

ACTUALITÉ

WAVEFIELD SYNTHESIS : PROMESSES ET LIMITES POUR LE FUTUR DE LA MISE EN ESPACE DU SON . . .

Après une quinzaine d'années de recherche et d'expérimentation, le principe de « reconstitution de scènes sonores par front d'ondes planes » dit Wavefield Synthesis (WFS) en arrive tout doucement à devenir une technique de spatialisation utilisable. Alors que les premiers développements commerciaux et les premières utilisations artistiques voient le jour (cf. reportage sur l'AES / précédent numéro), il est donc peut-être temps d'en connaître un peu plus...

LES LIMITES DE LA STEREOPHONIE

Pour la plupart des musiciens, des ingénieurs du son, des amateurs de musique et de création sonore, la stéréophonie constitue le principal pilier permettant le développement d'une «spatialité» du son. Cependant, cette technique, ou plutôt cet ensemble de techniques présente quelques limites. La principale est sans doute l'instabilité des positions des sons résultant de la position occupée par l'auditeur par rapport aux enceintes acoustiques. En stéréophonie, dès que l'auditeur n'est plus à même distance des deux enceintes, la perspective est biaisée et l'une des voies devient vite dominante. Une écoute stéréophonique optimale suppose donc que l'auditeur se place sur un axe défini par les points équidistants entre les deux enceintes («sweet spot»). Si cette limite est en général admissible pour une écoute domestique, elle constitue un réel problème lors d'une écoute collective. C'est d'ailleurs pour résoudre ce problème que le cinéma utilise depuis longtemps une enceinte centrale permettant la perception d'un son au centre de la rampe stéréophonique, cela quelle que soit la position de l'auditeur dans la salle. Mais, pour les concerts faisant appel à un dispositif de projection électroacoustique, ce problème n'est pas résolu. En effet, dans une configuration de type multicanal, à la fragilité de l'équilibre entre canal gauche et canal droit, s'ajoute celle d'un second équilibre entre enceintes frontales et arrières. Car, cette fois, la position idéale n'est plus restreinte à un axe, mais théoriquement à un seul point. Même si, en pratique, la plupart des auditeurs percevra des *mouvements* sonores, nette-

ment plus restreint sera celui des auditeurs susceptibles de percevoir *correctement* le positionnement et les déplacements de tous les sons.

Un autre défaut de la stéréophonie est qu'elle ne reproduit pas réellement la dimension de la profondeur. Certes, elle reproduit certains indices utilisés par la perception pour situer la distance d'une source sonore (rapport entre son avec et sans réverbération, intensité et richesse spectrale essentiellement). Mais ces indices fonctionnent surtout lorsqu'il s'agit de reproduire une scène sonore existante (captation). Sur des sons sans lien avec une réalité préexistante, comme les musiques électroacoustiques ou électroniques nous en proposent quotidiennement, ces indices ne fonctionnent pas toujours avec la même acuité. On parle d'ailleurs de «rampe stéréophonique», ce qui en dit long sur la limitation du positionnement apparent des sons à un écran sans relief et il reste impossible en stéréophonie *classique* de placer un son en-dehors de cette rampe – la perception d'*extra-largeur* ne s'obtenant qu'artificiellement par des procédés d'inversion de phase.

Enfin, une autre limite de l'espace stéréophonique est qu'il n'est pas du tout enveloppant, alors que, même dans le cas où une source sonore est en face de nous (musiciens en concert, interlocuteur en conversation), la perception sonore s'exerce sur 360° (sans compter la troisième dimension), notamment à cause des réflexions causées par l'acoustique du lieu. Ce problème a été partiellement résolu par les dispositifs en multicanal via l'adjonction des enceintes arrières, mais avec les limites expliquées plus haut, le multicanal fonctionnant, lui aussi, sur base de sources «fantômes».

Comme on le voit, certaines limites sont repoussées par le multicanal mais d'autres sont aggravées ou subsistent. D'autres solutions en matière de spatialisation sont donc recherchées.

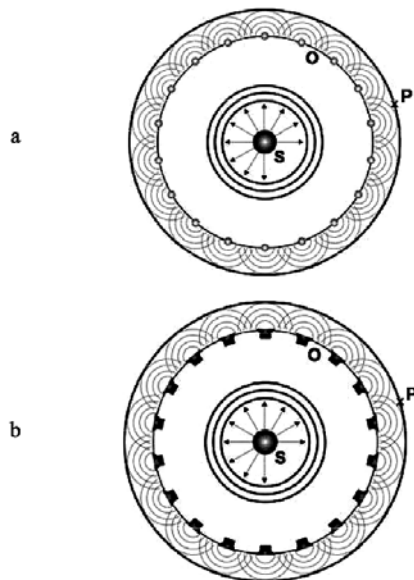
WAVEFIELD SYNTHESIS

Parmi les techniques qui sont ou qui ont été investiguées, la Wavefield Synthesis est non seulement prometteuse mais son application pratique est déjà testée intensivement et sa diffusion à grande échelle pourrait intervenir dans un futur proche.

Alors que la stéréophonie est essentiellement une photographie sonore (une «phonographie») dans le sens où elle fixe les positions et l'acoustique du lieu de captation avec le son capté, la WFS s'apparente plutôt à de l'holographie. En effet, elle reproduit un champ sonore complet et non une image de celui-ci, ce qui replace l'auditeur dans une situation d'écoute beaucoup plus proche de celle de l'écoute binaurale naturelle.

Si les premières publications à son propos remontent à 1987-88, elles s'inspirent entre autres des travaux sur la stéréophonie à trois points de Snow et Steinberg (1953) issus de la notion de «rideau sonore» développée intuitivement par ces deux chercheurs et Fletcher dès les années 30! L'idée du rideau sonore est de reproduire une scène sonore via un ensemble d'enceintes espacées de manière régulière devant les auditeurs.

D'un point de vue théorique, la WFS repose avant tout sur le principe de Huygens, du nom d'un savant hollandais du 17^{ème} siècle. Comme on le sait, lorsqu'une source sonore ponctuelle émet un son, les ondes à proximité de la source sont sphériques. Au plus on s'éloigne de cette source, au plus la surface des ondes devient plane. Le principe de Huygens indique qu'on peut obtenir une onde plus ou moins plane telle qu'on l'aurait avec une source lointaine (dite source primaire), en associant les fronts d'ondes sphériques d'une multitude de sources (dites sources secondaires) placées plus loin de la source virtuelle (ou plus près de l'auditeur) (voir figure 1).

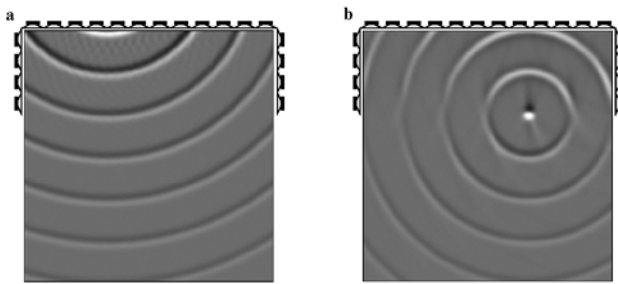


[Figure 1 : Principe de Huygens. Modèle théorique (a) et (b) application à la WFS (b). S = source primaire ; O = surface d'onde où sont placées les sources secondaires ; P = point où un front d'onde équivalent en (a) et (b) est obtenu]

Déjà appliqué aux ondes sismiques, c'est par les travaux de Berckhout, de Vries, Boone et leurs collaborateurs à l'université de Delft (Pays-Bas) que le principe de Huygens a été appliqué à la spatialisation.

La WFS se fonde sur le principe de Huygens en utilisant comme sources secondaires des haut-parleurs. Essentiellement en introduisant par traitement de signal des différences entre les signaux envoyés aux différents canaux d'un tel dispositif, on peut créer des fronts d'ondes très proches de ceux qu'on aurait avec une source acoustique réelle.

Avec un tel dispositif, une source est toujours localisée au même endroit quel que soit le point d'écoute de l'auditeur dans le lieu d'audition. De plus, la source peut être localisée par l'auditeur comme se trouvant devant ou derrière les haut-parleurs.

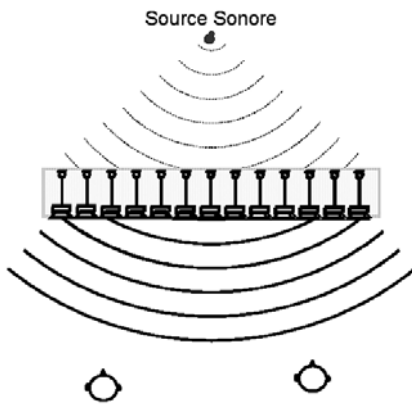


[Figure 2 : En théorie, la WFS autorise la localisation d'une source virtuelle derrière (a) ou devant (b) les haut-parleurs]

Les différences à appliquer entre les différents canaux peuvent être captées par enregistrement ou être créées artificiellement (par exemple au moyen d'un logiciel de spatialisation). Le nombre de canaux utilisé varie suivant les expérimentations menées, cela va de la quinzaine jusqu'au-delà de la centaine.

CAPTATION

Alors que la stéréophonie utilise un simple couple de microphones, l'idée du rideau sonore implique un nombre de microphones équivalant au nombre de haut-parleurs de reproduction, répartis de manière régulière autour de la scène sonore à capter et peu espacés entre eux.



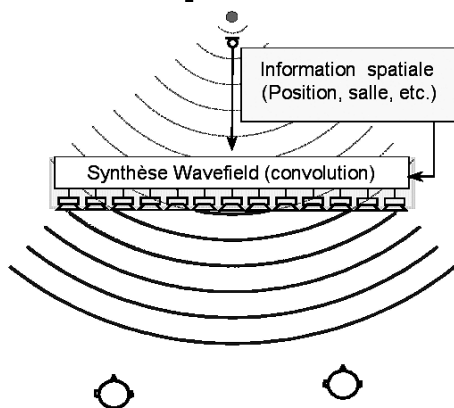
[Figure 3 : Le rideau sonore]

Cependant, s'il faut capter, transmettre ou stocker l'information correspondant à autant de canaux, la quantité de données recueillies constitue un obstacle à la mise en oeuvre de cette technique. C'est pourquoi, on se contente de capter la ou les sources, mais on consigne également pour chacune d'entre elles la position dans l'espace, de même que la «signature acoustique» du lieu de captation.

Le son est capté avec aussi peu d'influence de l'acoustique du lieu que possible. Les distances de captation sont donc plus réduites que celles qui sont utilisées en captation stéréophonique, de manière à diminuer le niveau de la réverbération et l'on se rapproche des usages de l'enregistrement en studio de la musique de variété.

La signature acoustique du lieu de prise de son est un ensemble d'échantillons de réponse impulsionnelle captés sur site avant ou après la captation principale. De tels échantillons sont déjà utilisés pour capturer la réverbération de salles et l'appliquer sur des sons via des logiciels ou des appareils. Dans le cas de la Wavefield Synthesis, les échantillons à enregistrer sont cependant plus nombreux. Il existe d'ailleurs un projet de normalisation des techniques de «capture» impul-

sionnelle visant à créer une base de données mondiale des acoustiques des salles. Cela permettrait de ne pas devoir refaire le même travail pour chaque prise de son. Bien que ce projet de base de données s'intéresse surtout aux formats en multicanal ou en Ambisonics, on imagine que les techniques développées puissent être appliquées, au moins en partie, à la WFS.



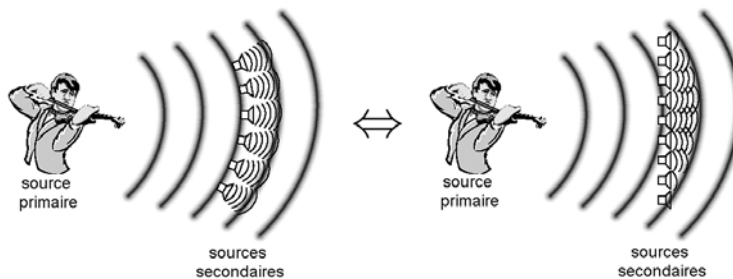
[Figure 4 : Captation séparée de l'information sonore et de l'information spatiale]

Dans le cas de sources mobiles (théâtre par exemple), un système dynamique de suivi de position est utilisable. De tels capteurs sont utilisés depuis bien longtemps dans les systèmes de réalité virtuelle (dits aussi de «réalité augmentée»).

Mais capter les sources séparément et selon une grande proximité pose certains problèmes que nous verrons plus loin et c'est pourquoi d'autres techniques de captation ont été proposées.

REPRODUCTION

La Wavefield Synthesis utilise donc un rideau de haut-parleurs placés à faible distance les uns des autres. En théorie, ces derniers devraient être placés suivant un alignement prenant la forme du front d'onde à obtenir mais en jouant avec des différences de temps et d'amplitude, on peut adopter des configurations droites ou circulaires aptes à restituer n'importe quels fronts d'ondes.



[Figure 5 : Placement des haut-parleurs suivant la forme du front d'onde original (à gauche) et placement aligné des haut-parleurs avec compensation (à droite)]

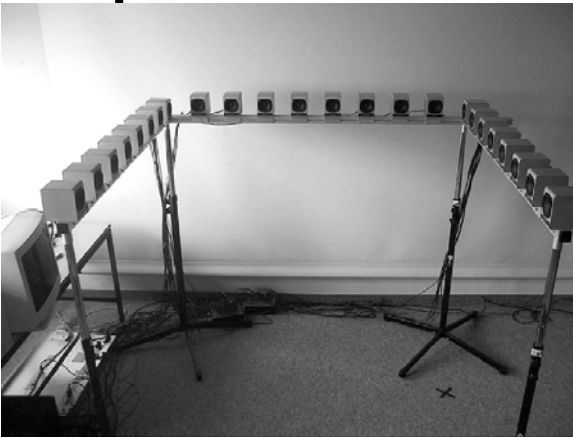
Un autre type de dispositif appelé panneau à actuateurs multiples crée la vibration d'une seule surface à partir de multiples transducteurs.



[Figure 6 : Panneau à actuateurs multiples]

Un système numérique de traitement de signal introduit avant ou pendant l'audition les différences requises entre canaux pour créer les ondes en concordance avec la localisation souhaitée. Dans le cas de la reproduction d'une captation, les trois informations (le son proprement dit, la position de la source et la signature acoustique de la salle) sont combinées afin de créer ces différences. On comprend dès lors pourquoi la technique est appelée «synthèse» puisque la scène sonore est régénérée, un peu comme on fabrique du jus de fruits à partir de concentré. Cette idée de synthèse permet (ou n'empêche pas selon le point de vue plus ou moins ouvert

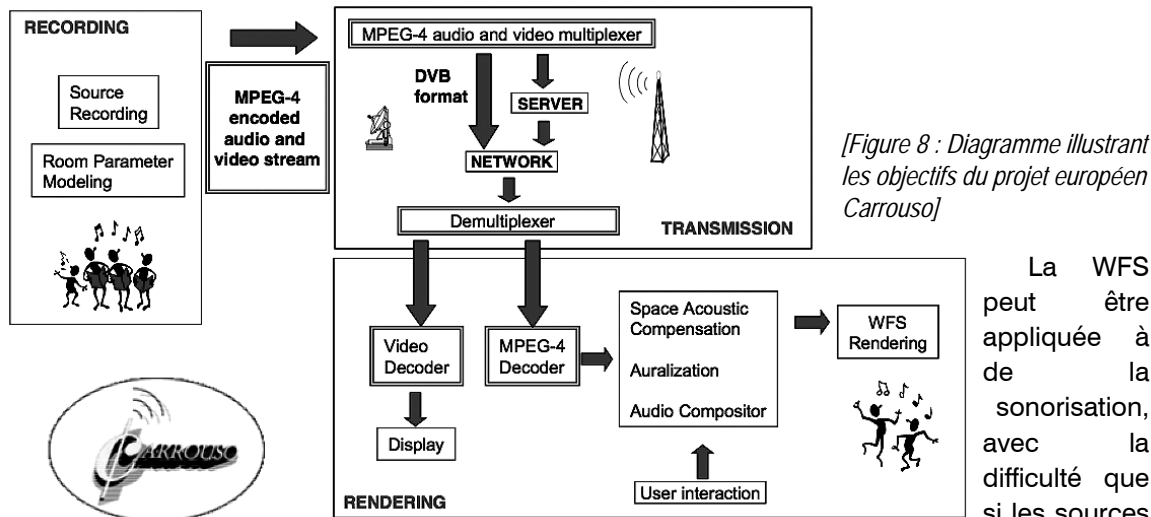
que l'on adopte) la manipulation assez large de la scène captée. Cela pose de nouveaux défis aux preneurs de son puisque le résultat est beaucoup plus ouvert, et aussi sujet à modifications, qu'il ne l'est en stéréophonie.



[Figure 7 : Disposition typique d'un dispositif d'écoute en WFS].

On peut imaginer la normalisation de l'envoi et de la combinaison des données. En ce qui concerne la transmission, les normes MPEG-4 et SAOL la permettent d'ores et déjà. C'est un des points auxquels s'est attaché le projet européen Carrouso qui a associé un ensemble d'universités (Delft, Erlangen, Thessalonique,

Ecole Polytechnique de Lausanne), de centres de recherche (l'IRCAM et l'IRT) et de sociétés (Studer, France Telecom, Thales,). L'ensemble était piloté par le Fraunhofer Institut für Digitale Medientechnologie (IDMT).



[Figure 8 : Diagramme illustrant les objectifs du projet européen Carrouso]

La WFS peut être appliquée à de la sonorisation, avec la difficulté que si les sources

sont mobiles, outre le système de suivi de position, la combinaison des différents signaux doit s'effectuer en temps réel.

COMPATIBILITE

L'élargissement du «sweet spot» ne requiert pas absolument une captation en Wavefield Synthesis. En fait, des programmes stéréo ou multicanal 5.1 peuvent eux aussi être écoutés sur un dispositif de Wavefield Synthesis en utilisant des points de panoramique virtuels (VPS – Virtual Panning Spots). Dans le cas de la stéréo, le système simule deux sources correspondant aux deux canaux de la stéréo, avec l'avantage qu'ici les deux sources sont virtuellement plus éloignées et l'on obtient donc des ondes planes avec lesquelles le «sweet spot» est donc un peu plus large, sans pour autant avoir un positionnement aussi bon qu'avec une captation adaptée à la WFS.

LIMITES

Certaines limites de la Wavefield Synthesis sont bien connues. Ainsi, en théorie, un système de captation et de reproduction sur trois dimensions pourrait être réalisé, mais cela requerrait non plus un ruban de haut-parleurs mais un mur complet, ce qui serait difficile à mettre en oeuvre tant d'un point de vue pratique que d'un point de vue économique. On se limite donc à la reproduction d'un plan horizontal.

D'autre part, alors que le principe de Huygens suppose un nombre infini de sources secondaires, le nombre de canaux est forcément limité, ce qui crée une limite dans le positionnement spatial et introduit la création d'artefacts.

L'espacement entre les haut-parleurs est également une contrainte importante qui limite la bande passante pour laquelle la localisation

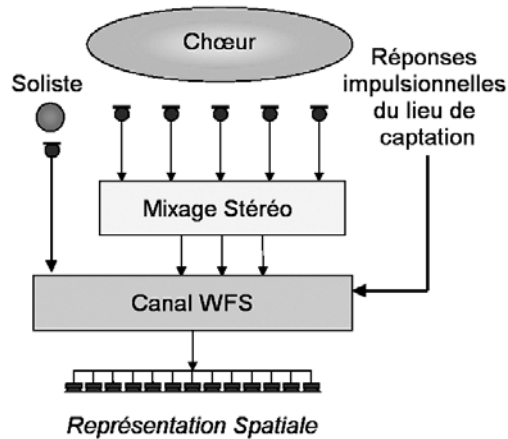
spatiale est correcte. Au-delà d'une certaine fréquence (spatial aliasing frequency), les sons ne sont plus localisables correctement. Les systèmes de Wavefield Synthesis actuels font appel à des haut-parleurs dynamiques similaires à ceux utilisés sur les systèmes stéréophoniques et les distances entre haut-parleurs varient donc de 10 à 20 cm environ (alors qu'idéalement la distance devrait être inférieure à la moitié de la plus petite longueur d'onde utilisée, soit environ 1 cm si on prend 20 kHz comme fréquence-limite). Diverses techniques sont investiguées pour résoudre ce problème, par exemple l'augmentation de la directivité des haut-parleurs, ou l'utilisation de haut-parleurs différents suivant le registre fréquentiel à reproduire (technique déjà employée dans l'immense majorité des systèmes d'écoute stéréophoniques).

La taille finie du champ des haut-parleurs crée des problèmes de diffraction aux extrémités du dispositif. Il est possible d'appliquer une pondération qui réduit le niveau des haut-parleurs au fur et à mesure qu'on s'approche des extrémités, mais cela restreint la zone où les auditeurs peuvent se placer. Une autre solution est d'élargir le champ des haut-parleurs sur les côtés, voire de créer un champ circulaire qui ne présente alors plus de bord.

Un autre problème est celui de l'interaction du dispositif d'écoute avec l'acoustique du lieu de diffusion. En effet, à moins de se trouver en chambre anéchoïde, l'acoustique risque d'interférer sur la localisation spatiale via la réverbération. Puisqu'un système WFS inclut, par nature, un système de traitement numérique du signal numérique (DSP), il est cependant envisageable d'intégrer dans ce système un dispositif de compensation acoustique fondé sur des mesures effectuées dans le lieu de diffusion lui-même.

Au-delà des limites techniques, il reste que la Wavefield Synthesis pose quelques questions pratiques. D'abord, si la captation de multiples sources et leur transmission est possible, elle pose tout de même un problème de quantité de données à manipuler, quantité qui croît à peu près linéairement avec le nombre de sources à reproduire. D'autre part, la captation à proximité des sources entraîne la création de sonorités pas toujours très naturelles. En effet, nous avons l'habitude de percevoir pas mal d'instruments de musique dans un champ diffus (réverbération) et cet aspect est lié intrinsèquement à notre expérience de la musique. Et la simple combinaison d'une prise de son de proximité avec la «signature» acoustique et une information de position semble ne pas donner des résultats probants. D'autre part, si on imagine que la Wavefield Synthesis devienne le futur de l'audio, un problème en matière de compatibilité avec les systèmes actuels (stéréo et multicanal) se pose. En effet, si une captation doit être réalisée, on imagine mal les

ingénieurs du son cumuler le placement d'un système WFS et d'un système stéréo (ou multicanal), sans compter la nécessité de devoir réaliser deux mixages distincts. C'est pourquoi, une pratique adaptée de la WFS tenant compte des pratiques actuelles et de l'expérience accumulée avec les techniques antérieures est à la fois souhaitable et possible. Ainsi, on imagine mal capter une chorale de 50 membres avec un micro sur chaque chanteur. On peut créer un mixage multicanal (voire stéréo) en réduisant le signal à transmettre sur cinq canaux par exemple. Ces cinq canaux constituent alors cinq sources virtuelles dans le dispositif d'écoute. On peut bien sûr augmenter le nombre de canaux suivant la source à capter, le genre, etc. On peut séparer les musiciens solistes des autres musiciens. On définit ainsi un compromis entre les avantages de la WFS et la praticabilité de la prise de son.



[Figure 9 : Diagramme illustrant une mise en œuvre plus réaliste de la captation]

Il reste à connaître les liens entre les multiples paramètres du système et la perception du son suivant ses diverses qualités et particulièrement de ses qualités spatiales.

«BIENTOT DANS UNE SALLE PRES DE CHEZ VOUS»?

Au-delà des fondements théoriques et des premières expérimentations, on se trouve donc dans une phase de développement et d'amélioration. Cependant, si on s'en réfère à la stéréophonie, celle-ci était déjà largement diffusée avant d'avoir été totalement comprise et ce n'est que petit à petit que tout son potentiel a pu être découvert et exploité. Si on excepte donc les défauts les plus limitatifs, il ne reste plus beaucoup de problèmes techniques entravant la diffusion à grande échelle de la Wavefield Synthesis.

L'institut Fraunhofer IDMT a d'ailleurs étudié la mise en œuvre pratique de la Wavefield Synthesis et a installé en 2003 un système d'écoute à 192 voies dans une salle de cinéma à Ilmenau (Allemagne). Une nouvelle compagnie, Iosono, a été créée afin de diffuser le système commercialement. Au mois de juillet, Iosono a réalisé une démonstration

devant quelques importantes firmes hollywoodiennes. Une première utilisation commerciale se ferait en 2005 dans un lieu dédié non plus au cinéma mais au spectacle vivant. Enfin, des négociations avec une firme active dans l'électronique grand public seraient en cours afin de proposer des systèmes utilisables chez les particuliers.



[Figure 10 : Gros plan sur les haut-parleurs du dispositif losono installé à Ilmenau]

Il faut noter enfin qu'un concert aura lieu précisément dans la ville de Ilmenau le 17 septembre, ceci avec l'Intuitive Music Ensemble de Weimar. Une bonne occasion d'écouter ce qu'un dispositif WFS peut apporter à la musique !

Roald Baudoux

(Audio Pro n° 80 / septembre 2004)

BIBLIOGRAPHIE

Wellenfeldsynthese – Ablösung der Stereophonie? - Helmut Wittek, Günther Theile, Markus Reisinger, IRT-Kolloquium 9. Dezember 2002,
http://www.irt.de/IRT/veranstaltungen/irt_kolloquium_09_12.pdf

Potential wavefield synthesis applications in the multichannel stereophonic world - Günther Theile, Helmut Wittek, Markus Reisinger
http://www.irt.de/wittek/hauptmikrofon/Banff_WFS_Theile_et_al.pdf

Acoustic rendering with wave field synthesis - Marinus M. Boone, Proc. Acoustic Rendering for Virtual Environments,
<http://www.bell-labs.com/topic/conferences/campfire/abstracts/boone.pdf>

Spatial perception in WFS rendered sound fields - Günther Theile,
<http://www.irt.de/wittek/hauptmikrofon/theile/DAGA04-Theile.pdf>

Sound Reproduction By Wave Field Synthesis - rapport de travail de Alfredo Fernández Franco, Sebastian Merchel, Lise Pesqueux, Mathieu Rouaud, Martin Ordell Sørensen, Faculty of Engineering and Science, Aalborg University,
<http://kom.aau.dk/~merchel/Sound%20Reproduction%20by%20Wave%20Field%20Synthesis.pdf>

Wave field synthesis and analysis using array technology - Diemer de Vries and Marinus M.Boone, Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,

http://www.soundcontrol.tudelft.nl/Research_projects/WFS_IEEE_99.pdf

High-Quality Acoustic Rendering with Wave Field Synthesis - S.Spors, H.Teutsch and R.Rabenstein, University of Erlangen-Nuremberg, Telecommunications Laboratory,

http://www.Int.de/LMS/publications/web/Int2002_21.pdf

LIENS INTERNET

Université Technique de Delft :

<http://www.soundcontrol.tudelft.nl/index.htm>

Fraunhofer IDMT : <http://www.idmt.fraunhofer.de/eng/english.htm>

Page dédiée au projet Carrouso à l'IRCAM :

<http://www.ircam.fr/produits/technologies/CARROUSO.html>

Institut für Rundfunktechnik : <http://www.irt.de>

Iosono : <http://www.iosono-sound.com>