

< Référence >

b-com-10-2017-090

< Objet >

Consultation Publique CSA « Avenir de la TNT »

< Date >

Rennes, le 19 octobre 2013

Conseil Supérieur de l'Audiovisuel

Consultation publique sur l'avenir de la TNT
39-43 quai André Citroën
75739 PARIS CEDEX 15

Madame, Monsieur,

Veillez trouver ci-joint la réponse de l'Institut de Recherche Technologique b<>com à la consultation publique concernant l'avenir de la TNT.

Grâce à ses innovations, l'Institut de Recherche Technologique (IRT) b<>com contribue à la transformation digitale européenne. Ses 230 chercheurs développent des outils, produits et services qui facilitent la vie quotidienne. Ils se concentrent sur deux grands domaines de recherche : l'hypermédia (images ultra haute définition, son spatialisé, contenus intelligents, confiance et identité numérique, réalités virtuelle et augmentée...) et les réseaux ultra haut débit plus agiles (cloud, cybersécurité, ultra haut débit mobile, résilience des réseaux, Internet des objets...). Parmi les multiples domaines d'application de ces technologies, la e-santé permet à b<>com de participer à la révolution numérique du système médical.

Issu d'un partenariat public/privé (à travers le Plan Investissements d'Avenir), l'IRT réunit les meilleurs experts issus du monde industriel et académique sur son campus de Rennes, et ses sites de Lannion, Brest et Paris. Parmi ses investisseurs privés, on peut citer les groupes Orange, TDF, Harmonic, ou des PME comme kerlink, Astellia. L'INRIA, Institut Mines Telecom ou encore ESC Rennes font partie de ses membres académiques.

La forte implantation de b<>com en région bretonne résulte de la longue histoire d'excellence de ce territoire dans le développement de technologies liés aux télécommunications et à l'audiovisuel.

Quoi qu'encre encore jeune, b<>com a développé grâce à ses projets de recherche propre une expertise solide en compression vidéo en déposant de nombreux brevets sur le sujet, dans la perspective de l'élaboration d'une future norme de compression au-delà de HEVC. b<>com a acquis une renommée internationale sur les problématiques associées au format HDR et plus particulièrement dans le contexte de la transition vers un futur « tout HDR ». Ainsi, l'IRT a développé et licencié une technologie de conversion SDR vers HDR qui a obtenu le prestigieux prix de l'innovation technologique lors de l'édition 2017 du NAB. Cette technologie est aujourd'hui considérée comme une référence sur le marché. Enfin, grâce ses projets de recherche propre et à ses contributions à des projets collaboratifs H2020 (exemple du projet Orpheus: <https://orpheus-audio.eu/>), de recrutements ciblés, b<>com a par ailleurs développé une compétence recherchée dans le domaine de l'audio, elle aussi reconnue à l'échelle mondiale. Dans ce contexte, b<>com a par exemple réalisé la keynote d'ouverture de l'édition 2017 du séminaire EBU « Object Based Audio » et était intervenant lors de plusieurs sessions focalisées « Next Generation Audio » lors de l'édition 2017 de l'IBC, qu'il s'agisse

de sujets liés à l'expérience comme aux technologies. Aujourd'hui, b<>com accorde une importance importante au soutien de l'écosystème français en matière d'innovation dans les technologies audio, en se mettant à la croisée des meilleures équipes académiques et de startups qui souhaitent développer leur croissance au travers du transfert de technologies. C'est par exemple le cas d'Aspic Technologies qui vient de rejoindre b<>com en tant que membre, sur un sujet connexe à cette consultation : l'audio pour la réalité virtuelle. Enfin, b<>com collabore de manière très proche avec des acteurs publics de l'audiovisuel comme France Télévisions et Radio France.

Notre réponse à la consultation est formalisée en considérant le fait que les facteurs de succès suivants devront être observés pour la future plateforme TNT sur le plan du déploiement de nouvelles technologies :

- permettre des investissements progressifs pour les diffuseurs et éditeurs,
- permettre une dynamique ouverte d'innovation, au-delà des cycles et des contraintes liés en particulier à l'emploi de technologies propriétaires, en favorisant l'utilisation de technologies ouvertes. Ainsi favoriser, sinon l'émergence de nouveaux acteurs qui pourront contribuer à l'évolution de cette plateforme, ancrée dans notre patrimoine audiovisuel, du moins créer de la valeur pour l'ensemble des acteurs parties prenantes, sur un marché horizontal,
- permettre de rester à un niveau de référence en Europe et dans le monde en termes de qualité d'expérience audiovisuelle, ce dans l'ensemble des dimensions d'expérience qui reposent sur l'emploi de technologies (immersion, personnalisation, interactivité, versatilité).

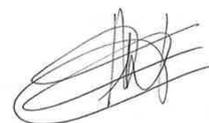
Ainsi, la réponse de b<>com à cette consultation se concentre volontairement sur les questions suivantes :

- Question #4 concernant les technologies de l'image
- Question #5 concernant les technologies audio. La réponse de l'Institut de Recherche Technologique b<>com à cette question a été élaborée avec la contribution de l'Institut Fraunhofer IIS (Institute for Integrated Circuits), partenaire de b<>com sur des problématiques conjointes de développement économique autour de nos technologies audio respectives
- Question #12 concernant la pertinence du choix DVB-T2/HEVC

Vous en souhaitant bonne réception, nous vous prions, Madame, Monsieur, de bien vouloir agréer nos cordiales salutations.



Bertrand GUILBAUD
Directeur Général



Ludovic NOBLET
Directeur Hypermedia

Réponse de l'Institut de Recherche Technologique b<>com à la consultation publique du Conseil Supérieur de l'Audiovisuel « Avenir de la TNT »

1. Question #4

« Parmi ces technologies, lesquelles sont stabilisées et disponibles commercialement ? Voyez-vous d'autres évolutions technologiques relatives à la qualité de l'image ? Quelles sont les technologies apportant le plus de gain en matière d'expérience de l'utilisateur par rapport à la HD ?

- Ces différentes fonctionnalités sont-elles rétrocompatibles avec les écrans qui ne les supportent pas ?
- Avez-vous des tests et des expérimentations en cours ? Quels enseignements en tirez-vous ?
- Les pouvoirs publics doivent-ils fixer les normes appelées à être utilisées en UHD-1 phase 2 ? »

1.1. Introduction

Après l'avènement des écrans de grandes tailles, l'exigence de qualité de l'utilisateur est toujours croissante. Dans ce cadre, l'Ultra Haute Définition (UHD) a fait son apparition. Il s'agit d'améliorer l'expérience utilisateur en augmentant la définition de l'image : on passe alors de 1920x1080 pixels en HD (Haute Définition) à 3840x2160 en UHD1 (première version UHD), et à 7680x4320 en UHD2 (seconde version UHD). En complément à l'augmentation de la définition, le nombre d'images par seconde peut atteindre 120 i/s et l'étendue des couleurs est également accrue (l'espace colorimétrique BT2020). **Afin d'obtenir des images encore plus réalistes, les formats vidéo HDR (High Dynamic Range) proposent une dynamique de la luminosité étendue.** Cela se traduit par plus de détails dans les zones les plus sombres et également dans les zones les plus éclairées de l'image.

1.2. La résolution

D'une certaine manière, l'augmentation de la résolution s'est faite sous la poussée des fabricants de téléviseurs qui ont introduit sur le marché leurs premiers produits compatibles, avant même que les contenus soient disponibles. Cependant, cette introduction a permis à l'industrie de se questionner sur l'évolution des formats audiovisuels en diffusion, et de proposer d'autres améliorations telles que l'augmentation du rythme image (HFR), l'accroissement de l'espace colorimétrique (WCG) ou bien encore la dynamique étendue du signal (HDR).

Il est souvent reproché au format UHD (3840 x 2160) d'être difficilement distinguable du format HD (1920 x 1080) lorsque la distance à l'écran est supérieure à 1,5 fois la hauteur de ce dernier. C'est d'ailleurs pour cette raison qu'un certain nombre d'acteurs du domaine se posent la question d'introduire les améliorations telles que le HDR ou le HFR sur des services au format HD. Cette approche très pragmatique semble pourtant aller à contresens de l'évolution des formats. L'histoire de la télévision montre une quête constante d'amélioration de la qualité d'image, et la résolution fait partie des facteurs à prendre en compte. **Par ailleurs, la combinaison des différentes améliorations apportées par l'UHD-1 phase 2, et en particulier le HDR, améliore la**

perception de l'augmentation de résolution. Enfin la taille des écrans augmentant en moyenne, le besoin en résolution va de pair.

Côté production, il faudra nécessairement une période d'adaptation afin que les fournisseurs/éditeurs de contenus puissent s'équiper et s'approprier ce format UHD qui est plus difficile à produire/filmer que la HD. Les capteurs actuels imposent une profondeur de champ réduite et le moindre mouvement de caméra introduit du flou, d'où l'intérêt de coupler l'UHD avec une fréquence d'acquisition plus élevée (HFR).

1.3. Le HDR

Le HDR est avant tout une remise à niveau des standards de télévision par rapport à l'évolution de la technologie des caméras et des écrans. Il ne faut pas oublier que les standards actuels ont été définis à l'époque des tubes cathodiques. Idéalement, on souhaiterait que l'image télévisée soit adaptée aux caractéristiques de la vision humaine et qu'elle procure le même rendu qu'une fenêtre ouverte sur le monde réel. Le HDR va dans cette direction même si les contraintes à la fois techniques et économiques ne permettent pas d'atteindre pleinement cet objectif. Un format HDR « optimal » posséderait les mêmes caractéristiques que la vision humaine. Il permettrait de restituer des images qui reproduisent totalement la vision humaine. Cependant, l'utilisation d'un format HDR « optimal » pose des problèmes en termes de transmission (bande passante) et en termes de restitution (consommation énergétique). Les formats et les écrans HDR actuellement proposés utilisent donc un codage sur un nombre limité de bits et sur une gamme de luminance réduite.

Le consensus actuel est d'utiliser un signal codé sur 10 bits et d'adresser des écrans procurant un maximum de luminosité allant jusqu'à quelques centaines de cd/m² (« nits »), voire supérieur à 1000 cd/m² (« nits »). Il y a actuellement deux fonctions de transfert (EOTF) prévues dans le standard ITU-R BT.2100 : PQ et HLG.

Dans le cas de l'EOTF PQ, un téléviseur standard (SDR pour « standard dynamic range ») ne pourra exploiter directement le signal HDR. Un opérateur diffusant des programmes en HDR devra adapter des techniques de diffusion simultanée de deux versions SDR et HDR du même programme (« simulcast »), ou bien recourir à des solutions s'appuyant sur l'envoi de métadonnées permettant la reconstruction d'une version à partir de l'autre.

Dans le cas de l'EOTF HLG, le signal est censé être rétro-compatible et permettre à un téléviseur SDR de ne prendre que la partie du signal le concernant, alors qu'un téléviseur HDR exploitera le signal complet. Il faut néanmoins être prudent sur l'étendue de cette rétrocompatibilité car le HLG est combiné avec l'espace colorimétrique BT.2020 qui n'est supporté que sur les nouveaux téléviseurs UHD-1 phase 1. Idéalement, pour assurer la rétrocompatibilité la plus large avec le parc de téléviseurs installés, il faudrait un signal HLG dans l'espace colorimétrique BT.709 mais cela reviendrait à diminuer la qualité du signal HDR, ce qui n'est pas non plus l'effet souhaité. On peut donc penser que même dans le cas d'un signal HDR HLG, **il sera nécessaire de disposer de fonctions d'adaptation qui permettent de réaliser la conversion BT.2020 vers BT.709 pour les téléviseurs SDR les plus anciens. Cette action devrait a priori être examinée par DVB.**

De la même façon, pour un signal HDR produit avec une référence normalisée (i.e. une référence en termes de maximum de luminance), se pose la question de l'adaptation aux caractéristiques réelles des téléviseurs HDR, diverses, tout en maintenant l'intention sinon artistique du moins l'intention en production. Ainsi, DVB se penche aussi sur la question de l'emploi et de la standardisation de

métadonnées pour ce faire. La position de b<>com est que, sur ce sujet des métadonnées, seule une solution standardisée permettra de s'affranchir de solutions propriétaires qui ont pour principal inconvénient que de créer un effet de dépendance technologique sur l'ensemble de la chaîne de valeur. **b<>com sera vigilant quant au fait que des solutions standardisées puissent aboutir sur le plan des métadonnées HDR.**

b<>com a beaucoup travaillé sur le sujet de la conversion de contenus SDR en HDR (PQ ou HLG). Il ne faut pas se tromper sur l'enjeu de cette solution ; il ne s'agit pas de montrer qu'un signal SDR est suffisant pour fournir un service HDR, et d'ailleurs, la plupart des téléviseurs HDR actuels proposent déjà une fonctionnalité native de conversion du signal SDR adaptée au mieux aux capacités de leur dalle. La technologie développée par b<>com trouve tout son sens dans un environnement mixte, où des signaux SDR côtoient des signaux HDR natifs. C'est par exemple le cas lors d'une captation réalisée à partir d'un ensemble de caméras hétérogène, certaines HDR et d'autres SDR, il est alors nécessaire de faire la conversion du SDR en HDR de manière à ce que le mélangeur traite des signaux de même nature. La conversion est également intéressante sur un programme linéaire diffusé en HDR pour lequel certains contenus (publicités, bande annonce, etc.) n'existent qu'en SDR. Afin d'éviter la commutation entre les deux formats qui serait déconcertante pour le spectateur, il est souhaitable de réaliser la conversion SDR vers HDR à la volée dans la régie finale. Un autre cas d'usage de la solution est la conversion des contenus SDR existants pour une utilisation future sur une chaîne diffusée en format HDR. **On constate donc que la solution b<>com est utile dans la période de transition vers le HDR et pour la conversion des archives.**

Néanmoins, même si cette conversion est faite avec la meilleure qualité possible et le respect de l'intention artistique, **le format HDR ne peut s'exprimer pleinement que quand il est natif**, c'est-à-dire capté à partir de caméra offrant une dynamique suffisante. Ce n'est qu'à cette condition que la diffusion de contenus en format HDR pourra se justifier.

1.4. Le HFR

b<>com a expérimenté des captations réalisées à 100 et 120 images par secondes (HFR) dès 2014 afin de se constituer une banque de séquences de test pour ses travaux sur la compression de débit ou pour l'étude des pré et post-traitements nécessaires. Certains travaux ont ainsi évalué la possibilité de recréer artificiellement le flou de bougé («motion blur») à partir d'une captation HFR lorsqu'elle est affichée sur un écran standard (LFR pour « Low Frame Rate»). Des travaux récents visent à adapter automatiquement le rythme image en fonction de la nature du contenu.

Jusqu'à présent, le format HFR a fait l'objet d'une réticence de la part des fabricants d'équipements grand-public, motivée principalement par l'augmentation importante du débit nécessaire sur les interfaces d'entrée-sortie. Il faudra donc attendre le déploiement du nouveau standard HDMI 2.1 pour que les premiers téléviseurs HFR voient le jour. Pourtant, l'apport du HFR peut être déterminant pour certains types de contenus, en particulier pour les images de sport. Côté production, le prix à payer est une bande passante plus élevée qui augmente les besoins en stockage, en capacité de calcul, etc. Par contre, la captation devient plus aisée notamment en résolution élevée (UHD) à condition que l'éclairage soit suffisant. **Il faut aussi noter que le transport du HFR est très peu pénalisant pour le débit compressé.** Le codeur, disposant d'images plus proches les unes des autres, est plus efficace et bien que le nombre d'images soit doublé, il n'y a que quelques pourcents de différence sur le débit compressé. Il n'y a sans doute pas d'intérêt à diffuser systématiquement l'ensemble des contenus en HFR, mais une adaptation sélective selon le type de programme peut apporter un vrai bénéfice en termes de fluidité et qualité d'affichage, notamment sur les contenus sportifs.

Enfin, **la rétrocompatibilité avec les écrans existants est assez simple à mettre en œuvre côté récepteur puisqu'une simple décimation d'images suffit**, avec éventuellement un post-traitement pour réintroduire un peu de flou de mouvement.

2. Question #5

« L'amélioration de la qualité sonore sur la TNT a-t-elle le même degré de priorité que l'amélioration de l'image ? Par l'emploi de quelles technologies ? Doit-elle nécessairement accompagner l'amélioration de la qualité de l'image ? Est-il nécessaire de rendre possible l'introduction de composantes sonores ou d'accessibilité supplémentaires sur la TNT ? »

2.1. Introduction

Le développement d'une nouvelle génération de technologies audio (formats comme codecs), connues sous l'acronyme « NGA » (Next Generation Audio), permettent d'envisager une évolution significative de l'expérience audiovisuelle sur la composante sonore, à la hauteur de l'évolution sur l'image (ultra haute-définition, high dynamic range, high frame rate par exemple). Au delà du sentiment d'immersion, souvent mis en avant en termes d'expérience audiovisuelle, ces nouvelles technologies permettent d'améliorer considérablement d'autres aspects de la qualité d'expérience comme l'accessibilité ou encore le confort d'écoute (audio description, réhaussement de dialogue par exemple). A l'heure d'imaginer le futur de la plateforme TNT française, ces points concernant la modernisation de l'expérience sonore nous semblent primordiaux et en accord avec l'ambition d'une plateforme populaire, délivrant des programmes au plus grand nombre et proposant une expérience audiovisuelle de référence, aussi bien en termes de qualité d'image que de qualité sonore.

Dans le but d'offrir la meilleure expérience audio aux téléspectateurs, la proposition de b<>com consiste à déployer la technologie MPEG-H 3D Audio qui est aujourd'hui le seul système de « Next Generation Audio » (NGA) prêt et entièrement fonctionnel sur des services réguliers (exemple du déploiement réalisé en Corée du Sud et détaillé plus bas dans ce document).

Dans cet optique, la réponse de b<>com est ainsi focalisée sur une description de la technologie MPEG-H Audio.

2.2. Le système MPEG-H Audio

Conçu pour compléter les évolutions technologiques relatives à la qualité de l'image, la norme internationale MPEG-H Audio apporte un ensemble de fonctionnalités nouvelles qui ont été développées par le comité MPEG à la suite d'un processus concurrentiel et collaboratif clairement défini. En offrant une expérience audiovisuelle immersive ainsi qu'en offrant les outils nécessaires à une restitution sonore personnalisée et interactive, le MPEG-H 3D Audio Low Complexity (LC) Profile Level 3 est le système audio le mieux adapté pour l'accompagnement des services vidéo Ultra Haute Définition.

Comme spécifié au sein de la norme DVB ETSI TS 101 154 v2.3.1 [1], **MPEG-H Audio remplit l'ensemble complet des exigences du NGA** définies par DVB dans le document «*Commercial Requirements for Next Generation Audio*». Le MPEG-H Audio fournit en outre un ensemble de **fonctionnalités uniques** qui font de lui le système le plus développé pour la distribution de contenu audio immersif, et ce **indépendamment du support de distribution** (*broadcast, streaming, hybrid delivery*).

2.2.1. Personnalisation et Interactivité

Le MPEG-H Audio met à disposition des téléspectateurs de nouvelles façons d'interagir avec le contenu audio et de le personnaliser selon leurs préférences. En effet, la flexibilité offerte par des objets audio, transmis indépendamment les uns des autres, autorise une modification contrôlée du mix audio au niveau du récepteur. L'utilisateur final se voit ainsi offrir la capacité de remplacer certains objets par d'autres, d'augmenter/réduire leur volume sonore ou encore de modifier leur positionnement dans l'espace 3D. Les métadonnées audio du MPEG-H permettent aux diffuseurs et éditeurs d'activer ou de désactiver les options d'interactivité et de définir strictement les limites dans lesquelles l'utilisateur peut interagir avec le contenu. Ainsi, ces capacités leur permettent d'envisager une innovation continue.

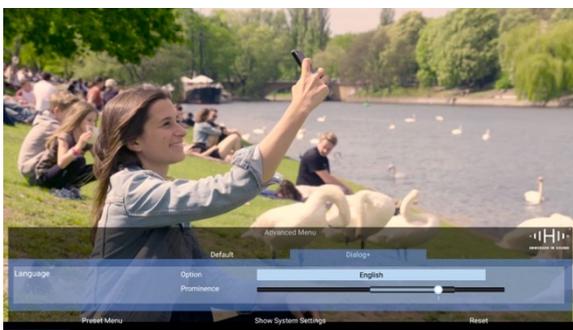


Figure 1 – Exemple d'interface utilisateur MPEG-H pour la sélection de *presets* et autres paramètres d'interactivité

De plus, les métadonnées audio du MPEG-H permettent de définir plusieurs versions d'un même contenu, aussi appelées *presets*. Ceux-ci décrivent précisément comment les différents signaux audio (canaux, objets et/ou Ambisonics) peuvent être présentés au téléspectateur. Choisir entre différents *presets* est le moyen le plus simple d'interagir avec le contenu. Certains paramètres d'interactivité plus avancés peuvent aussi être proposés au

téléspectateurs plus expérimentés. (voir Figure 1).

2.2.1.1. Services d'Accessibilité

En utilisant l'audio orienté objet (Object Based Audio, OBA) qui est partie intégrante du concept de « Next Generation Audio », le MPEG-H Audio offre **des services d'accessibilité avancés** tels que l'audio description, le multilingue, ou l'augmentation (relative) du volume des dialogues pour les malentendants ou les téléspectateurs exposés à un bruit ambiant important.

2.2.1.1.1. Audio Description

Le MPEG-H Audio permet la diffusion de l'audio description (AD) dans plusieurs langues qui peuvent être organisées en *presets*. Ces *presets* sont à la base de la sélection automatique faite au niveau du récepteur (priorité) ainsi que du choix offert à l'utilisateur pour une sélection manuelle.

Le MPEG-H Audio permet également de **déplacer spatialement le signal d'audio description vers un emplacement choisi par l'utilisateur** pour ainsi séparer spatialement le dialogue ou commentaire principal du signal d'audio description. Cela a pour effet **de considérablement améliorer l'intelligibilité des deux signaux audio**. Comme montré dans de nombreuses démonstrations techniques, toute interaction du téléspectateur s'effectue sans artefact et de manière parfaitement transparente, que ce soit en activant/désactivant l'audio description, en changeant la langue ou en déplaçant un objet audio vers la gauche ou la droite.

Enfin, le système permet de **continuer à utiliser le matériel et les outils existants** notamment pour la génération automatique de séquences de « ducking » (abaissement relatif du fond sonore pour mettre en valeur la voix), en tant qu'étape précédant l'encodage du signal.

2.2.1.1.2. Réhaussement de dialogue

Le MPEG-H Audio inclut **un réglage avancé du volume des dialogues** (*Dialogue Enhancement*, DE) soit par sélection automatique faite au niveau du récepteur (priorité) ou alors par une sélection manuelle de l'utilisateur.

Par exemple, le système prend en charge **la personnalisation du Dialogue Enhancement** au niveau de l'interface utilisateur avec la possibilité de choisir le volume, jusqu'à une valeur minimale/maximale contrôlée par le diffuseur. L'échelle de valeurs permises pour l'interaction peut être définie différemment pour différents contenus.

Le MPEG-H Audio est le seul codec NGA qui **compense automatiquement les changements de volumes** résultant des modifications de l'utilisateur afin de garantir **un niveau sonore constant**, non seulement entre les programmes mais aussi après chaque modification de l'utilisateur.

Le système MPEG-H Audio est également parfaitement compatible avec les solutions existantes de *Dialogue enhancement* et permet l'utilisation d'une seule et même interface utilisateur indépendamment du contenu. MPEG-H Audio permet notamment l'utilisation des solutions existantes d'extraction de dialogues en prétraitement.

2.2.1.2. Services Multilingue

En utilisant une seule configuration de canaux audio avec différents objets audio pour chaque langue ou audio description, le MPEG-H Audio **nécessite un débit nettement plus faible** que n'importe quel autre codec audio existant. Toutes ces fonctionnalités (par exemple, audio description et réhaussement de dialogue dans plusieurs langues) peuvent désormais découler d'un seul flux de manière beaucoup plus efficace (en termes de débit diffusé) comparativement à la diffusion simultanée de plusieurs mix complets comme cela serait nécessaire avec les codecs audio actuels. Cela simplifie notamment le processus de signalisation et de sélection requis du côté récepteur. **Table 1** montre une réduction de débit de 50% pour offrir un service 5.1 dans 3 langues différentes en utilisant MPEG-H par rapport à HE-AAC.

Table 1 Comparaison de débits entre codecs existants pour des services multilingues

5.1 Surround pour 3 Langues			
Débit MPEG-H Audio		Débit HE-AAC	
5.1 Surround	144 kbps	5.1 Surround en Langue 1	160 kbps
Langue 1	32 kbps	5.1 Surround en Langue 2	160 kbps
Langue 2	32 kbps	5.1 Surround en Langue 3	160 kbps
Langue 3	32 kbps		
Total	240 kbps	Total	480 kbps

2.2.1.3. *Personnalisation des programmes événementiels*

Les programmes événementiels, par exemple ceux liés au sport, sont des événements télévisuels importants de direct pour le grand public. Pour divers types de programmes, tels que les programmes sportifs, le MPEG-H Audio peut offrir des **options d'interactivité et de personnalisation supplémentaires**, telles que par exemple :

- Choix entre différentes sources de commentaires sportifs
- Écouter la *team radio* entre un pilote et son manager dans une course automobile
- Ou encore, se concentrer sur l'ambiance d'un stade, sans commentaire, lors d'un match de sport

2.2.2. *Son Immersif*

La diffusion et la reproduction du son immersif est une caractéristique clé des systèmes NGA. Contrairement au son 5.1 surround classique, le son immersif nécessite la reproduction de la 3^{ème} dimension sonore, pour constituer une ambiance homogène dans toute la pièce d'écoute. En effet, en élargissant l'image sonore dans la dimension verticale, le son peut maintenant provenir de toutes les directions (y compris au-dessus ou au-dessous de l'auditeur), et offre ainsi au téléspectateur une expérience plus enveloppante et réaliste du point de vue de la perception. Cela fait que, dans son expérience finale, le spectateur devient partie intégrante du contenu audiovisuel.

Le MPEG-H Audio est **le seul codec NGA capable de fournir un son immersif en utilisant n'importe quelle combinaison des trois formats audio suivants** :

- Multicanal,
- Objets Audio,
- Ambisonics (First/Higher Order Ambisonics, FOA/HOA)

Concernant le format HOA, il est important de noter qu'il permet de reproduire une scène audio complexe comportant de manière intrinsèque un nombre illimité d'objets (par exemple, le son d'un stade capté en tribune). Il permet également une rotation aisée du champ sonore ce qui en fait le format de choix pour la accompagner la vidéo à 360°. Ainsi, le format « Ambisonics » s'est imposé comme un standard incontournable pour la diffusion de contenus immersifs de type vidéo 360 ou VR.

Le MPEG-H 3D Audio Low Complexity (LC) Profile Level 3 permet de décoder simultanément jusqu'à 16 éléments audio (canaux, objets ou HOA), tandis que jusqu'à 32 éléments audio peuvent être transportés simultanément dans un même flux. **Par ailleurs, le codec MPEG-H Audio est le seul codec NGA aujourd'hui capable de décoder des signaux HOA.**

2.2.3. *Système Audio Multi-écrans*

Le MPEG-H Audio n'est pas uniquement un codec audio, mais **un système audio complet permettant de transporter et restituer de manière efficace la meilleure expérience audio possible, indépendamment du système de reproduction utilisé**. Pour cela, le système inclut des fonctionnalités de rendu et de mixage (e.g. *downmixing*), ainsi que des fonctions avancées de *Loudness* et *Dynamic Range Control* (DRC).

Le module de normalisation sonore assure un volume de reproduction constant entre les programmes et les différents chaînes, quels que soient les *presets* et systèmes de reproduction utilisés. A l'aide des informations de *Loudness* contenues dans le flux MPEG-H, le récepteur peut automatiquement effectuer une normalisation instantanée du volume sonore lorsque l'utilisateur bascule entre différents

réglages. Les informations de volume spécifiques à un *artistic downmix* peuvent être également fournies.

2.2.4. Connectivité des Périphériques

Les utilisateurs peuvent décider de connecter leurs appareils de diverses manières, comme connecter un boîtier décodeur à un téléviseur via HDMI ou connecter un téléviseur à un *home theater / audio-video recorder (AVR)*, barre de son via une interface HDMI ou SPDIF, et cela, tout en conservant l'interface d'interactivité utilisateur sur son appareil de choix (e.g. sur le téléviseur).

Pour de tels cas d'utilisation, le MPEG-H Audio fournit **une solution unique pour séparer le traitement de l'interactivité utilisateur de l'étape de décodage**. La structure en paquets du flux audio MPEG-H et l'interface standardisée pour les métadonnées d'interactivité permettent une architecture distribuée entre un appareil source (e.g. TV, STB) et un récepteur (e.g. AVR, barre de son), comme le montre la Figure 2:

- **Interface utilisateur située dans l'appareil source (TV ou STB),**
- Le résultat d'une interaction utilisateur est inséré, en tant que métadonnées, dans le flux audio MPEG-H de l'appareil source (TV ou STB),
- Le flux audio MPEG-H est alors transmis via HDMI, HDMI ARC ou SPDIF au périphérique de lecture (AVR ou barre de son),
- **Le décodage MPEG-H avec restitution sonore est effectué dans le périphérique de lecture** (AVR ou barre de son), prenant en compte les métadonnées de l'interaction utilisateur effectuée dans l'appareil source.



Figure 2 – Modèle d'interface utilisateur.

Dans ce cas, **aucun décodage ni transcodage n'est nécessaire dans l'appareil source** (TV ou STB), ce qui permet une **implémentation largement simplifiée** au niveau du périphérique source.

Le codec MPEG-H Audio est **le seul codec NGA qui permet la transmission d'un son immersif au périphérique final de lecture tout en permettant une interactivité dans l'appareil source (e.g. TV, STB)**. Tout autre codec NGA nécessite un décodage et une reproduction complète dans l'appareil source avant d'être réencodé dans un format soit différent ou identique au précédent pour la transmission par HDMI ou SPDIF au périphérique de lecture, ceci introduisant une charge de complexité significative au sein de l'appareil source.

2.2.5. Lecture Audio Immédiate

MPEG-H Audio est **le seul codec NGA qui autorise la mise en place de véritables points d'accès aléatoires (Random Access Points, RAP)** dans le flux audio. La norme MPEG-H 3D Audio spécifie ce qu'on appelle une *Immediate Playout Frame*, qui contient des informations sur les trames audio précédentes et ainsi permet une commutation transparente entre différents *presets*, configurations ou encore débits transmis.

Des changements de configuration précis (au sample près) peuvent être obtenus en utilisant un mécanisme pour tronquer les trames audio avant et après le point de transformation, comme illustré sur la Figure 3.

Le **MPEG-H Audio permet d'effectuer des changements de configuration transparents et précis**, comme lors d'insertions de publicités et ce sans augmenter le débit requis. Les RAP des flux audio et vidéo peuvent alors être alignés et découpés de manière transparente dans le domaine codé pour l'insertion de programmes locaux ou l'insertion de publicités, par exemple.

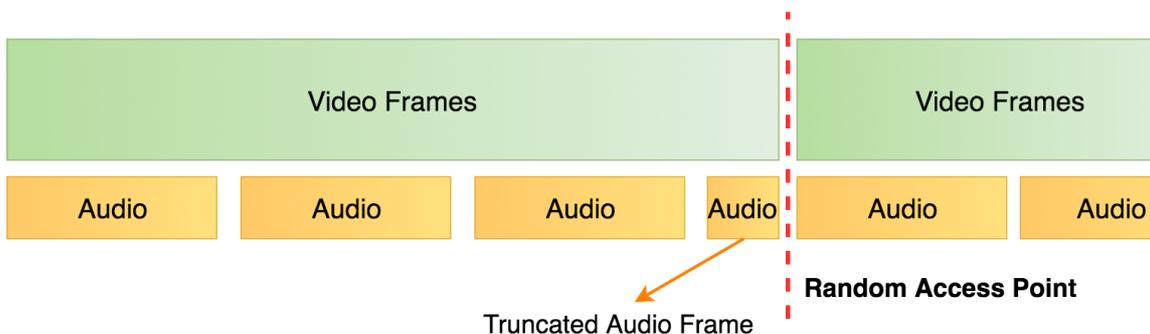


Figure 3 - Exemple d'alignement A/V au niveau des RAP.

2.2.6. Transport de Métadonnées Audio Robustes sur SDI

Les systèmes audio de nouvelle génération nécessitent des métadonnées supplémentaires qui doivent être transmises avec le contenu audio à travers les interfaces de production. Bien qu'il soit probable que les futures installations soient équipées d'une infrastructure IP, l'infrastructure actuelle utilise encore majoritairement les connexions SDI. MPEG-H Audio est le seul système NGA qui offre une solution robuste pour la transmission de métadonnées sur SDI afin de permettre l'ensemble des fonctionnalités NGA sur les infrastructures d'aujourd'hui.

Pour cela, les métadonnées MPEG-H Audio sont modulées et transmises en tant que canal PCM sur une piste SDI. Cette piste est synchronisée avec la *video frame rate* et reste robuste vis-à-vis des changements de volume et re-échantillonnage.

C'est pour cela que le MPEG-H Audio est **le seul codec NGA qui offre la possibilité d'activer toutes les fonctions NGA avec les infrastructures actuelles de diffusion sans aucun coût supplémentaire.**

Les diffuseurs peuvent ainsi décider librement d'activer progressivement les diverses fonctionnalités NGA, par exemple :

- **Phase 1** : Stereo et 5.1 surround avec des configurations prédéfinies et statiques. Dans cette étape initiale, aucune métadonnée n'est requise dans le workflow.

- **Phase 2** : Stereo et 5.1 surround avec des configurations flexibles et des métadonnées dynamiques. Les configurations audio peuvent être sélectionnées et définies à l'aide d'outils de création en post-production et les métadonnées dynamiques peuvent être fournies avec l'audio en PCM (sur SDI ou sur fichier).
- **Phase 3** : Production immersive en utilisant les outils de création et de *monitoring* pour l'audio 3D avec du HOA, OBA, et les métadonnées dynamiques fournies avec l'audio PCM

2.3. Standardisation

Finalisé en 2015, MPEG-H Audio [2] est une norme ISO internationale ouverte développée dans le cadre d'un processus concurrentiel et collaboratif clairement défini. Les principales étapes du processus MPEG sont :

- Appel d'offres
- Évaluation et sélection des propositions technologiques
- Phase collaborative pour les améliorations technologiques

Un nombre significatif d'entreprises ont contribué au développement de la norme MPEG-H 3D Audio, comme par exemple : Fraunhofer, Qualcomm, Technicolor, Sony, Samsung, Orange, Apple, ETRI, Dolby.

Toutes les phases intermédiaires, du début du projet jusqu'à la spécification finale, ont été examinées conformément aux règles ISO avec le vote des comités nationaux de tous les pays membres.

Les performances du MPEG-H Audio ont été soigneusement évaluées par MPEG et documentées dans le rapport de test de vérification [3]. Avec **plus de 15000 évaluations subjectives, provenant de 288 experts audio répartis dans 7 laboratoires d'essais indépendants**, MPEG-H Audio a fait ses **preuves en matière d'excellence** pour tous les cas d'utilisation évalués.

Le MPEG-H Audio a également été évalué avec succès lors de tests d'écoute et de fonctionnalités réalisés par le comité audio ATSC qui l'a ensuite **adopté dans la suite de normes ATSC 3.0** [4].

Enfin, **DVB a également sélectionné et inclus MPEG-H Audio dans la norme DVB ETSI TS 101 154 v2.3.1** [1] qui définit l'utilisation des codecs audio et vidéo pour les systèmes DVB. La signalisation appropriée pour les systèmes MPEG-2 TS DVB a été spécifiée dans la norme ETSI EN 300 468 [5].

2.3.1. Tests de Conformité et Implémentation de Référence

Un élément important du processus de standardisation MPEG est la disponibilité d'un code source de référence ainsi que de flux audio et de critères de test pour tester les décodeurs. **MPEG a défini environ 200 flux de test pour l'audio MPEG-H**. Ils sont importants pour **permettre les implémentations indépendantes** et assurer **l'interopérabilité** entre les périphériques.

En plus des tests de conformités et de l'implémentation de référence fournis par MPEG, le partenaire technologique de b<>com, Fraunhofer IIS, soutient les fabricants avec **des kits de test supplémentaires et un Trademark Program pour les récepteurs**.

Cela fait de MPEG-H Audio **le seul codec NGA qui soit un véritable standard ouvert permettant des implémentations indépendantes**.

2.3.2. Efficacité de Codage

La supériorité et l'efficacité du codage MPEG-H 3D Audio LC ont été démontrés dans les multiples tests effectués par MPEG [3]. Le format offre **une excellente qualité audio à faible débit**, comme indiqué dans le Table 2 pour certains scénarios typiques de diffusion.

Table 2 Recommandations de Débits pour une excellente qualité audio

Configuration		Débit
2.0 Stereo		64 kbps
5.1 Surround		144 kbps
5.1 Surround pour 3 Languages		240 kbps
	5.1 Surround	144kbps
	Langue 1	32 kbps
	Langue 2	32 kbps
	Langue 3	32 kbps
7.1 +4H Immersif avec 4 haut-parleurs élevés		384 kbps
22.2 Immersif		768 kbps

A noter que les résultats du test ATSC 3.0 peuvent être consultés dans le document DVB TM-AVC0803 [6], disponible pour les membres DVB. Ces résultats indiquent que **MPEG-H Audio offre la meilleure efficacité de codage pour une excellente qualité audio.**

2.4. Autres Zones Géographiques et Services

Le MPEG-H Audio a été sélectionné par la *Telecommunications Technology Association (TTA)* **en Corée du Sud comme le seul codec audio pour leur norme de diffusion UHDTV terrestre** TTAK.KO-07.0127 [7] qui est basée sur ATSC 3.0.

Le 31 mai 2017, la Corée du Sud a lancé son service TV UHD 4K et le système audio MPEG-H est devenu le premier et le seul déploiement NGA mondial sur un service régulier. Depuis le mois d'avril, un grand nombre de modèles de téléviseurs et de décodeurs de LG et de Samsung intégrant le MPEG-H 3D Audio LC Profile Level 3 sont disponibles dans les magasins.

La Figure 4 montre le calendrier officiel du déploiement du service UHDTV en Corée du Sud. Aujourd'hui, le service offre le système MPEG-H utilisant plusieurs objets et canaux. Dans les prochaines années des services supplémentaires seront activés, comme la réception mobile, la transmission hybride et l'insertion dynamique de publicités. Tous les éléments nécessaires à la chaîne de production de contenu et de transmission pour la diffusion sont en place et utilisés dans des conditions de fonctionnement réel. Partenaire technologique de b<>com, Fraunhofer est en contact

étroit avec les diffuseurs coréens et les différents équipementiers pour s'assurer que l'expérience acquise puisse se refléter dans les futurs développements du produit.

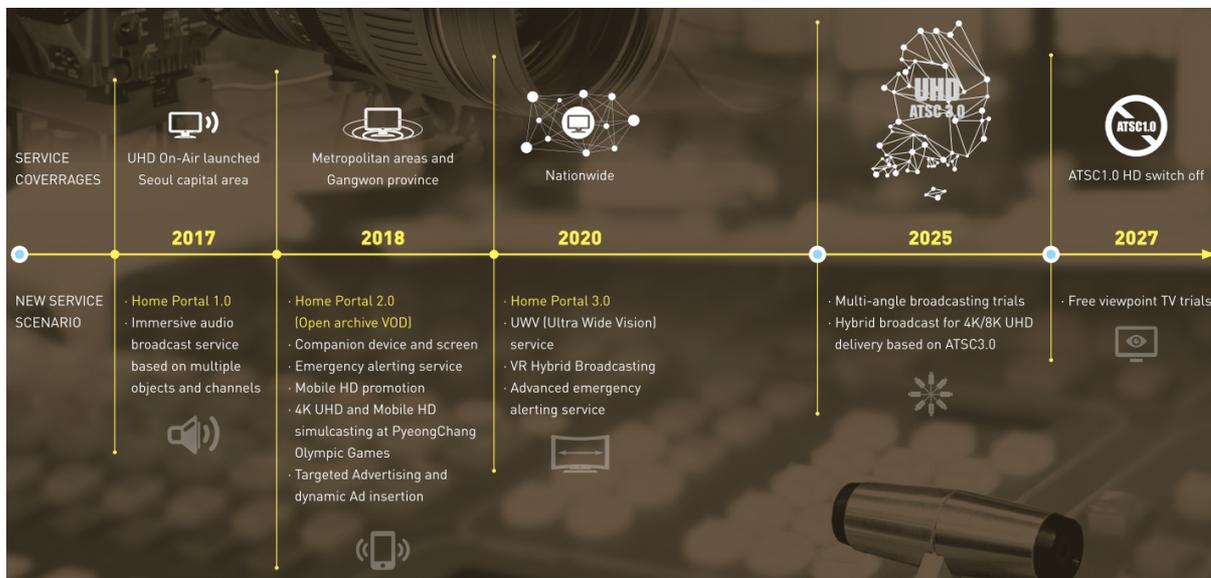


Figure 4 – UHDTV Déploiement en Corée du Sud.

Les jeux olympiques d'hiver 2018 seront diffusés en direct en MPEG-H Audio sur le service UHDTV Sud-Coréen.

A noter également que la Chine est actuellement en train de définir son système de télévision de nouvelle génération et le MPEG-H Audio est un candidat pour ce système.

2.5. Prise en Charge des Équipementiers

Les récepteurs MPEG-H Audio, capables de décoder et de reproduire **l'ensemble complet des fonctions NGA spécifiées par DVB**, sont disponibles sur le marché depuis avril 2017. De plus, les outils de production pour le Live, la post-production et l'encodage de contenu MPEG-H Audio sont disponibles chez divers fabricants, et actuellement utilisés en Corée du Sud.

Toutes les implémentations disponibles sur le marché aujourd'hui intègrent l'ensemble des fonctions NGA du MPEG-H 3D Audio LC Profile Level 3. Cela garantit une interopérabilité maximale et permet aux diffuseurs de ne commencer à utiliser aujourd'hui qu'un sous-ensemble des fonctions NGA (par exemple, seulement AD et le multilingue), sans créer de problèmes de compatibilité.

La Figure 5 fournit une vue d'ensemble des principaux fabricants qui ensemble proposent une gamme grandissante de périphériques compatibles MPEG-H :

- Outils de création et de *monitoring* en temps réel disponibles chez Jünger et Linear Acoustic. Ces appareils sont capables de créer du contenu MPEG-H Audio avec l'ensemble des fonctionnalités de NGA ;
- Encodeurs disponibles auprès de DS Broadcast, KaiMedia, PixTree et MediaExcel, avec la prise en charge de l'ensemble complet des fonctionnalités du MPEG-H 3D Audio LC Profile Level 3 ;
- Outils de post-production tels que DAW plug-ins de New Audio Technologies et DSPatial, et la prise en charge native du MPEG-H pour l'exportation audio chez Pyramix ;

- SoC pour STB / TV et AVR / Soundbar avec la prise en charge du MPEG-H Audio annoncée dans plusieurs SoC, comme par exemple MStar, Realtek et Analog Devices ;
- Téléviseurs et décodeurs de marques LG et Samsung prennent en charge l'ensemble des fonctions du MPEG-H 3D Audio LC Profile Level 3 avec jusqu'à 16 éléments audio pouvant être décodés simultanément et 32 éléments pouvant être transmis dans un seul flux.



Figure 5 – Équipementiers MPEG-H Audio.

À noter que plusieurs modèles de téléviseurs de la marque LG en Europe prennent déjà en charge le format MPEG-H Audio, avec l'ensemble des fonctionnalités NGA disponibles à partir du périphérique USB (see Figure 6)



Figure 6 – Téléviseurs LG avec MPEG-H Audio.

2.6. Licenses

Le format MPEG-H Audio est une norme ISO/IEC et par conséquent, la politique ISO/IEC relative aux droits de propriété intellectuelle (IPR) s'applique [8]. En résumé, ISO/IEC demande à ce que: "...any party participating in the work of ITU, ISO or IEC should, from the outset, draw the attention of the Director of ITU-TSB, the Director of ITU-BR, or the offices of the CEOs of ISO or IEC, respectively, to

any known patent or to any known pending patent application, either their own or of other organizations, ...”.

Les déclarations d'IPR pour le MPEG-H 3D Audio, reçues par ISO/IEC, sont publiées et peuvent être consultées sur la page Web ISO/IEC, quelques exemples sont fournis ici :

- *Fraunhofer, against license fee under FRAND conditions, submitted: Nov. 30, 2015*
- *Dolby, against license fee under FRAND conditions, submitted: Feb. 26, 2015*
- *Technicolor, against license fee under FRAND conditions, submitted: May 26, 2016*
- *Qualcomm, against license fee under FRAND conditions, submitted: July 07, 2016*

2.7. Conclusion

MPEG-H Audio est **le seul système NGA** qui est prêt aujourd'hui pour des services réguliers :

- MPEG-H Audio est une véritable **norme internationale ouverte**
- Le **seul service NGA déployé aujourd'hui** (Corée du Sud) utilise le MPEG-H Audio
- Les appareils grand public ainsi que diverses implémentations sont **aujourd'hui disponibles sur le marché**, tous prenant en charge **l'ensemble des fonctionnalités** du MPEG-H 3D Audio LC Profile 3, **telles que défini par la norme DVB**.
- Tous les éléments nécessaires à la **chaîne de production, transmission et reproduction** dans les appareils récepteurs sont **disponibles dans le commerce** auprès de différents fabricants.
- Les diffuseurs peuvent désormais passer au format MPEG-H Audio avec les contenus stéréo ou 5.1 existants, et peuvent progressivement **ajouter les nouvelles fonctionnalités étape par étape**, tandis que les appareils sur le marché prennent déjà en charge l'ensemble des fonctions NGA.
- MPEG-H Audio prend en charge **l'ensemble des exigences DVB NGA** et est particulièrement adapté pour répondre aux exigences de **services d'accessibilité**.
- MPEG-H Audio est le seul codec NGA permettant la compression de contenus audio au format ambisonique (FOA/HOA).
- MPEG-H Audio offre **la meilleure efficacité de codage** de tous les codecs audio actuellement disponibles.
- Le format MPEG-H Audio est une norme ISO/IEC et, par conséquent, la politique ISO/IEC relative aux droits de propriété intellectuelle (IPR) s'applique.
- L'adoption du MPEG-H Audio en Corée du Sud conduit à la mise en œuvre d'un écosystème MPEG-H Audio grandissant sur différentes plateformes, y compris les téléphones mobiles et tablettes.

3. Question #12

« Compte tenu des contraintes exposées ci-dessus et du calendrier dans lequel les évolutions présentées en partie 2 pourraient être introduites sur la TNT, êtes-vous d'accord sur la pertinence du choix des normes DVB-T2 et HEVC pour la modernisation de la plateforme TNT en France ? »

La réponse de b<>com à cette question ne traite que du sujet codec vidéo.

Par ailleurs, la réponse à cette question nécessite que de se projeter par rapport à un échéancier précis de migration puis de transition complète de la plateforme technologique TNT, du système actuel vers un système basé sur de nouvelles technologies au delà du socle technologique DVB-T/H.264. Cette réponse doit intégrer l'objectif d'optimisation de la ressource spectrale tout en

garantissant à l'ensemble des diffuseurs/éditeurs, le fait de pouvoir bénéficier des meilleures conditions possibles pour introduire une nouvelle expérience audiovisuelle, moderne, évolutive, tout en intégrant la disponibilité de technologies par rapport à cet échéancier.

Pour un calendrier de démarrage de migration se situant à horizon 2020, le choix HEVC semble le plus pertinent en considérant non seulement la disponibilité d'éventuels futurs standards mais aussi la maturité des implémentations industrielles pour un déploiement sur une plateforme telle que la TNT associée à un marché horizontal.

Aujourd'hui, une technologie de compression vidéo est en cours de développement au sein des organismes ITU/MPEG pour être candidate à la succession du codec HEVC. L'objectif principal, pour une même qualité d'image, est de réduire le débit nécessaire en transmission d'un facteur 2. b<>com, à travers ses projets de recherche propre, développe de la technologie en ce sens et sera partie prenante pour l'élaboration de ce futur standard de compression. Le calendrier de standardisation associé au développement de ce futur standard est le suivant :

- Appel à propositions (« Call for proposals ») : Octobre 2017
- Echéance de soumissions à cet appel : février 2018, avec une phase d'évaluation subjective des soumissions programmée pour mars/avril 2018.
- Evaluation formelle des soumissions : Avril 2018
- Disponibilité d'un futur standard envisageable pour la fin de l'année 2020

En considérant le temps requis par des organismes comme DVB pour élaborer et publier des spécifications d'usage de ce futur standard (18 à 24 mois), le temps requis pour l'industrie afin d'atteindre un niveau minimal de maturité d'implémentation comme de viabilité commerciale, une telle technologie ne sera que difficilement déployable avant un horizon 2023.

Ainsi, si un début de migration de la plateforme TNT est envisagé avec un horizon moins lointain que 2023-2025, le choix HEVC est effectivement pertinent.

References

- [1] TS 101 154 v2.3.1: Digital Video Broadcasting (DVB) – Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream.
http://www.etsi.org/deliver/etsi_ts/101100_101199/101154/02.03.01_60/ts_101154v020301p.pdf
- [2] "Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3: 3D audio," International Organization for Standardization (ISO), Geneva, Standard ISO/IEC 23008-3:2015, 2015.
- [3] N16584, MPEG-H 3D Audio Verification Test Report, <http://mpeg.chiariglione.org/standards/mpeg-h/3d-audio/mpeg-h-3d-audio-verification-test-report>
- [4] A/342 Part 3: ATSC Standard – MPEG-H System, March 2017.
- [5] ETSI DVB BlueBook A038 (2017-06), Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems (EN 300 468).
https://www.dvb.org/resources/public/standards/a038_dvb_si_spec_june_2017.pdf
- [6] TM-AVC0803, ATSC Response to DVB TM-AVC Liaison on ATSC 3.0 Audio
<https://www.dvb.org/search/results/only/documents/page/1/items/12/groupid/11/keywords/TM-AVC0803>
- [7] TTAK-KO-07.0127R1: TTA - Transmission and Reception for Terrestrial UHDTV Broadcasting Service, Revision 1, December 2016.
http://www.tta.or.kr/include/Download.jsp?filename=stnfile/TTAK.KO-07.0127_R1.pdf
- [8] ISO/IEC/ITU common patent policy: <http://isotc.iso.org/livelink/livelink/Open/6344764>