



Ears Wide Open

3 journées d'études - SFA, AES, Orange Labs

Rennes, les 11, 12 et 13 mars 2008



Programme des conférences

Lundi 11 mars 2008 :

Aspects fondamentaux de la spatialisation

-- 14:00-14:45 --

Sophie Savel (LMA-CNRS)

Evaluations psychoacoustiques des performances humaines de localisation auditive en environnements réels et virtuels

Contrairement à la fréquence, l'espace n'est pas directement codé par l'oreille. Pour estimer la position exocentrique d'une source sonore, le système auditif extrait un certain nombre d'indices acoustiques : différences interaurales de phase (signaux basse fréquence entretenus), de structure fine des enveloppes (signaux complexes) et d'intensité (signaux haute fréquence), et différences de coloration spectrale (issues de la morphologie de l'auditeur, "HRTFs", pour les signaux large bande). Depuis un siècle, la psychoacoustique s'est intéressée à l'évaluation des différents types de



résolutions et d'erreurs issues du traitement de chacun des ces indices (résolution gauche-droite, résolution azimutale absolue et différentielle, confusions avant-arrière, résolution de la hauteur) en situation d'écoute réelle. Plus récemment, les travaux portent sur la dégradation de ces performances en milieu bruyant et chez les malentendants, et sur leur dégradation en environnement sonore virtuel (sources filtrées par simulation des HRTFs personnalisées ou non-personnalisées). Les recherches les plus actuelles portent sur la mise en évidence des différences individuelles de performance dans ces différents environnements d'écoute, sur l'identification de l'origine de ces différences, et sur leur contournement par apprentissage perceptivo-cognitif.

-- 14:45-15:30 --

Nick Zacharov (Delta Danemark)

Méthodes d'évaluation des systèmes audio

-- 16:00-16:45 --

Danièle Dubois et Jean-Dominique Polack (LAM, Institut Jean Le Rond d'Alembert, Université Pierre et Marie Curie/CNRS)

Reproduction sonore et validité écologique : 10 ans de travaux au LAM

Depuis maintenant plus de 10 ans, le Laboratoire d'Acoustique Musical (LAM) effectue des études sur la prise de son et sa restitution, en particulier dans le cadre de thèses consacrées aux environnements sonores urbains. Ces études ont confirmé certains résultats techniques, comme la supériorité de la stéréo de phase pour tout ce qui concerne la localisation des sources et la lisibilité des scènes sonores, ou la supériorité d'Ambisonics pour l'immersion dans une scène sonore. Ces résultats seront rappelés dans l'exposé, mais n'en constituent cependant pas le point principal.

Nous insisterons en effet sur la contribution originale du laboratoire au thème de la journée à travers le développement d'une démarche d'évaluation de la prise de son, basée sur le concept de validité écologique. Ce concept, introduit par Gibson en 79



pour la vision, repose sur l'hypothèse selon laquelle "les valeurs et le sens de l'environnement peuvent être perçus directement", en fonction de l'affordance des stimulations. Transposée dans le domaine sonore, la restitution va s'attacher à reproduire les significations accordées aux sons. Cela revient à déplacer l'évaluation de la prise de son dans le champ de la psychologie, et à positionner la "vérité" physique, qui s'attache à reproduire le plus exactement possible un champ sonore, en regard d'une restitution des effets sonores, qui font "sens" pour une diversité d'auditeurs tels les artistes, les mélomanes, et ... les ingénieurs du son. Ce cadre psychologique impose en effet que ce soit l'auditeur qui serve de référence. La question est alors d'évaluer dans quelle mesure le jugement de l'ingénieur du son que la tradition allemande appelle un Tonmeister, qui se manifeste dans leur expertise à produire des effets sonores, est partagé par les auditeurs concernés.

À cette fin, nous avons développé une démarche expérimentale qui permette de tester en laboratoire quelles conditions de restitution induisent une écoute, des traitements cognitifs et des jugements perceptifs qui garantissent les effets voulus (reproduction de l'expérience quotidienne, effets réalistes, effets artistiques et émotionnels...). Cette démarche consiste à choisir dans un premier temps des séquences représentatives, puis à décider d'un système de reproduction sonore adapté aux différents objets sonores et aux effets souhaités. La « validité écologique » du dispositif technique est évaluée à partir d'une analyse sémantique des discours ou de jugements sur des échelles sémantiques qui donnent accès aux interprétations et ressentis des sujets. Cette démarche impose ainsi une coopération pluridisciplinaire d'acousticiens, d'ingénieurs du son mais aussi de psychologues, sémanticiens et psycholinguistes et sera illustrée à l'aide d'exemples de travaux réalisés au LAM, en particulier ceux de V. Maffiolo, C. Vogel, et de C. Guastavino.

Bibliographie :

Thèses

Guastavino C. (2002) Étude sémantique et acoustique de la perception des basses fréquences dans l'environnement sonore urbain. Thèse de Mécanique, Université Paris 6.

Maffiolo V (1999) Caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain, Thèse de l'Université du Maine.

Vogel C. (1999) Etude sémiotique et acoustique de l'identification des signaux sonores d'avertissement en contexte urbain., Thèse de mécanique, Université Paris 6.

Raimbault, M. (2003) Simulation des ambiances sonores urbaines : intégration des aspects qualitatifs. Thèse de l'Université de Nantes.



Publications (une bibliographie plus approfondie sera disponible lors des journées)

Dubois, D (2007) "From psychophysics to semiophysics: Categories as acts of meaning. A case study from olfaction and audition, back to colors". In Plümacher, Martina and Peter Holz (eds.) *Speaking of Colors and Odors*, 167–184.

Guastavino, C., Katz, B., Polack, J-D., Levitin, D., & Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica*, 91(2), 333-341.

Guastavino, C., & Katz, B. (2004). Perceptual evaluation of multi-dimensional spatial audio reproduction. *Journal of the Acoustical Society of America*, 116(2), 1105-1115.

Guastavino, C., Larcher, V., Catusseau, G. & Boussard, P. (2007). Spatial audio quality evaluation: Comparing transaural, ambisonics and stereo. Invited paper at the 13th International Conference on Auditory Display - ICAD 2007, June 26-29 2007, Montreal, Canada.

Maffiolo, V., Castellengo, M., Dubois, D. (1998). Qualité sonore de l'environnement urbain : sémantique et intensité. *Acoustique et technique*, 16 déc 1998.

Maffiolo, V., David, S., Dubois, D., Vogel, C., Castellengo, M., Polack, J.-D. (1997). Sound characterization of urban environment, in *Internoise 97*, Budapest, 25-27 Aout 1997.

Maffiolo, V., David, S., Vogel, C., Polack, J.-D., Castellengo, M. Dubois, D. (1997). Ambiances sonores représentatives d'une ville : le cas de Paris, 4ème Congrès français d'acoustique, Marseille, 14-18 avril 1997.

Raimbault, M., Dubois, D. (2003). Subject's opinions and attitudes about various urban soundscapes, *Actes du colloque Euronoise*, Naples 2003.

Raimbault, M., Dubois, D. (2003). Perceptual judgements about urban soundscapes, *Actes du colloque Euronoise*, Naples 2003.

Vogel, C., Dubois, D., Polack, J.-D., Castellengo, M. (1998) Perception of warning signals in urban context, in *Actes du colloque « Internoise 98 »*, Wellington.

Vogel, C., Polack, J. D., Castellengo, M., Dubois, D. (1999) Perception and meaning of warning signals in urban context, in *Actes du colloque « 137th Meeting of the Acoustical Society of America et Forum Acusticum »*, Berlin, 14-19 mars 1999, 1373



-- 16:45-17:30 --

Dave G. Malham (Music Research Centre, University of York)

Ambisonics : Bases fondamentales d'Ambisonics,

Le monde est un espace acoustique à trois dimensions. Très tôt dans l'évolution animale, la capacité à utiliser le son dans les trois dimensions pour prévenir ou informer est devenu d'une importance vitale pour la plupart des espèces animales. Jusqu'à la fin du dix-neuvième siècle, nous avons découvert des moyens pour transmettre le son à travers l'espace et le temps. Malheureusement, les techniques utilisées à l'origine perdaient la plupart des attributs spatiaux présents dans le champ sonore de départ. Depuis, les ingénieurs du son se sont évertués à rétablir ces attributs.

Cette présentation va s'intéresser à l'une des approches les plus prometteuses pour enregistrer et restituer un espace sonore 3D : Ambisonics. Nous balaierons les principes de base et l'histoire d'Ambisonics, en nous intéressant à la relation avec les premières expériences stéréophoniques d'Alan Blumlein, avant de proposer un bref survol de l'état actuel de la technologie.

The world is a three dimensional acoustic space. Very early on in the evolution of animal life, the ability to utilise sound in three dimensions to warn or inform became of vital importance to most animals. Towards the end of the nineteenth century, we discovered ways of transmitting sound through space and time. Unfortunately, the original techniques used lost most of the three dimensional attributes that were present in the original soundfields and audio engineers have been trying to restore the losses ever since. This presentation will look at one of the most promising approaches to recording and recreating 3-d acoustic space, Ambisonics. It will cover the basic principles and history of Ambisonics, looking at its relationship with the early stereo experiments of Alan Blumlein before giving a very brief overview of the current state of the technology.



Mardi 12 mars 2008 :

Présentation des systèmes d'enregistrement

-- 8:30-9:30 --

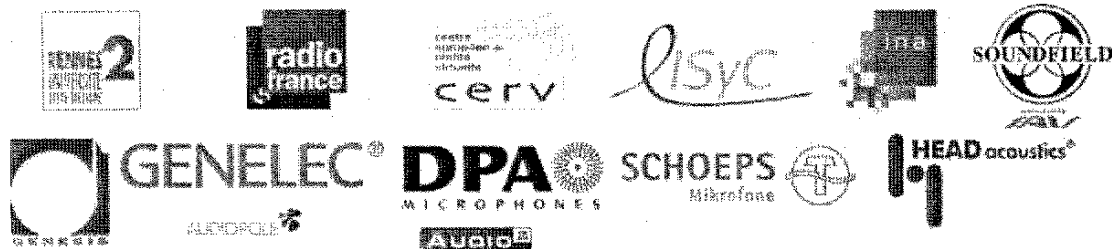
Michael Williams (AES France)

Multichannel Microphone Array Design (MMAD)

Le développement de systèmes de microphones est, pour l'enregistrement stéréo comme pour l'enregistrement multicanal 5.0 ou le Blu-ray 7.0, directement dépendant de la configuration du système de reproduction. Cette présentation décrit l'intersection entre les caractéristiques psychoacoustiques de l'environnement d'écoute multicanal et les caractéristiques physiques du dispositif microphonique, ceci dans le contexte où un microphone alimente directement une enceinte de diffusion, sans matriçage préalable. Quel degré de réalisme peut-on atteindre dans la reproduction naturelle du champ sonore?

L'utilisation du processus « MMAD » permet de choisir parmi un nombre presque infini de configurations des microphones pour répondre aux exigences de chaque situation d'enregistrement. Les caractéristiques de base de « Front Segment Coverage », avec en même temps le processus de « Segment Offset » pour obtenir le « Critical Linking », font partie intégrale du processus « MMA Design ». Cependant, la prise en compte d'autres critères de sélection est nécessaire afin d'assurer le choix optimal du système d'array le plus adapté aux circonstances de l'enregistrement.

Jusqu'à maintenant, la seule solution disponible pour le « main microphone array » pour un enregistrement multi-format est d'utiliser des « arrays » individuelles pour chaque format. Cette présentation montre comment il est possible de remplacer cette jungle de micros par un seul « array » de 5 ou même 7 microphones, qui peut générer des signaux directement compatibles avec 6 formats standard : mono, stéréo à deux et trois canaux, la quadraphonie à 4 canaux, le multicanal à 5 canaux et finalement le Blu-ray à 7 canaux. Le format de reproduction peut être choisi soit pendant la production



en fonction du support média utilisé, ou par le consommateur à partir d'un produit multimédia en fonction de leur propre configuration d'écoute.

-- 9:30-10:30 --

Jérôme Daniel (Orange Labs)

Higher Order Ambisonics (HOA)

Encore très peu connue du monde des ingénieurs du son, l'approche de spatialisation sonore "High Order Ambisonics" (HOA) a fait l'objet de développements théoriques et technologiques depuis une douzaine d'années. L'intérêt qu'elle présente réside en ses "3 visages": c'est en effet à la fois un format de représentation audio 3D générique et flexible, une technologie rationnelle et efficace de la captation sonore 3D, et une technologie de reconstruction de champ sonore adaptable à une grande variété de systèmes de restitution. HOA hérite ainsi des propriétés avantageuses de l'Ambisonics « traditionnel » (d'ordre 1), tout en réduisant les inconvénients et en repoussant les limites, notamment grâce à la notion de résolution spatiale accrue (ce qu'on entend par Ambisonics « d'ordre élevé »).

Au cours de cette présentation, un aperçu « vulgarisé » de l'approche HOA sera d'abord proposé, complété par un éventail d'angles de vue et d'analogies qui permettent d'appréhender l'approche depuis différents domaines techniques et scientifiques (ingé-son, traiteur du signal, acousticien...). Nous en illustrerons les bénéfices en termes de reproduction spatiale à diverses échelles : reconstruction « holophonique » sur zone étendue (ordres très élevés) ; reconstruction « binaurale » lors d'une écoute individuelle centrée ou au casque ; restitution compatible avec les configurations surround standard ; et stabilisation du rendu pour les écoutes de groupe grâce à une bonne séparation spatiale. Le principe des antennes microphoniques pour la captation 3D HOA sera ensuite exposé, donnant à l'occasion une nouvelle perspective au clivage trompeur entre systèmes microphoniques non-coïncidents versus coïncidents. Un état de l'art rendra compte des microphoniques existants pour la prise de son HOA (prototypes de laboratoire et produits commercialisés). Il décrira les performances actuelles ainsi que les perspectives d'amélioration.

Enfin quelques enjeux actuels seront évoqués : la marge d'amélioration des éléments technologiques (captation, restitution...) ; la standardisation du format ; l'évolution des outils de post-production ; les usages.



-- 11:00-12:00 --

Radio-France

Enregistrements et diffusions multicanaux à Radio France

-- 12:00-13:00--

Rozenn Nicol (Orange Labs)

Les Technologies Binaurales

Au quotidien, nous localisons les sources sonores qui nous entourent à partir des deux signaux acoustiques captés par les tympans. Si ces deux signaux suffisent, c'est qu'ils portent en eux toutes les informations nécessaires à une perception auditive de l'espace sonore dans les trois dimensions. De ce constat sont nées les technologies dites "binaurales".

Une prise de son binaurale utilise deux microphones placés à l'entrée du conduit auditif. Pour la restitution, l'idéal est d'écouter l'enregistrement binaural avec un casque, de telle sorte que les signaux enregistrés soient restitués à l'endroit même où ils ont été captés. Il convient de corriger la réponse du casque afin de garantir la transparence de la restitution et d'assurer un rendu correct et convainquant de la scène sonore. La restitution des signaux binauraux peut aussi se faire sur un dispositif de deux haut-parleurs, à condition d'annuler les trajets croisés qui interviennent entre chaque haut-parleur et l'oreille controlatérale.

L'encodage spatial des technologies binaurales est basé sur les indices naturels de localisation (différences interaurales de temps et d'intensité, indices spectraux). Les HRTF (Head Related Transfer Function), qui sont les fonctions de transfert modélisant la propagation acoustique entre une source sonore et l'entrée du conduit auditif, décrivent ces indices de façon compacte. Ce sont ces HRTF qui contrôlent la spatialisation de la scène binaurale, en exploitant principalement les différences de temps d'arrivée entre les deux oreilles et le filtrage directionnel introduit par les réflexions et diffractions sur le



pavillon, la tête et le haut du torse de l'auditeur. L'ensemble de ces phénomènes dépend fortement de la morphologie de l'auditeur. Il en résulte que la spatialisation binaurale se démarque par un encodage à la fois spectral et individuel de la direction des sources.

Ces différents aspects seront présentés et discutés dans l'exposé, avant de conclure sur les applications des technologies binaurales.

-- 18:45-19:45 --

Bergame Périaux (INA)

Etat des lieux sur la formation multicanal à INA Formation, cohérence de l'image sonore et prise de son multicanal

Depuis une dizaine d'années, l'INA Formation est impliquée dans le domaine du son multicanal au travers de son offre de formation multicanal complète et mise à jour chaque année. En effet, face à l'arrivée de la Télévision Haute Définition, au passage à la Radio Numérique, et au développement des nouveaux supports audio et vidéo, les entreprises et autres acteurs de l'audiovisuel voient leurs différents métiers évoluer et viennent acquérir à l'INA les nouvelles compétences associées à cette évolution technologique et culturelle.

L'INA Formation répond d'une part à une demande conséquente sur les différents aspects technologiques de cette évolution - la TVHD en est un bel exemple par sa chaîne de production et son nouveau format le Dolby E imposant de nouvelles méthodes de travail. Parallèlement, la radio numérique introduit aujourd'hui également de nouvelles normes de diffusion utilisées par le son 5.1 et rejoint très certainement les mêmes problématiques de diffusion que celles rencontrées par la Télévision. D'autre part, elle intègre les aspects production et réalisation multicanal qui induisent un changement culturel au travers d'une écoute nouvelle, mettant en scène un espace et des concepts de fabrication nouveaux et plus complexes à gérer.



Aujourd'hui, le point de départ d'une formation multicanal, est un passage obligé par l'analyse d'une nouvelle image sonore et ses nouveaux critères qu'on peut lui associer. Cette première étape facilite l'apport des connaissances sur la prise de son multicanal, sur les concepts de mixage et la réalisation dans ce nouvel espace sonore. L'analyse de l'image sonore multicanal sera ainsi développée durant l'intervention, accompagnée de la présentation des systèmes de prise de son les plus couramment utilisés dans les différents domaines de la production.



Jeudi 13 mars 2008 :

La spatialisation en contexte

-- 9:00-9:45 --

Alexis Baskind

Utilisation de techniques binaurales/transaurales mixtes en post-production 5.0 et 5.1: principes esthétiques et illustrations sonores

L'un des enjeux principaux du mixage est de donner à l'auditeur la plus grande lisibilité d'écoute possible dans la restitution d'une œuvre musicale. De ce point de vue, l'ouverture spatiale permise par la norme 5.1 ITU apporte un plus considérable par rapport à la stéréophonie. Cependant ce standard hérite d'une conception

cinématographique, et induit généralement une instabilité des éléments latéraux et arrière en raison du grand écart angulaire entre les haut-parleurs. Le travail présenté ici vise à pallier ce défaut, par l'adjonction aux méthodes traditionnelles de spatialisation (par quintuplés microphoniques et appoints monophoniques placés au panoramique d'intensité) de techniques basées sur la synthèse binaurale et le traitement transaural. La présentation du principe sera illustrée par des extraits issus de mixes de "Répons" (P. Boulez) et "Shadows" (P. Eötvös) réalisés par Jean-Marc Lyzwa, du CNSMDP.



GENELEC®

DPA
MICROPHONES

SCHOEPS
Mikrofone



HEAD acoustics®



-- 9:45-10:15 --

Olivier Warusfel (STMS-IRCAM-CNRS)

Cognition spatiale auditive

Les situations sonores immersives proposées par la réalité virtuelle, les jeux ou encore les installations sonores artistiques apportent un éclairage nouveau sur la perception spatiale auditive en soulignant l'importance de son interaction avec d'autres modalités sensorielles comme la vision et la proprioception. Sans prétention d'exhaustivité, nous nous proposons de décrire quelques pistes de recherches actuelles consacrées à la cognition spatiale auditive, notamment à travers l'étude de processus d'intégration multi-sensorielle. Ces études, originellement issues du domaine des neurosciences comportementales, soulignent l'importance des aspects sensori-moteurs dans la formation de notre représentation auditive de l'espace et relativisent les analyses basées sur les seuls indices acoustiques. Nous tenterons également de dégager, lorsque cela semble pertinent, les implications du point de vue technologique et des pratiques de création de contenus sonores interactifs.

-- 10:30-11:15 --

Cécile Le Prado (CNAM, ENJMIN)

L'espace comme simulateur d'écriture / De l'espace capté à l'espace restitué / Installations sonores et expérimentations interactives

En tant que compositeur, je travaille depuis très longtemps le son dans l'espace, mettant en rapport un paysage sonore imaginaire et un espace physique d'écoute.

L'espace capté, l'espace composé et l'espace restitué sont devenus des éléments incontournables de mon travail de composition. Mes installations sonores sont liées au paysage, à la cartographie et à la mémoire des lieux. Il s'agit dans la phase de composition de l'espace,



de chorégrapheur, de décrire le comportement des sons dans l'espace et dans le temps.

En partenariat avec le laboratoire d'acoustique de l'Ircam, j'ai pu à deux reprises réaliser des esquisses de prise de son au format ambisonic, mettant en regard la captation en multicanal et la restitution en 5.1 ou autres dispositifs sur 4 à n enceintes. À l'occasion du projet de recherche « Listen Lisboa » et auparavant dans le cadre de la pièce « Vocatifs », j'ai conjugué cet aspect avec une interaction entre l'auditeur, qui se déplace dans l'espace d'écoute, et sa perception du paysage sonore virtuel.

-- 11 :15-12 :00 --

Claude Bailblé

Problèmes de spatialisation en son multicanal

Avec l'invention du montage, le cinéma s'est émancipé de la scène frontale des origines, assemblant la suite de points de vue en une sorte de voyage spatio-temporel parfaitement centré. La diffusion monophonique – réellement ponctuelle, virtuellement omnidirectionnelle – a intégré le dispositif en aménageant une audibilité réduite, certes, mais totalement compatible avec les changements de plans. L'objet visuel y dicte sa place à la source sonore (in ou off). Aujourd'hui, le son multicanal se généralise dans le cinéma commercial, depuis – entre autres – La guerre des étoiles, Les aventuriers de l'arche perdue, Il faut sauver le soldat Ryan... Immersion spatiale, effets pyrotechniques ou telluriques, sons omniprésents, émotions garanties. Pourtant le système narratif mis en œuvre au temps du muet compose à l'écran un espace supra-directionnel (quoique toujours frontal) en désaccord possible avec l'espace sonore englobant (la "géode auditive" de la salle). On a cherché à résoudre les inévitables problèmes de scénographie qui en résultent en s'abstenant de croiser les espaces, en recherchant au contraire la convergence audio-visuelle. Il s'agit le plus souvent de retourner à la transparence du dispositif, tout en profitant habilement des possibilités expressives de la spatialisation. On pourrait aussi bien explorer la dissociation attentionnelle (œil/oreille), où se jouerait une nouvelle forme de montage horizontal, un peu à la manière du triple écran d'Abel Gance.

