

La TVHD

dossier technique

PAGE 6

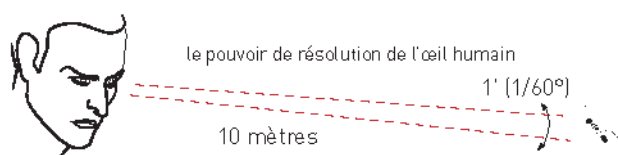


Patrice Bourcet, (Son 70 - ENLL), débute sa carrière à TDF au laboratoire d'acoustique du CERIM (Centre de Recherche d'Issy-les-Moulineaux) où il obtient son diplôme d'ingénieur CNAM en 1976 et réalise notamment un équipement de mesure instantanée des niveaux de bruit normalisés caractérisant les performances acoustiques d'un bâtiment. Il intègre ensuite le laboratoire audiofréquences où ses travaux sur les techniques numériques lui permettent de développer la technologie dite FM Synchrone expérimentée en première mondiale par Radio-France-Savoie à l'occasion des jeux olympiques d'Albertville et largement déployée aujourd'hui sur le canal FM réservé aux autoroutes : 107.7 MHz. Il crée en 1979 un cours de traitement du signal qu'il enseigne aux deux années son à l'ENLL jusqu'en 1991. En 1992, il fonde le département Traitement du Signal du Centre d'Études en Radiodiffusion et Radiocommunications implanté à Metz où il développe notamment un système de données cachées dans l'audio préfigurant le concept actuel de watermarking. Ayant déposé une dizaine de brevet lors de sa carrière au sein de la R&D de TDF, il développe actuellement une nouvelle activité de conseil.

A propos de TVHD

La qualité des différentes sources d'images numériques et leur restitution par des écrans de plus en plus grands et parfaitement plats, ne cesse de croître dans une dynamique de baisse des coûts. Les principaux fournisseurs de contenus audiovisuels ont déjà introduit l'image numérique haute définition dans leurs productions pour une part très significative en ce qui concerne le cinéma et en forte croissance pour la télévision. Enfin, l'argument qui consistait à dire, il y a 10 ans, que la haute définition n'apportait pratiquement rien compte tenu de la taille relativement réduite des écrans de télévision est en train de s'inverser. En effet aujourd'hui, les grands écrans se démocratisent et vont très rapidement révéler l'intérêt, voire créer la nécessité, d'une meilleure définition pour ne pas diminuer, mais bien au contraire, pour augmenter la résolution des images.

- La résolution d'une image détermine la distance minimale à laquelle on peut la regarder !
- Le pouvoir de résolution d'un système optique désigne sa capacité à distinguer des détails fins. Il est défini comme la distance angulaire minimale entre deux éléments d'un objet qui permet d'en obtenir deux images séparées (pouvoir séparateur); on dit alors que l'objet est résolu.
- Le pouvoir de résolution de l'œil humain est au mieux d'une minute d'angle, soit environ 100 km sur la surface de la lune vue de la Terre ou encore une fourmi à 10 mètres de distance.



- Les détails dont la finesse correspond à un angle de vue inférieur ne pourront pas être résolus. Ils seront confondus et leur éventuelle restitution au niveau des points (pixels) constituant l'image est strictement inutile.

- Si l'on considère que le pixel est le plus petit élément d'une image, la résolution de cette image est tout simplement le rapport entre le nombre de pixels qui la compose et ses dimensions réelles sur un support physique. Ce rapport s'exprime en nombre de pixels par unité de surface.
- Pour une résolution donnée, on peut donc calculer la distance minimale à partir de laquelle notre œil sera en limite de résolution, distance qui ne nous permettra donc pas de distinguer les pixels constituant l'image mais de la percevoir de manière continue.

Résolution (pixel/mm²)										
0,01	0,05	0,1	0,5	1	5	10	50	100	500	
34,38	15,37	10,47	4,86	3,44	1,54	1,09	0,49	0,344	0,15	
Distance (mètre)										

- À perception égale, la résolution peut donc diminuer comme le carré de la distance de vision. Il faut 100 p/mm² à 34 cm et seulement 1p/mm² à 3,4 m. Soulignons également que l'œil est moins exigeant pour la vision des couleurs.

La TVHD

dossier technique

7 PAGE

La même finesse de reproduction sera obtenue pour la même image sur un écran d'un mètre de diagonale (soit environ 86,4 x 48,6 cm), observée à 1,5 m de distance ; à condition que sa résolution soit au moins de 5 p/mm², soit 5x864x486 = 2099 520 pixels.

En limite de résolution, la définition d'une image détermine l'angle (solide) qu'elle occupe dans notre cône de vision !

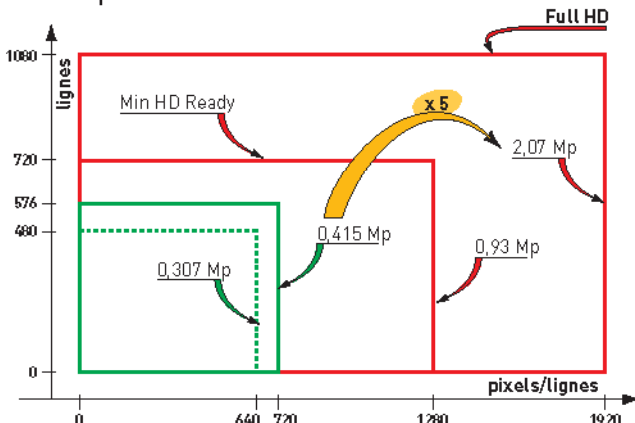
Une image possédant 1920 pixels en largeur et 1080 en hauteur aura une définition de 1920 par 1080 pixels, notée 1920 x 1080, soit : 2 073 600 pixels. Sur un écran de 86,4 x 48,6 cm la résolution correspondante sera de l'ordre de 5 p/mm² (= 2 073 600 / 864 x 486) et la distance limite de l'ordre de 1,55 m d'où un angle de vue horizontal d'environ 31,2 degrés.

Le tableau suivant donne l'angle de vue horizontal occupé par une image de 86,4 cm de large, pour différentes définitions horizontales, en conservant la distance limite de résolution propre à chaque cas de figure.

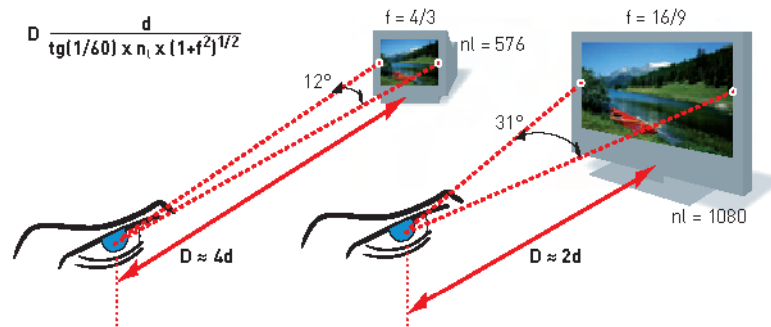
Définition horizontale (pixel)					
640	720	800	960	1280	1920
10,6	12	13,3	15,9	21,1	31,2
Angle horizontal (degré)					
TVSD			TVHD		

Comme on peut le constater, en limite de résolution, une plus grande définition augmente l'angle sous lequel l'image est vue. L'ordre de grandeur est d'environ 1 degré pour 60 pixels près de l'axe de vision.

À résolution identique, la définition du format 1920 x 1080, dit "full HD" est 5 fois plus grande que qu'en TVSD.



Comparaison des formats TVHD (en rouge) avec les formats de la TVSD actuelle (en vert, vert pointillé : format US). À résolution identique, la définition du format 1920x1080 est 5 fois plus grande.

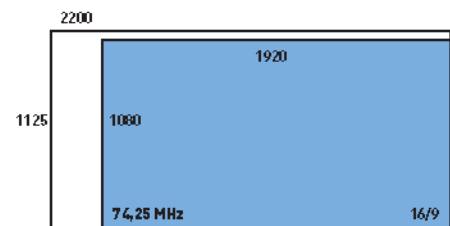
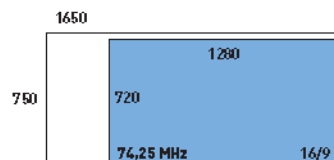


Comparaison TVSD/TVHD en limite de résolution. En pratique, le format "full HD" c'est donc une image 5 fois mieux définie que l'on peut regarder environ 2 fois plus près (relativement aux dimensions de l'écran) et sous un angle horizontal presque 3 fois plus large.

Quelles sont les normes de la TVHD aujourd'hui ?

Les deux normes fondatrices de la télévision haute définition et du cinéma numérique ont été écrites par la SMPTE.

La norme SMPTE 274 M, adoptée en 1995, regroupe les formats d'aspect 16/9 : 1920x1080 pixels utiles (dans un format 2200x1125) destinés à la télévision avec un balayage entrelacé (50 et 60 Hz) et ceux destinés au cinéma avec un balayage progressif (24, 25, 30 Hz et 60 Hz prévu pour le futur).



La norme SMPTE 296 M adoptée en 1997 définit le format d'aspect 16/9 : 1 280x720 pixels utiles (dans un format 1 650x750) pour la télévision.

Ces normes indiquent notamment les fréquences d'échantillonnage, luminance et chrominance ainsi que le nombre de bits de quantification qui déterminent le débit ou CDR (Common Data Rate) de l'interface de transmission des signaux vidéo appelée HD-SDI (Serial Digital Interface) et décrite dans la norme SMPTE 292 M.

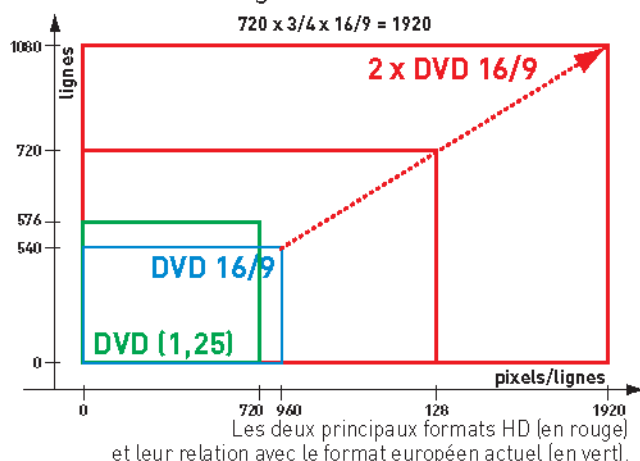
La TVHD

dossier technique

PAGE 8

Enfin, la quatrième version de la norme mondiale de spécification de l'image numérique haute définition de l'UIT (Union Internationale des Télécommunications), la recommandation UIT-R BT.709-4 de mai 2000, décrit le format 1080x1920 progressif, noté "p" (progressive) pour 24, 25, 30, 50 et 60 images par seconde, le format entrelacé, noté "i" (interlaced) pour 50 et 60 trames par seconde, ainsi que le format progressif segmenté noté "PsF" (Progressive segmented Frame), marquant l'avènement de ce format pour la TVHD. Il s'agit en fait du découpage d'une image "p" en deux segments A et B (pour ne pas les confondre avec des trames paire et impaire), de manière à pouvoir réutiliser le matériel de transmission et d'enregistrement prévu pour les images "i".

On remarquera que le format 4/3 (=1,33) du DVD (équivalent au PAL ou au SECAM) soit : 720x576 (de rapport 1,25) n'est pas un format "pixel carré". Si l'on applique la notion de "pixel carré" en conservant la définition horizontale, le nombre de lignes permettant de satisfaire cette contrainte est de $720/4 \times 3 = 540$ lignes. Le format 16/9 correspondant comporte alors : 960x540 pixels. Remarquons ensuite que le nombre de lignes correspondant à cette nouvelle définition horizontale pour un format 4/3 est égale à $960/4 \times 3 = 720$ lignes.



Ces remarques nous permettent de constater que le format 1920x1080 est un format dont les définitions horizontale et verticale sont respectivement égales au double de celles d'une version 16/9 de l'actuel DVD auquel on

aurait appliqué la notion de "pixel carré". Le format 1280x720 est, quant à lui, un format dont la définition verticale correspond à celle d'un retour en 4/3, à définition horizontale constante, du format DVD 16/9 précédemment défini. Sa définition horizontale est alors celle du format 16/9 calculé sur cette nouvelle base, soit $720 \times 16/9 = 1280$ pixels.

Les facteurs clés d'une bonne TVHD

- La pertinence des données conservées par le codeur perceptuel
- En matière de qualité, la définition d'une image n'est qu'un critère qui devient tout relatif lorsque l'image numérisée est codée pour être enregistrée, transportée, diffusée ou restituée.
- En supposant que l'on veuille enregistrer le flux numérique sortant de l'analyseur d'une caméra fournissant 25 images HD (1920x1080) par seconde, nous ne pourrions enregistrer qu'une vingtaine de secondes sur un DVD classique ! Il est donc impératif de réduire ce débit d'informations, c'est ce que va faire l'algorithme de réduction de débit, en éliminant les redondances mathématiques (compression sans perte) et les redondances psycho-visuelles (compression avec pertes).
- Le ou les codeurs que l'on interpose entre la caméra et le téléviseur vont éliminer jusqu'à 99,5 % des informations initiales.

capture image 1,5 Gbps →

redondance psycho visuelle

calcul simplifié débit binaire utile avant codage :

1 920 x 1 080 pixels
3 couleurs
8 bits par couleur } 16 777 216 couleurs
25 images par seconde
1 920 x 1 080 x 3 x 8 x 25 = 1 244 160 000 bps

information conservée
après réduction de débit :
MPEG-2 ≈ 1 %
MPEG-4 ≈ 0,5 %



Importance du codage perceptuel de l'image.
Tout l'art consiste à ne retenir que le minimum d'information, celles qui sont réellement perçues par notre œil.

L'absence de filtrage spatial ou l'identité des formats de capture et de restitution

- Les techniques numériques permettent de réaliser sans difficultés particulières toutes sortes de filtrages permettant de modifier la résolution d'une image pour qu'elle puisse

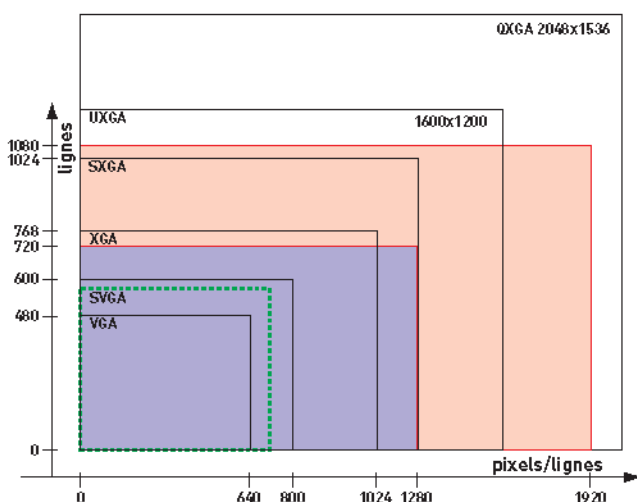
La TVHD

dossier technique

9 PAGE

être restituée sur n'importe quelle taille imposée par un support, dont les dimensions et la définition seraient différentes du support ayant servi à sa capture. Cette facilité ne doit pas nous faire oublier que tout traitement ne peut qu'altérer l'intégrité de l'information initiale et qu'il est toujours préférable de restituer une image dans sa définition et son format d'origine.

À titre d'illustration la figure suivante permet de comparer, à résolution identique, les formats "audiovisuels" et les formats "informatiques". Seul le format VGA correspond au format de télévision utilisé aux Etats-Unis. Les différences proviennent du fait que les formats audiovisuels ou cinématographiques (encore différents) résultent de compromis techniques ou historiques, tandis que les formats informatiques cherchent surtout les multiples des mots informatiques de 8, 16, 32 ou 64 bits utilisés pour les traitements d'images.



Comparaison des formats "audiovisuels" et des formats "informatiques" (en noir).

Les interfaces qui respectent le signal

L'interface DVI (Digital Visual Interface) est de plus en plus répandue pour les écrans de télévision et informatique. Elle permet de transporter un signal numérique non compressé, mais pas de son associé. L'atténuation du signal posant problème, les câbles DVI sont généralement limités à 2 m.

La génération suivante se dénomme HDMI (High-Definition Multimedia Interface). Elle permet de transporter un signal allant jusqu'à 5 Gbps, contenant un signal vidéo non compressé (de la résolution vidéo standard jusqu'à

1080p) avec un son multivoies associé. La connectique associée existe jusqu'à 15 m. De plus, l'interface HDMI est compatible avec HDCP (High-bandwidth Digital Content Protection), le système de protection des droits d'auteurs pour la diffusion (ce système ne protège pas la copie).

L'absence de filtrage temporel ou l'identité des méthodes temporelles d'analyse et de restitution

Le balayage entrelacé

Depuis 1936, toutes les normes de télévision "cathodique" utilisent le système d'analyse/restitution entrelacé quelle que soit la définition de l'image. On peut considérer qu'il s'agit de la première application de la notion de codage perceptuel tirant partie des caractéristiques propres aux visions centrale et périphérique de l'œil humain.

L'image est analysée en deux trames, une pour les lignes paires puis l'autre pour les lignes impaires, d'où le qualificatif de balayage entrelacé puisque l'image n'est complète qu'après 2 balayages imbriqués successifs. Au lieu de transmettre 50 images par seconde, on ne transmet donc que 50 demi-images (ou trames) par seconde. Toutefois, la qualité perçue n'est pas réduite dans le même rapport car on tire partie du fait que la vision centrale de l'œil (la fovéa) est plus sensible à la définition qu'au mouvement alors que c'est le contraire pour la vision périphérique. Ainsi, lorsque la vision centrale (le regard) fixe une partie de l'écran, sa faible sensibilité temporelle ne lui permet pas de distinguer la présentation successive des deux trames, d'où la perception d'une définition nominale correspondant à la somme des lignes paires et impaires. Simultanément, la faible acuité de la vision périphérique ne lui permet pas de distinguer la trame paire de la trame impaire d'où une impression globale de scintillement atténuée, compte tenu du fait que la fréquence trame est deux fois plus élevée que la fréquence image.

Pour notre œil, tout se passe donc comme si on avait transmis N images (vision périphérique) possédant n lignes (vision centrale) par seconde, alors que l'on a transmis en réalité N trames de $n/2$ lignes par secondes.

La réduction de bande passante est donc l'avantage décisif du mode entrelacé qui conserve

La TVHD

dossier technique

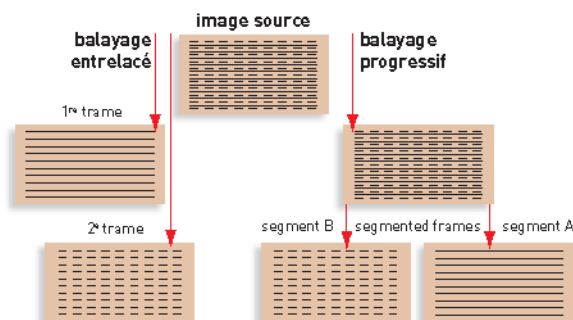
PAGE 10

une bonne analyse du mouvement. En contrepartie, un certain scintillement est encore perceptible notamment sur les surfaces blanches compte tenu de la brillance des écrans de télévision qui est nettement plus élevée qu'au cinéma, de plus un détail présent sur une seule ligne accentue le phénomène du fait qu'il n'est présenté qu'une trame sur deux.

Le balayage progressif

Si l'économie de bande passante est une exigence pour les images diffusées, le problème ne se pose pas en informatique puisque l'unité centrale et l'écran d'un ordinateur sont reliés par un fil. De plus, la scission d'une image en deux trames complique inutilement les traitements logiciels que l'on peut imaginer sur des images complètes (remarque qui s'applique également aux algorithmes de plus en plus complexe mis en œuvre pour réduire le débit). Chaque image est donc prise dans son intégralité en un seul balayage ce qui rapproche ce mode dit progressif de la prise de vue cinématographique. L'apparition du balayage progressif dans les équipements vidéo apporte de nouveaux avantages tout en conservant la possibilité d'utiliser certains équipements conçus pour l'entrelacement grâce au mode "PsF".

Outre sa plus grande compatibilité avec le monde de l'informatique, l'absence de scintillement et la stabilité de l'image sont les avantages décisifs du balayage progressif. En contrepartie, la bande passante nécessaire est élevée et l'on perçoit quelquefois un effet stroboscopique ou de saccade comparable à celui du cinéma.



Balayage entrelacé ou progressif et progressif segmenté.

Plus de lignes ou plus d'images ?

Le format d'échange international d'images numériques pour le cinéma en 24 images/seconde, le 1080/24p est considéré comme insuffisant pour le sport où il est nécessaire d'avoir 50 ou 60 images (ou trames pour la TV) par seconde pour les mouvements rapides.

La solution considérée comme le stade ultime de la HD est à l'évidence le 1080p/50 (ou mieux encore le 1080p/60 pour les pays 60 Hz) soit 50 (60) images "progressives" de 1080 lignes par seconde. Toutefois, compte tenu de la bande passante nécessaire en production (3 Gigabits/s), ce format est encore hors de portée commerciale. Pour rester raisonnable en termes de bande passante, il faut donc revenir au balayage entrelacé soit le 1080i/50. On constate toutefois que le format 720p/50 (qualifié "EDTV" pour Enhanced Digital TV dans la norme ATSC), plus économique en nombre de ligne (donc moins cher au niveau de l'écran de restitution) et compatible avec le monde de l'informatique, est très apprécié aux États Unis qui, rappelons-le, sortent d'un système de télévision analogique qui ne comporte que 480 lignes actives, le NTSC (que certains plaisantins traduisent par Never Twice The Same Color).

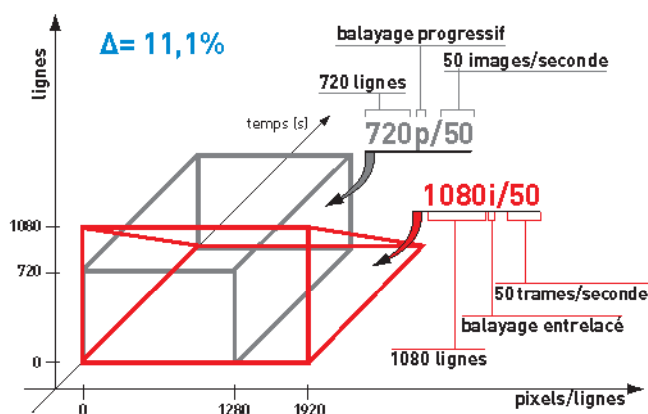
En fait, dans les toutes premières recherches sur la télévision, il a été démontré que, toutes choses égales par ailleurs, le balayage entrelacé donnait de meilleurs résultats que le balayage progressif. Cependant, le débat porte aujourd'hui entre un système entrelacé qui possède plus de lignes alors que son concurrent progressif comporte plus d'images par seconde, les deux formats étant compatibles avec le débit de la liaison HD-SDI à 1,485 Gbps.

Le schéma tridimensionnel ci-dessous permet une comparaison des 2 principaux systèmes en introduisant, au-delà du format proprement dit, la dimension temporelle relative à l'évolution de l'image dans le temps.

La TVHD

dossier technique

11 PAGE



Représentation tridimensionnelle des deux principaux formats de TVHD (l'entrelacement qui ne peut pas être représenté exactement est symbolisé par un demi-volume)

Bien que la question de la qualité du rendu entre ces deux possibilités ne puisse pas être explorée de façon convaincante, compte tenu de la rareté des grands écrans plats conçus pour le balayage entrelacé, il est convenu de considérer que la qualité perçue est équivalente. En effet, entre du 1080i/50 et du 720p/50, la réduction est de 33 % en vertical (et proportionnellement en horizontal) ce qui aboutit à une différence totale faible de 11 %. Le choix du 720p/50 peut être motivé par le compromis qu'il représente entre un format standard et un format de type 1080p/50 qui reste exigeant en bande passante. Le 720p/50 a été ajouté dans la recommandation UIT-R BT.1543 pour compléter le texte qui autorisait déjà le format 720p/60.

Rappelons, en guise de conclusion provisoire, que l'affichage optimal d'une séquence d'images correspond avant tout à une restitution conforme à son mode de capture en respectant aussi bien les spécifications spatiales que les spécifications temporelles de cette dernière.

Le son 5.1

La qualité perceptuelle d'un signal de télévision dépend à la fois de la qualité de la vidéo et de celle de l'audio. Autrement dit, pour une expérience utilisateur optimale, la qualité des images est prépondérante mais la qualité du son est également déterminante.

La TVHD s'entend ainsi de préférence avec un son 5.1. Attention toutefois, la configuration spatiale des enceintes d'un système 5.1 est d'importance puisqu'elle conditionne directement la qualité du rendu et le réalisme des effets sonores.

Plusieurs normes de codage multi-voies existent. Par exemple la norme du format audio MPEG Surround a été ratifiée durant l'été 2006 par l'AES. Elle permet la diffusion de son multi-voies pour seulement quelques kbps additionnels. Le son 5.1 sera ainsi porté à un débit de 64 à 96 kbps, soit un gain d'un facteur 2-3 par rapport à un codage aacPlus traditionnel.

Vocabulaire multi-voies

Le terme "son multicanaux" (ou "multi-voies") désigne l'utilisation de plusieurs pistes audio en vue de la restitution sur un système comportant plusieurs enceintes. Il existe une terminologie associée, constituée de deux chiffres séparés par un point (2.1, 5.1, 6.1, 7.1, etc.). Le premier chiffre indique le nombre de canaux principaux destinés à être restitués sur une enceinte chacun, tandis que le second désigne la présence d'effets basse fréquence destinés à être restitués sur un caisson de basses. Ainsi, 1.0 correspond à un son audio monocanal et 2.0 correspond à une source sonore stéréo. La TVHD est en principe associée à un son 5.1.

Le principe consiste à encoder, en aacPlus par exemple, le cœur du signal mono ou stéréo et d'y accoler les informations complémentaires de multi-voies. Le signal résultant reste donc compatible avec les récepteurs uniquement mono ou stéréo. L'avantage principal de cette technique est ainsi de permettre la diffusion multi-voies pour peu de débit additionnel et en évitant le simulcast du signal mono ou stéréo.

Patrice Bourcet (Son 70)

Bibliographie & références

Extrait d'une fiche Savoir + éditée par la veille technologique de tdf :
"La TVHD, le pourquoi du comment"

