 2003.

[Warusfel 2004a] Warusfel O., Corteel E., Misdariis N., Caulkins T., « Reproduction of sound source directivity for future audio applications », Proc. ICA 2004.

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

Résumés publiés dans des revues à comité de lecture

Actes de congrès sans comité de lecture

[Deschamps 03b] Deschamps M., « Investigation of interaction between recording or mixing parameters and the spatial subjective attributes in the frame of 5.1 multicannel », 115th convention of the Audio Engineering Society, New York-USA, (2003).

[Kuhn03] Kuhn C., Pellegrini R., Leckschat D. and Corteel E., "An Approach to Miking and Mixing of Music Ensembles Using Wave Field Synthesis", 115th convention of the Audio Engineering Society, New York - USA, (2003).

[Väänänen 03a] Väänänen R., « User interaction and authoring of 3D sound scenes in the Carrouso EU project », Proc. 114th convention of the Audio Engineering Society, (2003).

[Van Zon04] van Zon R.J.A., Corteel E., de Vries D., and Warusfel O., Multi-Actuator Panel (MAP) loudspeakers: how to compensate for their mutual reflections? Proc. 116th convention of the Audio Engineering Society, (2004)

Travaux universitaires

Mémoires de maîtrise, DEA et d'écoles d'ingénieurs

[Blum 03], Blum A., « Étude de la plasticité du système auditif en localisation - Application au problème de l'individualisation en synthèse binaurale », Mémoire de DEA, Université Aix-Marseille, Sept 03.

[Caulkins 03b], Caulkins T., « Restitution d'un effet de salle cohérent avec le lieu d'écoute dans le cadre de la Wave Field Synthesis », Rapport de DEA ATIAM, Université Marseille. Juillet 03.

[Nogues 03b] Nogues M., « », Mémoire d'ingénieur, Conservatoire National Supérieur de Musique de Paris. Juill 2003.

[Verron 03] Verron C., « Selection de fonction de transfert binaurales individuelles à partir d'un protocole de mesure simplifié »; Mémoire d'ingénieur. Septembre03

Thèses

Baskind A., « Modèles et méthodes de description spatiale de scènes sonores », Thèse de doctorat, Université Paris 6, 9 décembre 2003

Rapports de recherche

Väänänen R., Projet Carrouso IST- IST-1999-20646 – Deliverable D8. User Interface – Jun 03

Katz B., Vandernoot G., Warusfel O., Projet Listen IST-1999-20646 – Deliverable D4.3 Individual adaptation. Jun 03

Rio E., Rupprechter B., Warusfel O., Projet Listen IST-1999-20646 – Deliverable D4.4 Sound Pick-up. Jan 03

Warusfel O., Projet Listen IST-1999-20646 – Deliverable D4.5 Prototype of the Rendering Engine, Jan 03

Warusfel O., Goiser A., Projet Listen IST-1999-20646 – Deliverable D8.1 Display evaluation; Spet 03

Rio E., Vandernoot G., Warusfel O., Projet Listen IST-1999-20646 – Deliverable D8.2 Evaluation of binaural rendering. Jun 03

Warusfel O. et al. Projet Listen IST-1999-20646 – Deliverable D8.3 Modelling: Visual/Auditory coherence, modelling and authoring. Jun 03

Eckel G. & al, Projet Listen IST-1999-20646 – Deliverable D1 Final Report, Dec 03

Conférences invitées

Journée "Méthodes d'enregistrement et de restitution sonore pour l'évaluation de la perception des bruits de transports", organisée par la SFA, le groupement de recherche CNRS "bruit des transports", et le LVA (Laboratoire Vibrations Acoustique, INSA de Lyon). 15 septembre 2003.

« *Technique de reproduction binaurale : encodage multi-canal et décodage individuel*
Intervenant : Emmanuel Rio

Diffusion de connaissances

Corteel E., Roux S., Väänänen R., Warusfel, O. « Demonstration d'un système de reproduction de champ sonore basé sur la Wave Field Synthesis », 114th Conv. of the Aud. Eng. Soc., Amsterdam, Mars 2003.

Warusfel O., Nicolas F., Misdariis N., « Contrôle de la directivité et source sonore virtuelle », Colloque Son Multicanal », Ircam, Oct. 2003.

Corteel E., Roux S., Warusfel, O. « Demonstration d'un système de reproduction de champ sonore basé sur la Wave Field Synthesis », Nicéphore Days, Châlon/Saône, Nov. 2003.

Warusfel O., « Acoustique virtuelle et Spatialisation », La semaine du son, Ircam, Jan. 2004.

Corteel E., « Système de reproduction sonore du futur », La semaine du son, Ircam, Jan. 2004.

Colloques et séminaires

Baskind A., « Extraction automatique de descripteurs de scène sonore spatiale » – Séminaire Ircam – Avril 2003.

Corteel E., "Création et manipulation de scènes sonores pour la Wave Field Synthesis", Cycle de conférences consacré aux questions de la spatialisation sonore, ENS Louis Lumière, décembre 2003.

Viaud-Delmon I., Warusfel, O. « Traitement des informations multisensorielles en réalité virtuelle : application à la psychopathologie » – Séminaire Ircam – Avril 2003.

Väänänen R., Delerue O. « Outils auteur pour la spatialisation », Séminaire interne Ircam, juin 03.

Viaud-Delmon I., Warusfel, O. « Traitement des informations multisensorielles en réalité virtuelle : application à la psychopathologie » Ecole Réalité Virtuelle – Marseille. Mai 2003

Emissions radiophoniques et télévisées, entretiens journalistiques, animations

Entretien pour le quotidien "Le Monde" mercredi 2 juillet 2003