

AVC/H.264

Un système de codage vidéo évolué
pour la HD et SD

Paola Sunna
RAI CRIT

Il faut un débit d'environ 270 Mbit/s pour émettre un signal numérique vidéo non compressé compatible avec la recommandation UIT-R Rec. BT. 601 (télévision en définition standard). La TVHD exige un débit nettement plus élevé. La transmission dans des canaux de radiodiffusion classiques est impossible sans appliquer des techniques de compression évoluées.

Comment fonctionnent les technologies actuelles de codage compatibles avec la transmission TVHD, en particulier le système AVC/H.264 ?

En janvier 1988, le MPEG (*Motion Picture Expert Group*) réunissait quelques experts de l'ISO/CEI pour définir un système normalisé de codage de contenus cinématographiques. *MPEG-1* est ainsi la première norme présentée pour les applications multimédias. Elle cédera ensuite la place à la norme *MPEG-2* pour les applications du domaine de la radiodiffusion.

L'algorithme MPEG-2 compresse le contenu vidéo comme suit :

- **Compression sans perte**

La technique est basée sur l'exploitation de la redondance spatiale/temporelle dans les images (c'est-à-dire la corrélation entre pixels adjacents dans chaque trame et entre des

images ou des trames successives) et sur le codage à longueur variable.

- **Codage psychovisuel**

L'information image invisible à l'œil humain est éliminée.

- **Compression avec perte**

Certaines informations image visibles à l'œil humain (averti) sont éliminées pour accroître le facteur de compression. Elles ne peuvent pas être reconstituées côté récepteur, entraînant dès lors certaines imperfections de l'image.

MPEG-2 a recours à la *transformée en cosinus discrète* (TCD) pour réduire la redondance spatiale. En réalité, elle diminue la corrélation spatiale entre

pixels et contribue à la randomisation de la distribution statistique des échantillons, améliorant ainsi l'efficacité du codage à longueur variable. Les coefficients TCD correspondent rigoureusement au contenu spectral de l'image et, comme la sensibilité de l'œil humain n'est pas uniforme pour tous les angles et fréquences dans le plan horizontal, MPEG-2 réalise le codage psychovisuel à l'aide d'une matrice de quantification appropriée (les coefficients en rapport avec les hautes fréquences et les directions diagonales font l'objet d'une quantification plus sommaire que les autres).

En ce qui concerne la réduction de la redondance temporelle, MPEG-2 améliore l'approche prédictive en utilisant la technique dite de la *compensation de mouvement*.

La norme MPEG-2 est structurée en *profils et niveaux* et, pour chacun d'eux, le débit binaire maximum que le décodeur doit pouvoir traiter est clairement défini.

Une fois normalisé, MPEG-2 a été très largement utilisé (DVB, ATSC, Forum DVD, etc.). À l'instar de la plupart des autres méthodes de codage, MPEG-2 définit uniquement la syntaxe du flux binaire et les caractéristiques du décodeur – laissant aux fabricants le loisir d'appliquer l'algorithme côté codeur. Pour ces raisons, les codeurs MPEG-2 affichent des performances différentes selon le type d'algorithme affecté à la compensation de mouvement, les valeurs de la matrice de quantification et le dispositif de contrôle du débit – tandis que les paramètres tels que la taille des blocs, le codage entropique et le nombre d'images utilisées pour les prédictions ne peuvent pas être modifiés afin de garantir la compatibilité avec les décodeurs MPEG-2.

MPEG-4 partie 2

Le premier successeur de la norme MPEG-2 est *MPEG-4 partie 2* publiée par l'ISO en 1999. Comme dans le cas de MPEG-2, l'efficacité du codage est étroitement liée à la complexité du matériel de source et à la mise en œuvre du codeur. MPEG-4 a été

définie pour des applications multimédias à faibles débits binaires, puis étendue aux applications du domaine de la radiodiffusion. Une évaluation officielle subjective indique¹ que le codage MPEG-4 partie 2 offre un gain d'efficacité de 15 à 20 % par rapport à MPEG-2. Pour les applications DVB, ce gain n'est pas suffisant pour justifier une déstabilisation du monde MPEG-2 – en raison de l'incompatibilité de MPEG-4 partie 2 avec MPEG-2.

AVC/H.264

En 2001, soucieuses de mettre au point un système de compression plus efficace, les instances de normalisation ISO/CIE (MPEG) et l'IUT conjuguent leurs efforts au sein du groupe de travail JVT (*Joint Video Team*) chargé de développer le système de codage AVC (*Advanced Video Coding*). En 2003, le système AVC est intégré en tant que *partie 10* à la norme *MPEG-4* (ISO/IEC 14496-10) et repris sous l'appellation *H.264* à l'UIT.

En septembre 2004, le consortium DVB modifie la norme TS 101 154² (*Implementation guidelines for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*) afin d'y inclure également *AVC/H.264*.

Profils disponibles

Le système *AVC/H.264* ne produit pas de flux binaire conforme à MPEG-2, aussi son adoption exigera-t-elle de nouveaux codeurs et décodeurs. Le schéma AVC inclut différents profils :

- **Profil de base** – Pour les applications bout en bout à faible délai.
- **Profil X** – Pour les applications mobiles et e-streaming.

- **Profil principal** – Pour les applications de radiodiffusion définition standard (SD).

- **Profil supérieur** – La norme H.264/AVC porte principalement sur la vidéo de « qualité divertissement » basée sur 8 bits/échantillon et un échantillonnage de la chrominance 4:2:0. Pour répondre aux besoins des applications les plus exigeantes – conçues par exemple pour la contribution et la distribution de contenus, le montage studio et la postproduction – le projet conjoint a été étendu en vue d'augmenter les capacités de la norme standard. Au départ appelées *professional extensions*, elles ont été rebaptisées *fidelity range extensions* ou *FRExt*. Lors de sa mise au point, le projet *FRExt* a abouti à quatre nouveaux profils appelés collectivement *profils supérieurs* :

- **Profil supérieur (HP)** – Prend en charge la vidéo 8 bits et l'échantillonnage 4:2:0, et répond aux besoins du consommateur utilisant du matériel haut de gamme et des applications performantes qui exigent une vidéo haute résolution sans devoir utiliser des formats de chrominance ou une précision d'échantillonnage étendus.
- **Profil supérieur 10 (Hi10P)** – Prend en charge la vidéo 4:2:0, jusqu'à 10 bits par échantillon.
- **Profil supérieur 4:2:2 (H422P)** – Prend en charge l'échantillonnage de la chrominance jusqu'à 4:2:2 et 10 bits par échantillon.
- **Profil supérieur 4:4:4 (H444P)** – Prend en charge l'échantillonnage de la chrominance jusqu'à 4:4:4 et 12 bits par échantillon. En outre, il permet un codage zonal sans pertes ainsi qu'une transformée des couleurs résiduelles pour le codage de vidéo RVB évitant les erreurs de transformation de l'espace couleur.

Ces profils prennent tous en charge les fonctions du profil *principal* et ainsi qu'une taille de bloc pour transformées adaptatives et des matrices adaptées à la quantification perceptuelle.

1. Ken McCann : DVB+MPEG-4=? Présentation PowerPoint à DVB World 2002.
2. http://webapp.etsi.org/workprogram/Report_WorkItem.asp?WKI_ID=21401

Tableau 1 – Efficacité/complexité de la norme AVC/H.264 par rapport à la norme MPEG-2
(Source : www.m4if.org)

Profil	Efficacité par rapport à MPEG-2	Accroissement de la complexité du décodeur
Base	Environ 1,5 fois	Environ 2,5 fois
X	Environ 1,75 fois	Environ 3,5 fois
Principal	Environ 2 fois	Environ 4 fois

Le profil *supérieur* sera selon toute vraisemblance intégré aux spécifications de plusieurs grandes applications à court terme, notamment :

- la spécification vidéo *HD-DVD* du Forum DVD ;
- la spécification vidéo *BD-ROM* de la Blu-ray Disc Association ;
- les normes *DVB* pour la radiodiffusion en Europe.

Le profil supérieur pourrait se substituer rapidement au profil principal dans l'industrie. En effet, il offre une meilleure efficacité de codage par rapport à celle du profil principal, sans trop accroître la complexité de mise en œuvre.

Efficacité du codage

Selon le rapport des essais de vérification de AVC/H.264³, l'efficacité du codage est nettement supérieure à celle de MPEG-2. Les principaux avantages de AVC/H.264 sont les suivants :

1) Compensation de mouvement

AVC/H.264 utilise des blocs de dimension et de forme variables, MPEG-2 utilise des blocs fixes 16 x 16. Il est donc possible d'enregistrer un gain d'efficacité jusqu'à 15%.

AVC/H.264 permet une estimation du vecteur de mouvement de l'ordre

du quart de pixel, MPEG-2 du demi-pixel (gain d'efficacité jusqu'à 20%).

AVC/H.264 utilise jusqu'à cinq trames pour l'estimation du mouvement, deux trames pour MPEG-2 pour les images interpolées avec un gain du débit binaire situé entre 5 et 10%.

2) Réduction de la redondance spatiale

AVC/H.264 utilise une *transformée entière* (au lieu de la transformée en cosinus discrète du MPEG-2), ce qui réduit ainsi l'impact des erreurs d'arrondi.

3) Quantification

AVC/H.264 intègre un nombre plus élevé de niveaux de quantification : 52 contre 31 pour MPEG-2.

4) Codage entropique

AVC/H.264 utilise des techniques de codage plus complexes qui sont plus efficaces que le codage à longueur variable du MPEG-2.

5) Filtrage des blocs

AVC/H.264 utilise un filtre adaptatif pour réduire la pixellisation qui peut sérieusement détériorer la qualité de l'image finale MPEG-2.

L'efficacité supérieure de la norme AVC/H.264 – une réduction du débit binaire obtenue en maintenant la même qualité d'image subjective – a un prix : une complexité accrue tant au niveau du codeur que du décodeur.

Le *tableau 1* comporte quelques données relatives à la complexité du *décodeur*.

Le *codeur* AVC/H.264 est huit fois plus complexe que le codeur MPEG-2. Heureusement, la puissance des circuits intégrés a augmenté d'un facteur 100 depuis le début de la compression MPEG-2, et des processeurs codeur/décodeur AVC en temps réel sont actuellement en phase de développement.

Principaux concurrents

Ces dernières années, de nombreuses technologies propriétaires prometteuses (Windows Media, Real Video, On2, Sorenson, etc.) ont vu le jour afin de distribuer des contenus multimédias sur des canaux de transmission à bande étroite.

Ces systèmes de compression ont tout d'abord été utilisés pour des flux vidéo à faible débit et des images de plus petite taille (SQCIF, QCIF et CIF). Toutefois, leur vocation a été récemment élargie pour garantir une efficacité du codage sur un éventail de débits binaires – allant de quelques kbit/s à des dizaines de Mbit/s – afin de prendre en charge les télévisions standard et haute définition.

Windows Media 9 / VC-1

Windows Media 9 de Microsoft était un système propriétaire jusqu'en mars 2004 lorsque la société de Redmond a présenté à la SMPTE un document intitulé « *Proposed SMPTE Standard for Television: VC-9 compressed video Bitstream format and decoding Process* » afin de normaliser la technologie Windows Media 9. La norme a été

3. ISO/IEC JTC1/SC29/WG11 MPEG2003/N6231, décembre 2003, Waikoloa, Hawaii, USA.

Tableau 2 – Principales caractéristiques des normes MPEG-2, H.264 et VC-1

	MPEG-2	AVC/H.264	VC-1
Profils	Simple Principal Supérieur 4:2:2	Base Principal X FRExt	Simple Principal Évolué
Source	Entrelacée Progressive	Entrelacée Progressive	Entrelacée Progressive
Format de chrominance	4:2:0, 4:2:2	4:2:0, 4:2:2, 4:4:4	4:2:0
Type d'algorithme	Hybride (redondance spatiale/ temporelle)	Hybride (redondance spatiale/ temporelle)	Hybride (redondance spatiale/ temporelle)
Structure de codage	Hiérarchique Profiles@Levels	Hiérarchique Profiles@Levels	Hiérarchique Profiles@Levels
Type d'image	I, P, B	I, P, B, SP, SI	I, P, B, BI
Structure du macrobloc	16x16 Y 8x8 C _b 8x8 C _r	16x16 Y 8x8 C _b 8x8 C _r	16x16 Y 8x8 C _b 8x8 C _r
Précision de l'estimation de mouvement	Jusqu'à ½ pixel	Jusqu'à ¼ de pixel	Jusqu'à ¼ de pixel
Filtre en boucle pour réduire la pixellisation	Non	Oui	Oui
Type de transformée	Virgule flottante (TCD)	Entière	Entière
Codage entropique	VLC	CAVLC CABAC	VLC

appelée VC-1 et Windows Media 9 est aujourd'hui une application logicielle de VC-1.

Il existe trois profils dans la recommandation VC-1 :

- **Simple** – Il vise les flux Internet bas débit et les applications peu complexes telles que les télécommunications mobiles ou la lecture sur ordinateurs de poche.
- **Principal** – Il vise les applications Internet haut débit telles que la diffusion en continu, la transmission de contenus cinématographiques via IP, ou la télévision/vidéo à la carte via IP.
- **Évolué** – Il vise les applications radiodiffusion telle que la télévision numérique, la lecture de DVD HD sur ordinateur et la TVHD. C'est le

seul profil VC-1 qui prend en charge les contenus *entrelacés*. En outre, il contient les éléments syntaxiques requis pour transmettre ses propres flux binaires vidéo dans des systèmes génériques tels MPEG-2 Transport ou Programme Streams (ISO/IEC 13818-2).

La question de la licence reste ouverte pour le VC-1 parce que le MPEG-LA tente toujours de déterminer s'il existe des détenteurs de brevets et, dans l'affirmative, quel en est leur nombre.

Dirac

Dirac (voir page 59) est un autre codec vidéo qui tente de concurrencer les codecs performants analysés dans cet article, mais son développement n'en

est qu'à ses débuts. Il utilise les ondelettes, la compensation de mouvement et le codage arithmétique. Il a été mis sur pied par le département R&D de la BBC, non pas comme produit mais comme outil de recherche. Une version expérimentale du code, rédigé en C++, a été publiée en mars 2003, dans le cadre d'un contrat de licence de logiciel à source ouverte, et peut être téléchargée à <http://sourceforge.net/projects/dirac>.

Comparaison

Le tableau 2 compare les caractéristiques des normes MPEG-2, AVC/H.264 et VC-1. Les normes AVC/H.264 et VC-1 semblent très semblables mais, à l'heure où cet article était sous presse (février 2005), aucune donnée offi-

cielle n'était disponible sur l'efficacité et la complexité de la compression – même si Microsoft prétend que la performance de la norme VC-1 est comparable à celle de la norme AVC/H.264, tout en étant moins complexe à mettre en œuvre.

Conclusions

La norme MPEG-2, utilisée dans les domaines de la télévision numérique et des applications de stockage depuis plus de dix ans, doit faire face à la concurrence de nouveaux systèmes de compression vidéo efficaces. *AVC/H.264* est issue des derniers développements en matière de norme de codage vidéo et affiche une performance deux fois plus élevée que MPEG-2.

Microsoft obtient une performance similaire avec Windows Media 9 qui est en passe d'être adoptée par la SMPTE comme la norme de codage vidéo *VC-1*. Toutefois, d'autres systèmes de compression voient le jour – et visent aussi à jouer un rôle important dans le développement de nouvelles applications et de nouvelles opportunités commerciales, aussi bien dans le domaine de la radiodiffusion qu'en dehors (TVHD, diffusion vidéo en direct, télévision large bande, visioconférence, applications médicales et scientifiques, etc.).

Enfin, l'efficacité du codage n'est pas le seul facteur à influencer le choix du codec pour remplacer MPEG-2. Il reste d'importantes questions relatives aux droits de propriété intellectuelle avant de pouvoir mettre la dernière main à de solides plans commerciaux, basés sur cette nouvelle génération de codecs plus efficaces.



Paola Sunna est diplômée en électronique du *Politecnico* de Turin (1997). Sa thèse portait sur l'évaluation objective de la qualité vidéo du codage MPEG-2. Elle travaille depuis la fin de ses études au CRIT, le centre de recherches et d'innovations techniques de la RAI. Ses activités portent plus particulièrement sur les essais de laboratoire des méthodes d'évaluation de la qualité vidéo pour les applications de radiodiffusion, de diffusion web et 3G.

Elle préside depuis 2002 le groupe de projet B/VIM (Vidéo dans le multimédia) de l'UER qui a établi la nouvelle méthode subjective d'évaluation de la qualité multimédia, SAMVIQ.