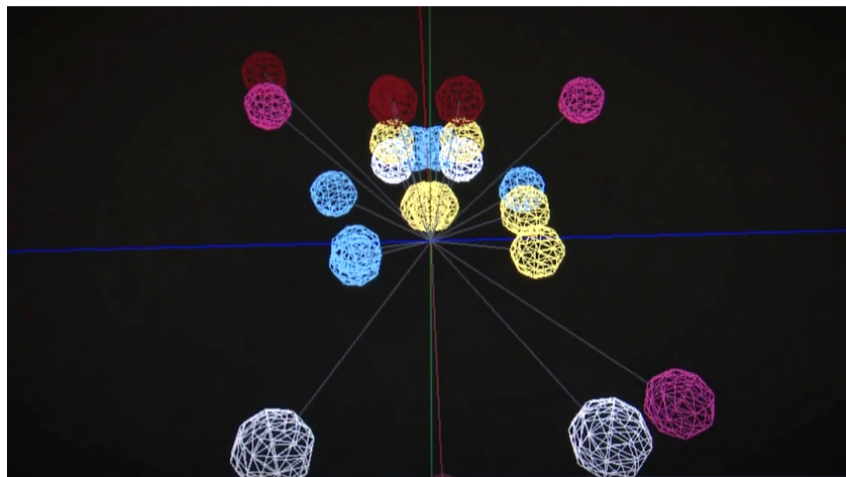


---

## Audio orienté objets et réseaux IP : le grand virage s'amorce

Publié le 28 septembre 2017 par [Olivier Jouinot](#)

 [Twitter](#)



Les 17 et 18 mai 2017 à l'[EBU](#) de Genève, France Télévisions organisait un séminaire entièrement consacré à l'Audio Orienté Objets, l'occasion pour les principaux acteurs de présenter leurs solutions, d'échanger autour des bonnes pratiques... et préparer l'avenir. S'étaient réunis les principaux broadcasters européens (BBC R&D, RAI, NRK, ORF TV, Radio France, Radio Télévision Suisse, Swedish

Radio), les leaders industriels que sont Dolby Laboratories, Fraunhofer IIS, DTS Inc., Qualcomm, Merging Technologies, New Audio Technology, ainsi que des laboratoires spécialisés (IRT, IRCAM, b<>com...). Le salon IBC fermant tout juste ses portes et avant de distiller les dernières tendances, retour sur ce moment et les différentes présentations, l'occasion de vérifier si les promesses ont été tenues quelques mois plus tard sur les stands des principaux leaders.

*par Vincent Dabouineau – France Télévisions – innovations&développements*

## L'Audio Orienté Objets

Dans la chaîne audiovisuelle, l'audio orienté objets (abrégé OBA, pour **Object-Based Audio**) dissocie la partie production de contenu de son mode de restitution chez le consommateur final. De fait, ce nouveau paradigme "speaker agnostic" met un terme à la course aux armements vis-à-vis du nombre de haut-parleurs nécessaires à la restitution du signal audio (5.1, 7.1, 9.1, 11.1 et 22.2) ou de leur emplacement. L'OBA est produit une seule fois puis distribué sur n'importe quel système de rendu. Qu'il soit mono, stéréo (au casque ou non), sur système multicanal "traditionnel" ou barres de son, le rendu devient adaptatif grâce à la transmission d'informations de positionnement – azimut, distance, élévation- des objets contenus dans le flux source. Sur des installations hétérogènes, l'OBA tend à respecter l'intention artistique originale, tout en offrant des capacités de personnalisation de l'écoute au niveau du récepteur, ainsi qu'une rationalisation importante des débits côté diffusion.

Cette dissociation de bout en bout amène énormément d'avantages, récapitulons :

- le contenu n'est produit qu'une seule fois pour la distribution et la consommation multiplate-forme ;
- en bout de chaîne, l'accès aux ressources médias est facilité, l'écoute est personnalisable. Sont présentés à l'utilisateur du contenu et des métadonnées qui lui permettent de prendre la main sur le mixage, jusqu'à un certain niveau toutefois ;
- des environnements sonores audio immersifs sont créés grâce à des métadonnées de position, celles-ci guident la reproduction des éléments de la scène sonore. Les utilisateurs de casques audio sont particulièrement privilégiés, car les métadonnées de position facilitent la création d'une stéréophonie binaurale immersive à la volée.

Décrire le contenu audio pour l'adapter à sa consommation, voici donc l'enjeu des technologies Audio Orienté Objets présentées au cours du séminaire. La pierre angulaire pour réaliser cette description est un schéma de métadonnées ouvert, normalisé ITU-R BS.2076-0 [41] : l'[Audio Definition Model \(ADM\)](#)

## L'Audio Definition Model

Mis au point au sein de l'Union Européenne de Radio-télévisions (UER – EBU) à partir de 2011, sur la base de travaux convergents entre France Télévisions et la BBC R&D, l'ADM est un standard ouvert qui vise à assurer la compatibilité entre les formats audio *channel-based*, *object-based* et *scene-based* en utilisant une représentation XML. L'ADM offre un moyen pour décrire les métadonnées audio de sorte que chaque piste individuelle d'un fichier ou d'un flux soit correctement rendue, traitée ou distribuée. Le modèle est divisé en deux parties, l'une relative au contenu et une autre au format. La première décrit le contenu audio, comme la langue des dialogues et le loudness. La partie format contient des informations techniques nécessaires pour que l'audio soit décodé ou rendu correctement, comme les coordonnées tridimensionnelles d'un objet sonore (azimut, distance, élévation) ou l'ordre d'un composant HOA (*High Order Ambisonics*).

Une option philosophique va également distinguer l'ADM de ses formats concurrents : l'abandon du concept de format multicanal. Là où les formats *scene-based* et *object-based* sont vus comme une extension du format multicanal historique (*channel-based*), l'ADM considère simplement qu'un canal équivaut à un objet fixe dans l'espace et le temps.

Le standard propose une série de balises ADM telles que *audioTrackFormat* (décrivant le format dans lequel se trouvent les données), *audioTrackUID* (liant de manière unique une piste ou un élément avec un enregistrement d'une scène audio), *audioPackFormat* (un regroupement de canaux audio), etc. Sont également documentées les bonnes pratiques concernant l'utilisation des ID, des systèmes de coordonnées et des descriptions de paramètres orientés objets. Enfin, une série d'exemples d'utilisation de l'ADM pour les 3 modes *channel*, *object* et *scene-based*, conclut la recommandation, code d'exemple XML et diagrammes UML à l'appui. Les métadonnées spécifiques à la réalité virtuelle telles que l'indication de présence d'un contenu diégétique ou non-diégétique ne faisaient pas partie de la recommandation au moment de l'écriture.

Dans le cas où les contenus audio sont stockés dans un fichier au format **BWF** (*Broadcast Wave Format*), la description XML peut être ajoutée à ce fichier via l'usage de l'*axml chunk* prévu dans le format et, dans la plupart des cas, directement intégrée dans le *chna chunk* prévu pour l'embarquement de codes ADM simplifiés.

À l'heure actuelle, l'ADM est un format plus orienté production que scénarios OTT, mais ceci pourrait évoluer dans un futur proche. Pour le streaming, un format de conteneur qui présente à la fois les paquets audio et vidéo (compressé et non compressé) ainsi que les métadonnées ADM serait souhaitable. A noter la soumission, en fin d'année 2016, d'une version dynamique des données ADM, autorisant l'insertion de ces métadonnées dans un flux. Cette solution doit permettre l'adoption de l'ADM pour la production en direct.

---

## Déroulé du séminaire

L'introduction de Simon TUFF (BBC) fut remarquée, avec un retour en image sur le parcours des pionniers qui ont produit des avancées technologiques concrètes dans le monde de l'audio. Depuis la fin du XIXème siècle, ces héros ont aidé à l'avènement des possibilités **NGA** (**N**ext **G**eneration **A**udio) que nous connaissons aujourd'hui. La normalisation a suivi :

---

**Ludovic NOBLET (institut b<>com)** ouvre officiellement l'événement, une présentation très axée sur les services innovants au profit de l'expérience utilisateur. Après quelques pastilles sur des sujets brûlants comme le renforcement des dialogues, b<>com estime que le futur des médias se confond avec les usages réalité virtuelle et augmentée :

Distillant quelques pistes vis-à-vis des nouvelles expériences possibles, l'OBA en Radio par exemple, Ludovic revient *in fine* sur la démonstration de Réalité Augmentée de b<>com proposée sur le NAB au printemps 2017. Une mise à jour étiquetée "*cognitive loops*" augmente encore le niveau d'immersion, elle est exposée sur l'espace démo' du séminaire.

---

**Florian CAMERER (ORF)** fait le point sur le vocabulaire utilisé dans le monde de l'Audio Orienté Objets. Retour sur l'état de l'art depuis l'époque du multicanal, rappelant que les objets sont présents depuis l'étape de production ; un discours volontairement très pédagogique sur les tenants et aboutissants de ce qu'on appelle désormais OBA. Il détaille les différents scénarios possibles jusqu'à l'étape ultime : convaincre les entités décisionnaires chez tous les diffuseurs, pour la mise en œuvre d'une chaîne de bout en bout qui tienne compte d'une modification des habitudes et des savoir-faire dès l'étape de production.

---

**Matthieu PARMENTIER** (France Télévisions – innovations&développements) définit le rôle de l'audio immersif sur la perception humaine et propose une nouvelle méthode d'évaluation pour des tests perceptifs OBA. C'est l'un des sujets phare de l'été qui agite le microcosme audio 3D, à l'heure où différentes approches s'affrontent :

---

**Michael WEITNAUER (IRT)** se focalise sur les formats et les standards dédiés aux workflows OBA, regrettant que les manques actuels soient en partie responsables d'un passage en production très timide... L'interopérabilité est au cœur de la problématique d'interconnexion des plates-formes, des différents codecs utilisés et des équipements existants. La cohabitation entre les formats ADM / BWF, AC-4 ou MPEG-H pose question. Pour rendre l'écosystème vertueux, les moteurs de rendu commerciaux (Dolby, Fraunhofer, DTS) devraient supporter nativement l'ADM en entrée/sortie. Selon l'IRT, il reste à résoudre une équation qui mélange le loudness orienté objets, la réverbération 3D et sa description dans l'ADM (adaptées aux objets audio), et la normalisation de profils ou à minima de bonnes pratiques pour un format de transport de l'ADM (rendre le schéma XML *streamable*) :

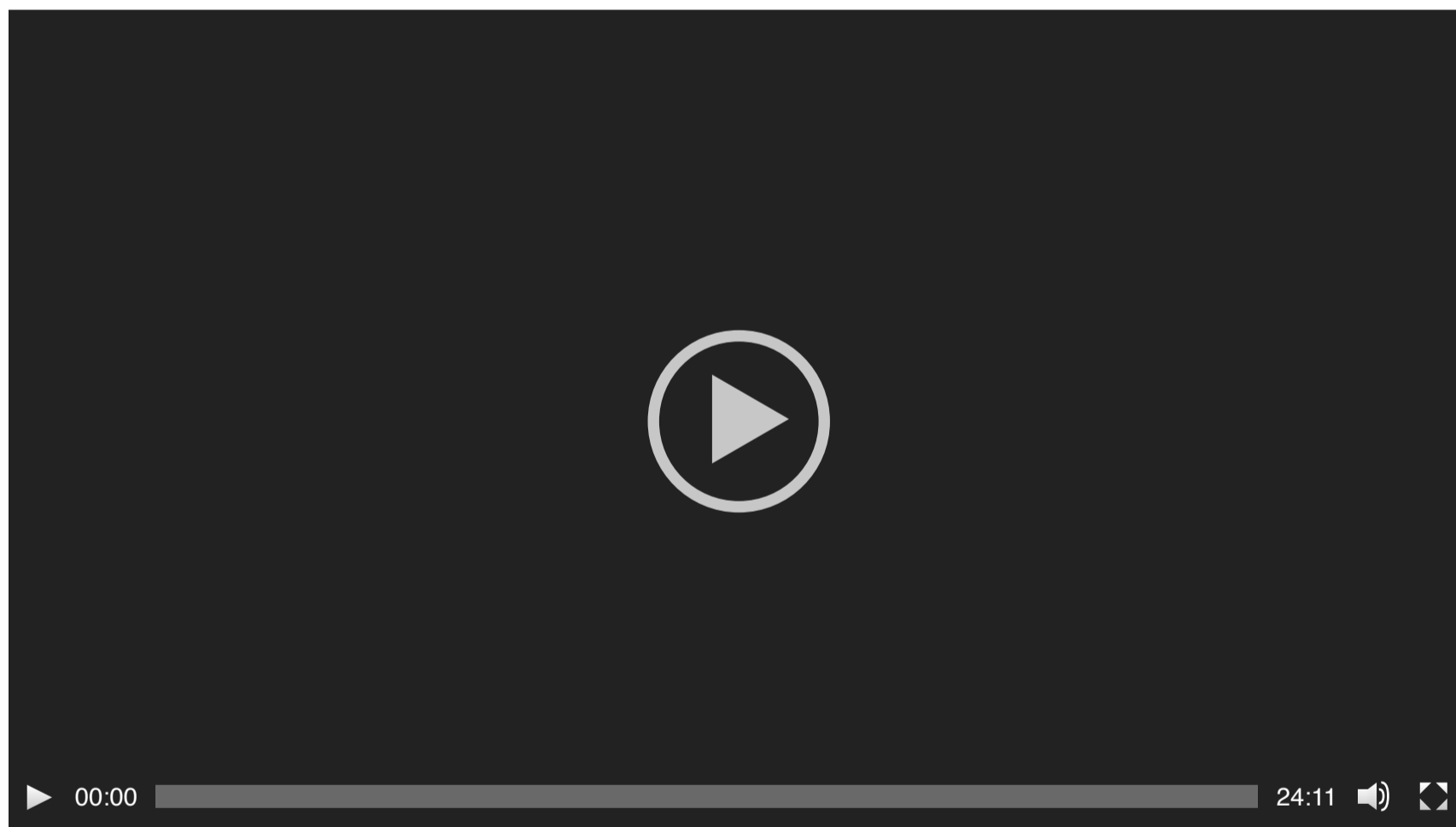
Si la plupart des constructeurs ont déjà rejoint l'une des alliances promoteurs d'un format de transport IP, tous attendent la structuration complète des infrastructures audiovisuelles autour de l'IP avec l'arrivée du standard SMPTE 2110 qui prévoit la cohabitation de flux audio, vidéo et métadonnées dans un même réseau, asservis au même protocole de synchronisation.

La publication de ce standard est attendue pour la fin d'année 2017. SMPTE 2110 reprend notamment le standard AES-67 sur le transport d'audio sur IP, lui-même établi en combinant les meilleures caractéristiques des grandes alliances privées (Ravenna, Livewire, Dante, etc.).

Dans l'attente de cette fusion des innovations : métadonnées dynamiques, audio orienté objets et protocoles de transport sur IP, les fabricants de console temporisent à raison. Pour l'heure, les premières solutions de production en direct sont donc spécifiques à la technologie d'encodage utilisée. Dolby, Fraunhofer (en partenariat avec Jünger et Linear Acoustics) et DTS proposent chacun leur unité de création des métadonnées et de contrôle du monitoring afin d'offrir aux ingénieurs du son la possibilité de simuler le rendu téléspectateur. Avec l'arrivée des formats audio 3D et la poussée importante du nombre d'auditeurs au casque, les différentes versions à contrôler connaissent une inflation galopante. Si l'on y ajoute les déclinaisons dynamiques liées à l'activation des **DRC** (**D**ynamic **R**ange **C**ontrol) automatisés.

---

**La question du traitement de la réverbération 3D** dans un écosystème OBA est très surveillée par la **Conservatoire National de Paris** côté production, et par l'**IRCAM** avec son approche plus laboratoire. Leur présentation successive finit de convaincre que l'immersion se joue en grande partie sur les capacités du système de rendu à reproduire une réverbération qui corresponde ou simule un environnement de captation réel, la spatialisation des objets s'en trouve fortement accentuée :



---

## Format scene-based

La sémantique liée au format “*scene-based*” inclut les concepts de captation, représentation et restitution de champs sonores. Dans cette catégorie on retrouve les termes “*Sound Field*”, scène sonore, “*Ambisonics*” et “*High Order Ambisonics – HOA*”.

À l'origine de ce format on trouve les travaux de Michael GERZON sur la décomposition en ondes harmoniques sphériques d'un champ sonore en 3 dimensions. En fonction de la précision de cette décomposition, le champ sonore sera représenté par une ou plusieurs composantes, l'ensemble constituant une description isotropique (équivalente en précision, dans toutes les directions) de la scène vue d'un point unique d'écoute. La monophonie

omnidirectionnelle correspond à la représentation de base, à laquelle s'ajoute des composantes de plus en plus nombreuses suivant l'ordre de définition choisi.

En 3 dimensions, l'ordre 1 correspond au format Ambisonics (ou format B). Il est constitué d'une composante monophonique omnidirectionnelle (W) à laquelle s'ajoutent 3 composantes bidirectionnelles représentant la répartition avant/arrière (X), gauche/droite (Y) et haut/bas (Z) du champ sonore. Au delà de l'ordre 1, le nombre de composantes des formats HOA 3D obéissent à la règle  $(\text{ordre} + 1)^2$  soit 9 composantes pour l'ordre 2, 16 pour l'ordre 3, 25 pour l'ordre 4, etc. On trouve également des formats HOA 2D qui nécessitent naturellement moins de composantes.

En matière de captation, ces travaux ont abouti à la création de microphones Ambisonics et HOA, tels que ceux des marques Soundfield (propriété du groupe TSL), DPA et **mh acoustics**.

Microphone HOA 3D em32 Eigenmike © mh acoustics

En matière de restitution, le maillage de haut-parleurs nécessaire dépend également de l'ordre utilisé. Ils sont généralement répartis sur une demi-sphère placée tout autour et au dessus de l'auditeur en 3D, ou bien un cercle parfait en 2D.

Système de restitution HOA en 2 dimensions

Ces formats “*scene-based*” 2D et 3D s'avèrent particulièrement utiles pour représenter un champ sonore sans discrimination de précision. Contrairement aux formats multicanaux standardisés qui privilégient un nombre accru de haut-parleurs derrière et autour de l'écran, les formats “*scene-based*” sont notamment utilisés dans les applications nécessitant la rotation de la scène sonore : réalité virtuelle à 360° et head-tracking dans le cadre d'une écoute binaurale.

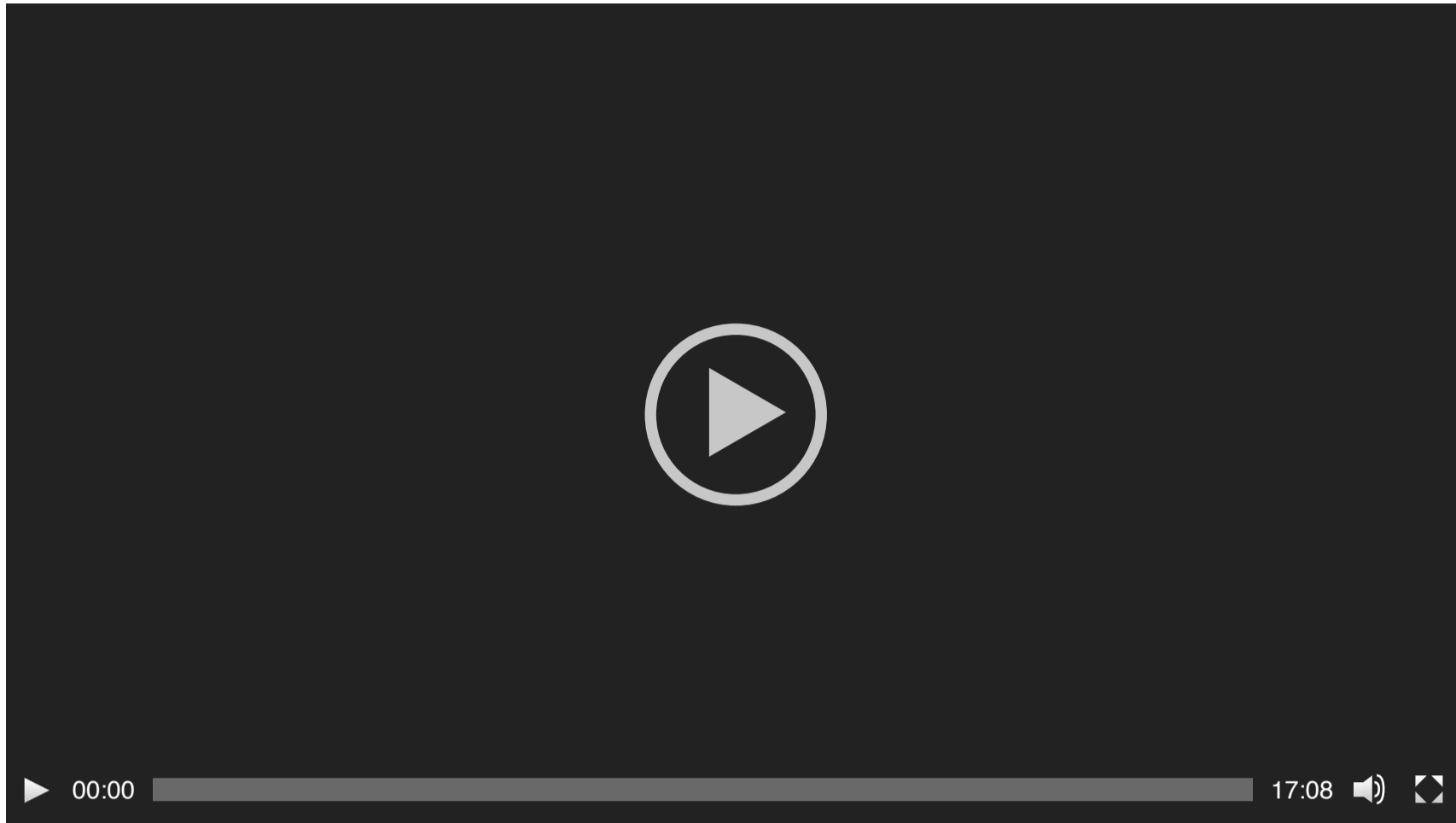
En restitution, les formats HOA s'associent volontiers aux technologies de Wave Field Synthesis (WFS) pour former une zone d'écoute optimale la plus large possible. A Paris, l'espace de projection de l'IRCAM combine une couronne WFS de plusieurs centaines de haut-parleurs pour la restitution de n'importe quelle format horizontal – multicanal ou HOA 2D – avec une demi-sphère parsemée d'une autre centaine de haut-parleurs disposés selon la répartition HOA 3D pour restituer la troisième dimension.

Schéma d'implantation des haut-parleurs HOA (cubes jaunes) et WFS (barres jaunes) à l'espace de projection de l'Institut de Recherche Coordonnée Acoustique Musique à Paris © IRCAM

---

**Ferdinando OLIVIERI pitch pour Qualcomm sur l'association d'audio scene-based et object-based.** La philosophie de départ et les fonctionnalités offertes par cette approche -choix d'un **PoV** (Point of View), déplacement de la scène audio, zoom dans la scène- s'intègrent parfaitement dans des scénarios VR. Qualcomm évoque néanmoins les problématiques de transport d'une telle solution dans les infrastructures SDI,

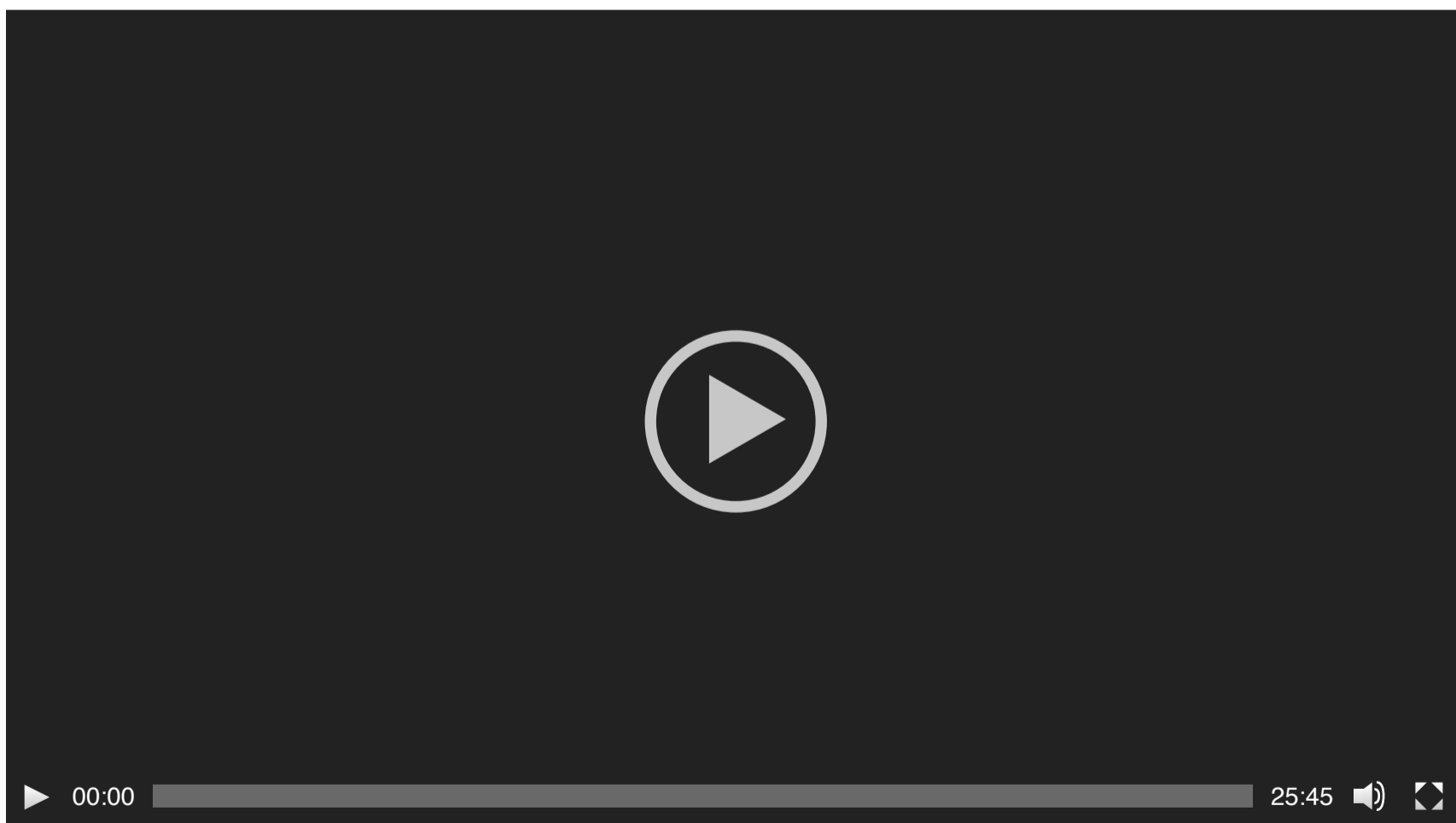
en fonction du nombre d'objets notamment, et de la bande passante disponible. Le meilleur des deux mondes - SBA pour la composante immersive et OBA pour une sélection d'objets en particulier- séduit largement l'assistance :



Cependant, réaliser un master en HOA n'est pas encore trivial et Qualcomm propose naturellement des plugins pour créer et rendre ce type de contenus sur des stations de travail. L'accent est mis sur la capacité d'approvisionnement de tous les dispositifs, depuis le home-cinéma 5.1, la barre de son 3D, le rendu binaural au casque et jusqu'au dispositif de réalité virtuelle à partir du même master HOA. Pour l'instant les plugins de rendu binaural ne sont malheureusement pas "personnalisables".

---

**Du côté des poids lourds de l'industrie OBA**, le séminaire fut une belle opportunité pour décrire leurs chaînes de bout en bout. **Tim ADDY présente les outils de production objets de Dolby, Atmos et AC-4 :**



Le codec AC-4 est la nouvelle génération de compression numérique pensée par Dolby pour réduire notablement le débit à qualité égale, à un niveau comparable aux débits évoqués ci-dessus. Ce codec est standardisé par l'ETSI, l'ATSC 3.0 et est en cours d'adoption par DVB.

Associé à sa technologie Atmos (audio orienté objets et contenus immersifs), l'AC-4 / Atmos propose pratiquement la même offre technologique que MPEG-H 3D, moyennant un vocabulaire différent

(“*presentation*” au lieu de “*preset*”, “*bed*” au lieu de “*multicanal*” etc.), et à l’exception du rendu binaural, qui diffère.

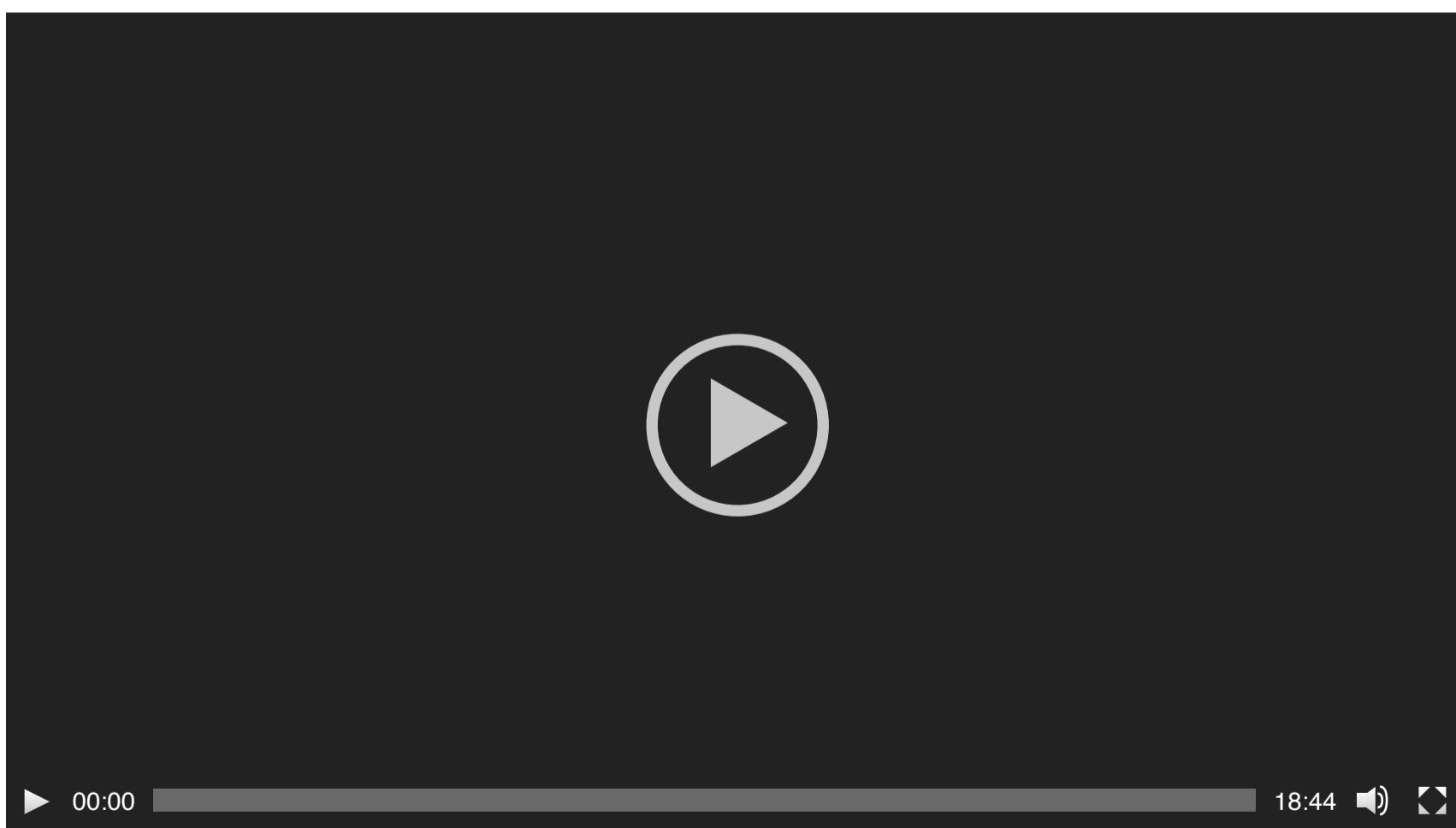
Pour l’instant Dolby ne croit pas à la personnalisation de l’écoute binaurale et ne propose pour l’instant qu’un seul type de process binaural censé convenir à tous. En revanche, la solution Dolby Headphones pensée pour la réalité virtuelle permet d’associer pour chaque objet un rendu au choix parmi 3 types de diffusion spatialisée : pas de réverbération, premières réflexions ou effet de salle. Cette solution favorise la sensation d’extériorisation des sources sonores et vient compenser la non-individualisation du rendu binaural.

Du côté des équipements de production, le processeur de monitoring DP590 et les encodeur DP591/décodeur DP592 abandonnent la connectique AES de leurs prédécesseurs au profit du MADI (DP590) et du 3G-SDI (DP591/592) pour mieux répondre à l’inflation du nombre de canaux à traiter. Dolby poursuit les consultations auprès de ses primo-utilisateurs pour faire évoluer l’interface de ses outils de création des mixages Atmos. Les contraintes s’avèrent nombreuses devant le nombre de possibilités offertes. On retiendra les 2 grands principes que Dolby met en œuvre sur la totalité de sa chaîne de production AC-4 / Atmos :

- Les présentations : une présentation correspond à un “preset” activé par l’utilisateur final, tel que “version française” ou “version originale”. Grâce aux objets on peut créer plusieurs présentations alternatives telles qu’une version dont le niveau des dialogues est renforcé, ou encore une retransmission sportive sans commentaire journalistique. La multiplication des présentations d’un même programme n’a théoriquement pas d’impact sur le débit du flux AC-4 / Atmos, tant qu’elles partagent les mêmes sources.
- Les exceptions de rendu : l’ingénieur du son peut choisir d’exclure certaines sources dans le cas d’une écoute utilisateur particulière. Par exemple, les sources sonores en élévation utilisées pour le rendu de la 3ème dimension à destination des barres de son 3D Atmos ou de l’écoute binaurale au casque peuvent être supprimées d’un rendu 5.1 ou stéréo.

---

**Adrian MURTAZA** introduit pour **Fraunhofer IIS** la partie authoring et utilisation des outils de rendu **MPEG-H Audio** :



Dopé par le lancement des chaînes Ultra HD coréennes qui ont choisi cette technologie d’encodage audio orienté objets, le consortium MPEG-H 3D offre plusieurs solutions pour le direct et la post-production. À l’instar du Dolby AC-4 / Atmos, MPEG-H 3D distingue également les notions de programme (= présentations) et les exceptions de rendu.

Après le prototype de processeur Jünger à l'interface beaucoup trop chargée, destiné à simuler le rendu téléspectateur depuis la régie de production, c'est l'américain Linear Acoustic qui lance son AMS (Authoring and Monitoring System), un outil à l'ergonomie beaucoup plus aboutie qui n'est pas sans rappeler le DP590 de Dolby. Au menu, un processeur capable de créer les métadonnées MPEG-H 3D et les associer à 15 canaux audio. L'occupation de ces 15 canaux permet d'envisager une ambiance 9.1, des objets "dialogues stéréo" dans 2 langues différentes ainsi qu'un objet "commentaire d'audiodescription".

Contrairement à Dolby, MPEG-H ne commercialise pas directement d'encodeur hardware mais noue des accords d'intégration avec les fabricants d'encodeurs vidéo. Autre partenariat technologique, l'allemand New Audio Technologies a fait évoluer son plugin de panning 3D "Spatial Audio Designer" compatible Pro Tools, Pyramix, Nuendo... Désormais ce plugin permet de signaler l'usage d'objets et génère le flux complet de métadonnées MPEG-H destiné à l'encodeur.

Les performances indicatives pour le transport d'audio 3D compressé surclasse les technologies précédentes pour une qualité réputée "bonne" :

- 32 kbps pour une voie monophonique ;
- 64 kbps pour une paire stéréophonique ;
- 128 kbps pour un mixage 5.1 ;
- 160 kbps pour un mixage 5.1.4 ;

MPEG-H 3D fait partie des codecs audio retenues par la future solution de diffusion hybride américaine : l'ATSC 3.0, et en cours d'adoption par DVB dans le cadre de la création de profils de diffusion Ultra HD phase 2 pour l'Europe. La solution américaine ATSC 3.0, également adoptée en 2017 par la Corée du Sud pour le lancement de plusieurs chaînes Ultra HD diffusées en numérique terrestre sur les principales villes du pays, a permis à Fraunhofer d'assurer les premiers déploiements mondiaux d'audio orienté objets.

---

**DTS:X et DTS:UHD** : DTS:X a été créé pour le marché du Blu-Ray disc, en compressant (réduction de débit) les éléments sonores issus des mixages cinéma orientés objets utilisant le format de métadonnées maison MDA pour Multi-Dimensional Audio.

DTS:UHD répond aux besoins en diffusion grand public (TV, streaming) avec des performances annoncées équivalentes aux faibles débits proposés par MPEG-H 3D et AC-4 / Atmos. Ses standardisations ETSI et DVB sont en cours, le détail des fonctionnalités disponibles fera l'objet de démonstrations lors des salons professionnels du second semestre 2017.

Fadi MALAK intervient pour le groupe DTS avec une présentation concentrée sur les usages grand public et très axée sur l'accessibilité.

Personnalisation de l'écoute avec le système DTS OBA en développement © DTS, Inc.

En partenariat avec l'université de Stanford (Manchester), le

concept de Clean Audio et le projet éponyme sont détaillés, la présentation fait la part belle à la personnalisation de l'écoute :





▶ 00:00 19:00 🔊 🗨

## Focus sur les barres de son

Pour pénétrer dans les foyers, les systèmes multicanaux 3D s'appuient sur plusieurs innovations combinant le **Wave Field Synthesis** et le **Beam Forming** au sein de produits grand public de type barres de son ou réflecteurs. La qualité du rendu de ces systèmes dépend étroitement des conditions acoustiques de la pièce d'écoute, et notamment de la symétrie des murs et des propriétés de leurs revêtements.

## Technologie Wave Field Synthesis

La synthèse d'onde sonore – en Français – est un procédé de restitution sonore industrialisé dans les années 2000 pour permettre notamment l'élargissement de la zone d'écoute. En effet, plus le nombre de canaux augmente, plus la zone d'écoute idéale se restreint. L'avènement des formats multicanaux a rapidement posé problème dans les grandes salles de cinéma, où le public placé sur les bords de la salle ne peut pas bénéficier d'un rendu correct de la bande-son du film.

Le Wave Field Synthesis utilise de nombreux haut-parleurs placés sur un même plan pour synthétiser l'onde sonore émise par une source virtuelle, notamment en jouant sur des décalages temporels en adressant successivement les haut-parleurs de plus en plus excentrés.

fig.1 – principe du Wave Field Synthesis © Iosono

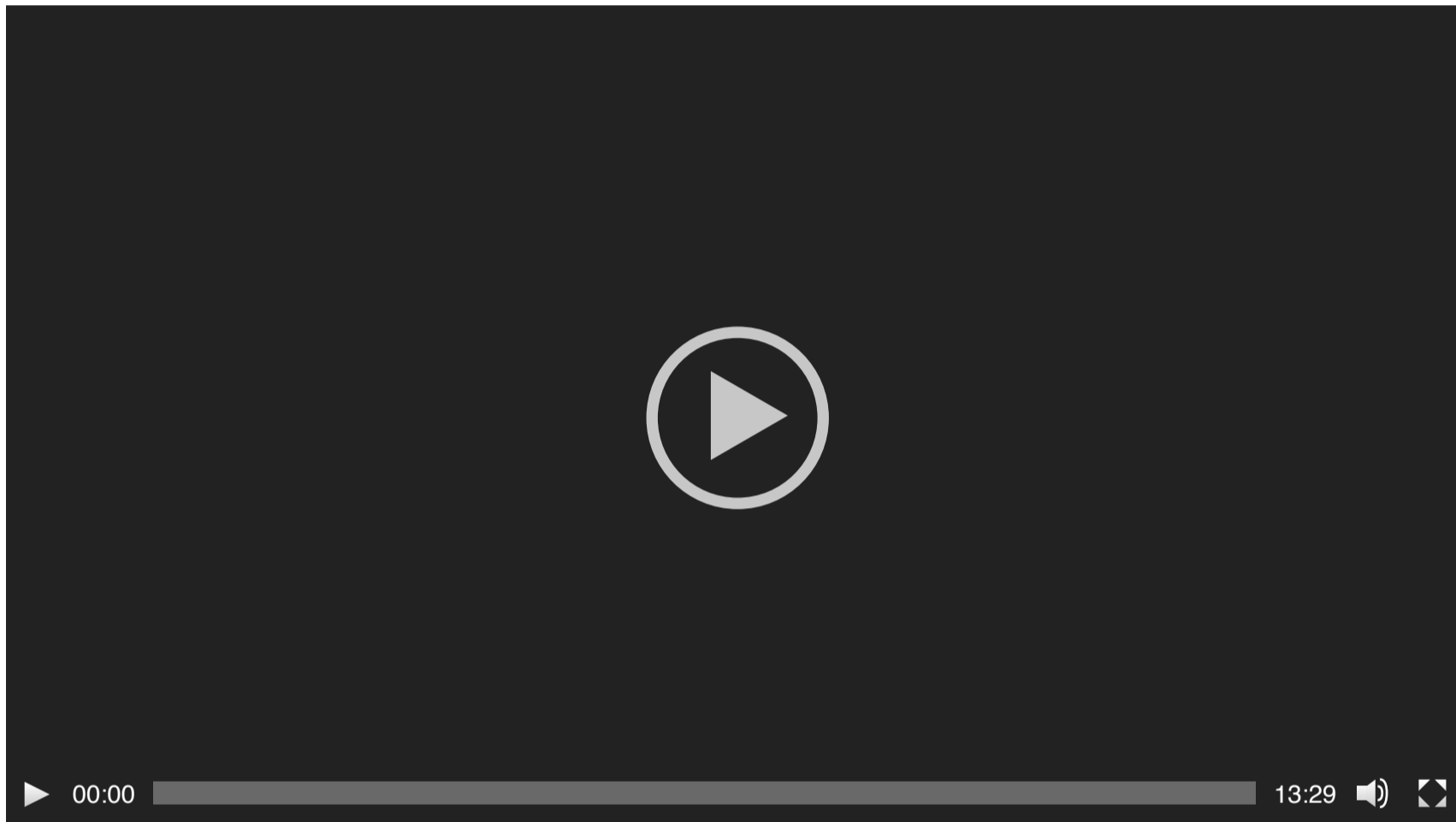
Ce procédé permet également de fabriquer des ondes planes, lesquelles correspondent à une source émise à l'infini. Pour ce faire les multiples haut-parleurs sont excités en même temps. Des retards sont néanmoins utilisés d'un haut-parleur à l'autre dans un même plan, pour donner un angle à ces ondes planes. En reproduisant autant d'ondes planes que de canaux d'un format stéréophonique ou 2D, la zone d'écoute idéale se trouve étendue à presque l'ensemble de la surface comprise entre les haut-parleurs.

fig.2 – système Wave Field Synthesis appliqué à la reproduction d'un format en 2 dimensions, et ci-dessous intégré en salle de cinéma au TCL Chinese Theater d'Hollywood © Iosono-Barco

Le Wave Field Synthesis est l'une des techniques utilisées pour la fabrication des barres de son, ces équipements

intégrés qui simplifient la mise en œuvre du son multicanal auprès du grand public. L'industrialisation de cette technologie est principalement l'œuvre du projet de recherche européen Carrouso qui vit l'émergence des marques Iosono (aujourd'hui propriété de BARCO, fabricant de matériel pour les salles de cinéma) et Sonic Emotion, qui développe principalement ses produits dans le domaine des salles de spectacles et de l'électronique grand public, notamment ce marché très prometteur des barres de son. Côté grand public, SAMSUNG et YAMAHA conservent une avance certaine avec des modèles équipés ou non de satellites, vendus sous la barre psychologique des 1 000,00 €.

Côté mise en pratique, Hervé DÉJARDIN explicite le mixage WFS pratiqué à Radio France et la diffusion linéaire en salle dans le cadre des sessions "Du cinéma pour vos oreilles". Retour également sur les différents rendus immersifs proposés sur le portail audio 3D du groupe, dont le player nouvOson et sa vaste sélection de choix d'écoute (stéréo, binaural, multicanal 5.1, 7.1, 8.0 ou ambisonic) :



## Technologie Beam Forming

L'idée d'utiliser un ensemble de haut-parleurs pour façonner le son à l'aide de *beam forming* existe depuis de nombreuses années, mais jusqu'à présent, il était difficile de le mettre en pratique. Le *beam forming* repose sur la mise en œuvre de haut-parleurs qui reproduisent le même signal d'entrée mais de façons différentes :

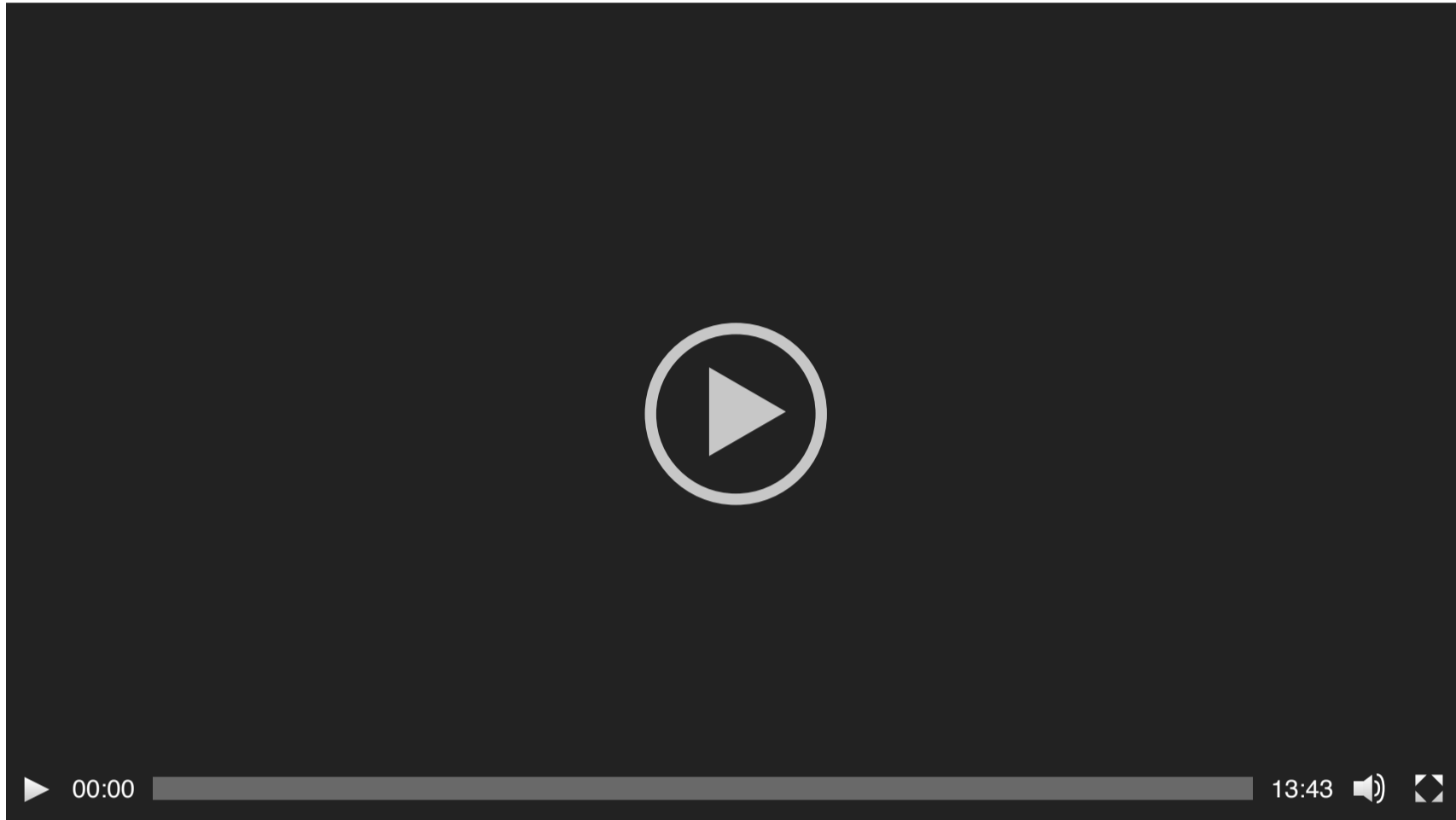
- en le retardant légèrement ;
- en le jouant à des volumes différents ;
- en utilisant des effets d'annulation croisée (technologie transaural).

Ces combinaisons de réglages permettent au système de contrôler la taille, la forme et le sens de l'onde acoustique. À la différence du procédé WFS et ses ondes planes, le beam forming utilise la diffusion d'ondes cardioïdes contrôlées pour un rendu homogène en tout point de la pièce, la personnalisation de l'écoute en plus. Du reste, assurer un rendu cohérent pour les basses fréquences (taille de membrane relativement grande) et les hautes fréquences (faible distance entre les haut-parleurs) est une tâche ardue, à cause de la plage de longueurs d'ondes très vaste du spectre audio. La réconciliation de ces deux exigences passe généralement par la mise en place d'un nombre relativement important d'enceintes, chacune contrôlée individuellement. La mise en forme dynamique de l'onde acoustique devient alors très coûteuse en terme de process (traitement du signal en temps réel), impensable pour des applications grand public jusqu'alors. Mais avec la chute du coût des chipsets DSP, la technologie est devenue assez économique pour être appliquée dans des produits grand public.

# Proposition de l'université de Southampton

Les scientifiques audio 3D de l'Institut de recherche sonore et vibratoire de l'Université de Southampton sont à l'origine d'une solution innovante qui associe du *head tracking* avec une barre de son "maison", en vue de proposer une écoute immersive personnalisée pour différents auditeurs dépourvus de casques, indépendamment de leur position dans la pièce . En entrée, le procédé reste compatible avec les principales technologies OBA du marché (Dolby, Fraunhofer, DTS).

Le procédé n'en est qu'à l'état expérimental mais la démonstration est pleine de promesses :



Ce contenu a été publié dans [audio](#), [Audio](#), [OBA](#). Vous pouvez le mettre en favoris avec [ce permalien](#).