

RECOMMANDATION UIT-R BT.1687-1

**Réduction du débit binaire vidéo pour la distribution* en temps réel
de programmes numériques sur grand écran destinés
à être présentés en salle**

(Question UIT-R 15/6)

(2004-2006)

Domaine d'application

La présente Recommandation porte sur la réduction du débit binaire vidéo pour la distribution en temps réel de programmes numériques sur grand écran (LSDI, *large-screen digital imagery*) destinés à être présentés en salle et donne des exemples de débits binaires pour des signaux compressés MPEG-2 et MPEG-4/AVC utilisés dans des applications LSDI. Les systèmes d'imagerie numérique pour la distribution de programmes LSDI en bande de base sont définis dans la Recommandation UIT-R BT.1680.

L'Assemblée des radiocommunications de l'UIT,

considérant

- a) qu'il est souhaitable que la distribution et la présentation de programmes numériques sur grand écran (LSDI) respectent le contenu créatif et la qualité d'image du programme de distribution¹, dans toute la mesure permise par les conditions de présentation du programme considéré;
- b) que la Recommandation UIT-R BT.1680 préconise d'utiliser les systèmes d'imagerie numérique respectant les spécifications applicables aux programmes LSDI destinés à être présentés sur grand écran, et devrait être utilisée pour la distribution de ces applications;
- c) qu'il est nécessaire, dans le cadre de la distribution des programmes LSDI, de réduire le débit binaire des signaux de programme, afin de maintenir la valeur du débit binaire à une valeur acceptable et donc diminuer le temps et le coût de transmission, sans toutefois détériorer la qualité de programme telle que perçue par l'observateur;
- d) que la résolution du système visuel humain est beaucoup plus grande pour les détails de luminance que pour les détails de couleur;
- e) que l'ISO/CEI JTC 1/SC 29/WG 11 a réalisé et présenté des études approfondies relatives au débit binaire requis pour assurer une compression vidéo virtuellement transparente² à l'aide de divers algorithmes de compression et différents systèmes d'imagerie³;

* L'UIT définit le terme «distribution» comme suit: «Acheminement de programmes de télévision lorsque plus aucun traitement de post-production n'est prévu».

¹ Le terme «programme de distribution» est utilisé ici pour désigner le programme obtenu après adaptation (en contenu et en qualité) du programme d'origine au média de distribution.

² L'UIT définit la «réduction transparente de débit» comme un «processus de réduction de débit qui n'affecte pas la qualité subjective des séquences audio ou vidéo».

³ Voir par exemple le Document ISO/CEI JTC 1/SC 29/WG 11 – MPEG2003/N6231, décembre 2003, Waikoloa – Report of The Formal Verification Tests on AVC (ISO/CEI 14496-10 – UIT-T Rec. H.264).

- f) que ces tests ont confirmé que le codage MPEG-2 de combinaison MP@HL à l'aide d'un codeur optimisé⁴ à 20 Mbit/s est virtuellement transparent vis-à-vis de la qualité du signal source pour les systèmes d'imagerie indiqués dans la Recommandation UIT-R BT.1680 qui présentent une fréquence d'image de 25 ou 30 Hz et un balayage progressif ou entrelacé;
- g) que la compression vidéo MPEG-2 est aujourd'hui largement disponible et utilisée et qu'elle est conforme à la politique de l'UIT en matière de propriété industrielle;
- h) que tout récemment, un nouveau système de compression vidéo établi d'après la norme MPEG-2 et dénommé MPEG-4/AVC (Recommandation UIT-T H.264) a été mis au point et qu'il offre une efficacité de compression environ deux fois supérieure à celle de la norme MPEG-2, même au stade des applications initiales de ses codecs;
- j) que l'on a élaboré, tout récemment, des extensions de la première version du système de codage source MPEG-4/AVC, appelées *Fidelity Range Extensions* (FRExt) (Extensions de profil), qui présentent une efficacité de codage encore meilleure avec le codage des systèmes d'images 1080/1920p à 50 et à 60 images/s;
- k) que l'utilisation du système de codage source MPEG-4/AVC peut se révéler être une option intéressante pour les applications LSDI, en raison du débit binaire net beaucoup plus faible qui est nécessaire pour distribuer des programmes LSDI, bien que les codecs MPEG-4/AVC soient plus complexes et qu'ils devraient par conséquent être un peu plus onéreux;
- l) que la norme MPEG-4/AVC offre une grande souplesse pour le choix des paramètres de codage, tels que l'échantillonnage 4:4:4, 10 ou 12 bits par échantillon et des fréquences d'image supérieures à 72 Hz;
- m) que plusieurs pays mettent actuellement en œuvre des services pour la présentation de programmes LSDI destinés à être projetés en salle ou prévoient de le faire dans un avenir proche,

recommande

- 1 d'utiliser, pour les premières mises en œuvre des programmes LSDI destinés à être présentés en salle, un signal vidéo numérique en bande de base conforme à la Recommandation UIT-R BT.1680 au niveau des interfaces d'entrée et de sortie de la chaîne de distribution LSDI;
- 2 d'utiliser, étant donné que la résolution du système visuel humain est plus grande pour les détails de luminance que pour les détails de couleur, un codage vidéo 4:2:0 ou, de préférence, un codage vidéo 4:2:2 pour la distribution de programmes LSDI destinés à être présentés en salle;
- 3 d'utiliser de préférence, étant donné que le signal de programme LSDI ne fait généralement pas l'objet d'un post-traitement d'image de type créatif après sa distribution, une réduction binaire intertrames pour la distribution de programmes LSDI destinés à être représentés en salle, puisque celle-ci assure une efficacité de compression supérieure à celle du codage intratrame; il est reconnu que, lorsqu'une insertion de programme local (liste de décisions de montage ou liste de diffusion) est requise, le codage intratrame est préférable et qu'en pareil cas, la valeur minimale du débit binaire vidéo peut être plus élevée;

⁴ Le terme «codeur optimisé» est utilisé ici pour désigner un codeur MPEG-2 installé au niveau de l'équipement et offrant des solutions évoluées, c'est-à-dire optimisées vis-à-vis du codeur «modèle de référence» MPEG-2 TM5, mis en œuvre au niveau logiciel et adopté par l'ISO/CEI JTC 1/SC 29/WG 11 comme référence de mesure de la performance.

4 d'envisager d'utiliser dans un proche avenir, étant donné que la réduction du débit binaire utilisée pour la voie de distribution devrait être virtuellement transparente vis-à-vis de la qualité du programme de distribution dans les conditions de présentation prévues, la méthode de réduction du débit binaire intertrame MPEG-2 avec la combinaison MP@HL (HiQ) et une valeur nette minimale du débit vidéo de l'ordre de 20 Mbit/s, pour la distribution en temps réel de programmes LSDI destinés à être présentés en salle (l'utilisation du codage 4:2:2 @ HL nécessitera un débit binaire légèrement plus élevé);

5 d'envisager également l'application de la méthode de réduction du débit binaire MPEG-4/AVC comme une autre solution intéressante, en particulier lorsqu'il est souhaitable d'utiliser un débit binaire net beaucoup plus faible pour la distribution de programmes LSDI et que le recours à des codecs un peu plus complexes est acceptable;

NOTE 1 – La Recommandation UIT-T H.264 est disponible en version électronique à l'adresse suivante: <http://www.itu.int/md/R03-SG06-C-0211/fr>.

6 d'analyser, en temps voulu, la possibilité d'utiliser des codecs encore plus efficaces en termes de compression et susceptibles d'apparaître dans l'avenir, lorsque ces dispositifs seront conformes à la politique de l'UIT en matière de propriété industrielle et lorsqu'ils auront été testés de manière approfondie, qu'ils seront disponibles à grande échelle et qu'ils seront de préférence interopérables avec les codecs MPEG-2 et MPEG-4/AVC.

Appendice 1 (Informatif)

Codage source de divers systèmes d'imagerie LSDI définis dans la Recommandation UIT-R BT.709 sur la base de la norme MPEG-4/AVC (Recommandation UIT-T H.264): Exemples de paramètres et d'outils de base

Le présent Appendice donne des exemples de paramètres et d'outils de base utilisés dans la méthode de codage source avec compression MPEG-4/AVC (Recommandation UIT-T H.264) applicable à divers systèmes d'imagerie LSDI définis dans la Recommandation UIT-R BT.709-5 (Partie 2). Il indique également les débits binaires requis pour l'acheminement quasi imperceptible des signaux ainsi codés à la source.

Membre de la famille de la Rec. UIT-R BT.709	Paramètres et outils de base MPEG-4/AVC	Débit binaire pour un transport quasi imperceptible
1 920 × 1 080 × 24/25p	<p style="text-align: center;">Niveau 4</p> <p>Outils de codage Supérieur 10</p> <p>Outils du profil principal X</p> <p>Format de chrominance 4:2:0 X</p> <p>Profondeur d'échantillonnage 8 bits X</p> <p>Transformée 4 × 4/ transformée 8 × 8 X</p> <p>Matrices d'échelle de quantification X</p> <p>Commande séparée des paramètres de quantification de C_B et de C_R X</p> <p>Format vidéo monochrome X</p> <p>Profondeur d'échantillonnage 9 et 10 bits X</p> <p>Format de chrominance 4:2:2</p> <p>Profondeur d'échantillonnage 11 et 12 bits</p> <p>Format de chrominance 4:4:4</p> <p>Transformée du signal de couleur résiduel</p> <p>Codage prédictif sans perte</p>	<p>9-11 Mbit/s – Qualité contribution</p> <p>7-10 Mbit/s – Qualité distribution</p>
1 920 × 1 080 × 60/50i	<p style="text-align: center;">Niveau 4</p> <p>Outils de codage Supérieur 4:2:2</p> <p>Outils du profil principal X</p> <p>Format de chrominance 4:2:0 X</p> <p>Profondeur d'échantillonnage 8 bits X</p> <p>Transformée 4 × 4/ transformée 8 × 8 X</p> <p>Matrices d'échelle de quantification X</p> <p>Commande séparée des paramètres de quantification de C_B et de C_R X</p> <p>Format vidéo monochrome X</p> <p>Profondeur d'échantillonnage 9 et 10 bits X</p> <p>Format de chrominance 4:2:2</p> <p>Profondeur d'échantillonnage 11 et 12 bits</p> <p>Format de chrominance 4:4:4</p> <p>Transformée du signal de couleur résiduel</p> <p>Codage prédictif sans perte</p>	<p>10-15 Mbit/s – Qualité contribution</p> <p>8-12 Mbit/s – Qualité distribution</p>

Membre de la famille de la Rec. UIT-R BT.709	Paramètres et outils de base MPEG-4/AVC	Débit binaire pour un transport quasi imperceptible
1 920 × 1 080 × 60/50p	<p style="text-align: center;">Niveau 4.2</p> <p>Outils de codage Supérieur 4:2:2</p> <p>Outils du profil principal X</p> <p>Format de chrominance 4:2:0 X</p> <p>Profondeur d'échantillonnage 8 bits X</p> <p>Transformée 4 × 4/ transformée 8 × 8 X</p> <p>Matrices d'échelle de quantification X</p> <p>Commande séparée des paramètres de quantification de C_B et de C_R X</p> <p>Format vidéo monochrome X</p> <p>Profondeur d'échantillonnage 9 et 10 bits X</p> <p>Format de chrominance 4:2:2</p> <p>Profondeur d'échantillonnage 11 et 12 bits</p> <p>Format de chrominance 4:4:4</p> <p>Transformée du signal de couleur résiduel</p> <p>Codage prédictif sans perte</p>	<p>18-20 Mbit/s – Format d'émission</p> <p>NOTE – Il se peut que l'échantillonnage sur 10 bits n'augmente pas le débit binaire par rapport à l'échantillonnage sur 8 bits.</p>
1 920 × 1 080 × 24/25p	<p style="text-align: center;">Niveau 4</p> <p>Outils de codage Supérieur 4:2:2</p> <p>Outils du profil principal X</p> <p>Format de chrominance 4:2:0 X</p> <p>Profondeur d'échantillonnage 8 bits X</p> <p>Transformée 4 × 4/ transformée 8 × 8 X</p> <p>Matrices d'échelle de quantification X</p> <p>Commande séparée des paramètres de quantification de C_B et de C_R X</p> <p>Format vidéo monochrome X</p> <p>Profondeur d'échantillonnage 9 et 10 bits X</p> <p>Format de chrominance 4:2:2</p> <p>Profondeur d'échantillonnage 11 et 12 bits</p> <p>Format de chrominance 4:4:4</p> <p>Transformée du signal de couleur résiduel</p> <p>Codage prédictif sans perte</p>	<p>8-10 Mbit/s – Mode «cinéma» haute qualité</p>

Pièce jointe 1

Informations sur la méthode de codage source du MPEG-4/AVC [Sullivan et autres, 2004]

La Recommandation UIT-T H.264/MPEG-4 (Partie 10) – Codage vidéo évolué pour les services audiovisuels génériques, communément appelée H.264/AVC, est la dernière en date de la série des normes internationales relatives au codage vidéo. Il s'agit actuellement de la version la plus aboutie et la plus performante de cette norme, qui a été élaborée par une Equipe mixte sur la vidéo (JVT, *Joint Video Team*) composée d'experts du groupe d'experts du codage vidéo (VCEG, *Video Coding Experts Group*) de l'UIT-T et du Groupe d'experts pour les images animées (MPEG, *Moving Picture Experts Group*) de l'ISO/CEI.

Comme pour les normes antérieures, cette version représente actuellement le meilleur compromis possible entre l'efficacité du codage, la complexité de la mise en œuvre et les coûts, compte tenu de l'état actuel des techniques de conception des systèmes VLSI (unité centrale de traitement (CPU), processeur de signaux numériques (DSP), circuit intégré à applications spécifiques (ASIC), réseau prédéfini programmable (FPGA), etc.).

La norme ainsi créée permet un gain d'efficacité du codage d'un facteur au moins égal à deux (en moyenne) par rapport à la norme MPEG-2, et ce à un coût raisonnable.

En juillet 2004, cette norme a fait l'objet d'une nouvelle modification, connue sous le nom de *Fidelity Range Extensions* (extensions de profil (FRExt), Amendement 1), qui offre une efficacité de codage encore supérieure à celle de la norme MPEG-2, pouvant atteindre pas moins de 3:1 pour certaines applications essentielles.

Même si elle avait un grand nombre d'applications, la version initiale de la norme H.264/AVC (telle qu'elle a été achevée en mai 2003) portait essentiellement sur la vidéo de «qualité grand public» (échantillonnage sur 8 bits et échantillonnage de la chrominance 4:2:0). En raison des délais impartis pour son élaboration, cette norme ne pouvait pas répondre aux besoins des applications professionnelles les plus exigeantes et n'était donc pas adaptée aux très hautes résolutions vidéo. En effet, pour les applications de contribution et de distribution de programmes, le montage studio et la postproduction, il est parfois nécessaire:

- d'échantillonner le signal vidéo source sur plus de 8 bits pour une précision accrue;
- d'utiliser, pour la représentation en couleur, une plus haute résolution que celle qui est généralement utilisée dans les applications grand public (c'est-à-dire un échantillonnage 4:2:2 ou 4:4:4 au lieu d'un format d'échantillonnage de chrominance 4:2:0);
- de mettre en œuvre des fonctions de montage à la source, par exemple le mélange alpha (opération consistant à mélanger plusieurs scènes vidéo essentiellement pour les bulletins d'information météorologique, en combinant la vidéo d'un présentateur à celle d'une carte ou à une image de radars météorologiques);
- d'utiliser des débits binaires très élevés;
- d'utiliser une très haute résolution;
- d'offrir une très haute fidélité, même en représentant sans perte certaines parties de la vidéo;
- d'éviter les erreurs d'arrondi lors de la transformation de l'espace chromatique;
- d'utiliser une représentation couleur rouge-vert-bleu (RVB).

Le projet FRExt a abouti à une série de quatre nouveaux profils, appelés collectivement profils supérieurs:

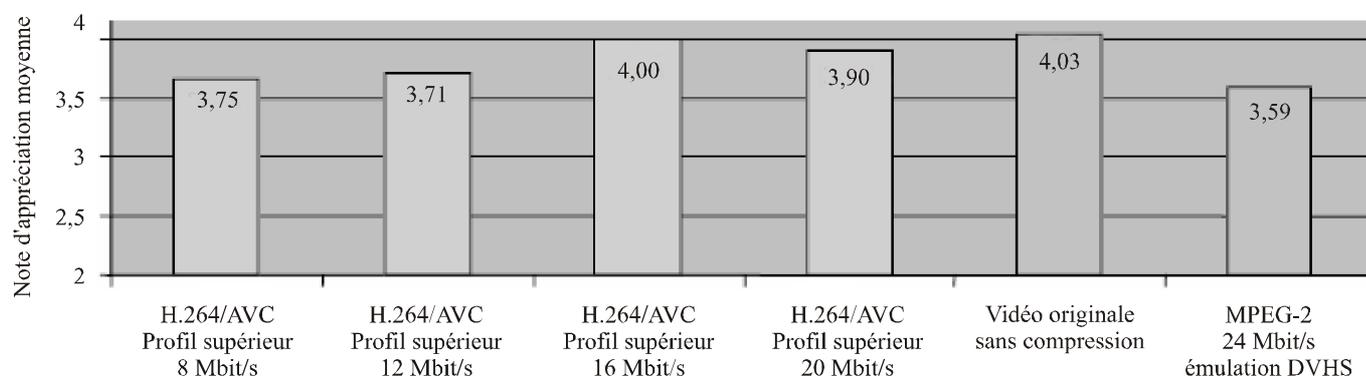
- Le profil supérieur (HP) prend en charge la vidéo à 8 bits avec un échantillonnage 4:2:0 et répond aux besoins des consommateurs utilisant du matériel haute gamme ainsi qu'à une des applications nécessitant une vidéo haute résolution, sans qu'il soit nécessaire d'utiliser des formats de chrominance ou une précision d'échantillonnage étendus.
- Le profil supérieur 10 (Hi10P) prend en charge la vidéo 4:2:0 jusqu'à 10 bits de précision de représentation par échantillon.
- Le profil supérieur 4:2:2 (H422P) prend en charge l'échantillonnage de la chrominance 4:2:2 jusqu'à 10 bits par échantillon.
- Le profil supérieur 4:4:4 (H444P) prend en charge l'échantillonnage de la chrominance 4:4:4 jusqu'à 12 bits par échantillon. En outre, ce profil permet un codage sans perte efficace ainsi qu'une transformée entière du signal couleur résiduel pour le codage vidéo RVB tout en évitant les erreurs de transformation dans l'espace chromatique.

Etant donné que la spécification FRExt est encore relativement nouvelle, et que certains de ces avantages sont plus subjectifs qu'objectifs, il est un peu plus difficile d'en évaluer les capacités. A cet égard, il est intéressant de rendre compte du résultat d'une évaluation subjective de la qualité qui a été menée à bien par l'Association Blu-ray Disc (BDA). Les résultats résumés sur la Fig. 1 sont extraits du rapport d'essai [Wedi et Kashiwagi, 2004]. Les essais, effectués sur un film de 24 images/s, avec balayage progressif 1 920 × 1 080, ont donné les résultats nominaux suivants (résultats qui ne sauraient être considérés comme rigoureusement exacts du point de vue statistique):

- Le profil supérieur de la spécification FRExt a permis d'obtenir une meilleure qualité vidéo théorique que la norme MPEG-2 avec à peine un tiers des bits (8 Mbit/s au lieu de 24 Mbit/s).
- Le profil supérieur de la spécification FRExt a permis d'obtenir une qualité vidéo théoriquement imperceptible (c'est-à-dire difficile à distinguer de l'original n'ayant pas fait l'objet d'une compression) à 16 Mbit/s seulement.

FIGURE 1

Comparaison entre les normes MPEG-2 et H.264

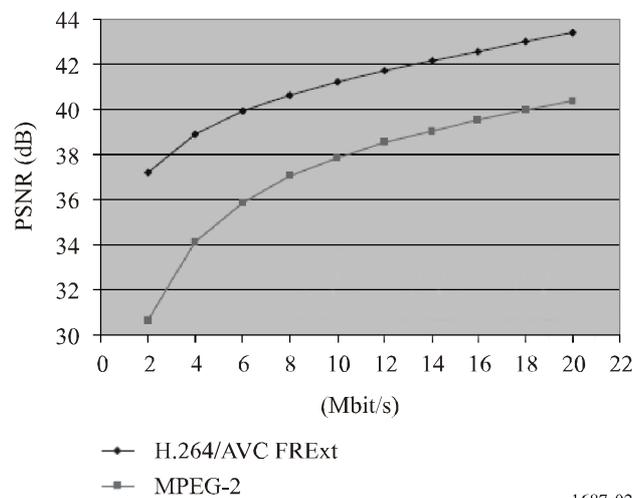


Le niveau de qualité (3.0), considéré comme suffisant pour les systèmes haute définition dans cette organisation, a été nettement dépassé avec un débit d'à peine 8 Mbit/s. Là encore, on a constaté que la méthode de codage H.264/AVC employée lors de ces essais n'était pas optimale. En conséquence, il sera sans doute possible de ramener le débit binaire nettement au-dessous de 8 Mbit/s, tout en restant au-dessus du niveau de qualité 3.0, avec une qualité suffisante comparable à une «qualité haute définition» acceptable pour cette application exigeante.

Le résultat d'un test de comparaison objective (valeurs de crête du rapport signal/bruit, PSNR, *peak signal-to-noise ratio*) effectué par la société FastVDO⁵ est représenté sur la Fig. 2. Ces résultats objectifs confirment l'excellente qualité du profil supérieur. (Là encore, les utilisations sous-optimales des images B font que la qualité indiquée sur le graphique est en deçà de la réalité concernant le système FRExt.)

FIGURE 2

Comparaison du PSNR



1687-02

Références bibliographiques

- SULLIVAN, G.J., TOPIWALA, P. et LUTHRA, A. [2004] Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions. Document présenté à la Conférence de la SPIE sur les applications du traitement d'images numériques (XXVII). Séance spéciale consacrée aux progrès réalisés en ce qui concerne la nouvelle norme H.264/AVC.
- WEDI, T. et KASHIWAGI, Y. [2004] Subjective quality evaluation of H.264/AVC FRExt for HD movie content. Document JVT-L033 de la JVT.

⁵ FastVDO est une société spécialisée dans les techniques de communication média et les logiciels d'infrastructure qui a son siège à Columbia, MD, Etats-Unis d'Amérique.