

TÉLÉVISION STÉRÉOSCOPIQUE – PROFIL MULTIVUES MPEG-2

(1998)

1 Description générale du profil multivues (MVP, *multi-view profile*) MPEG-2

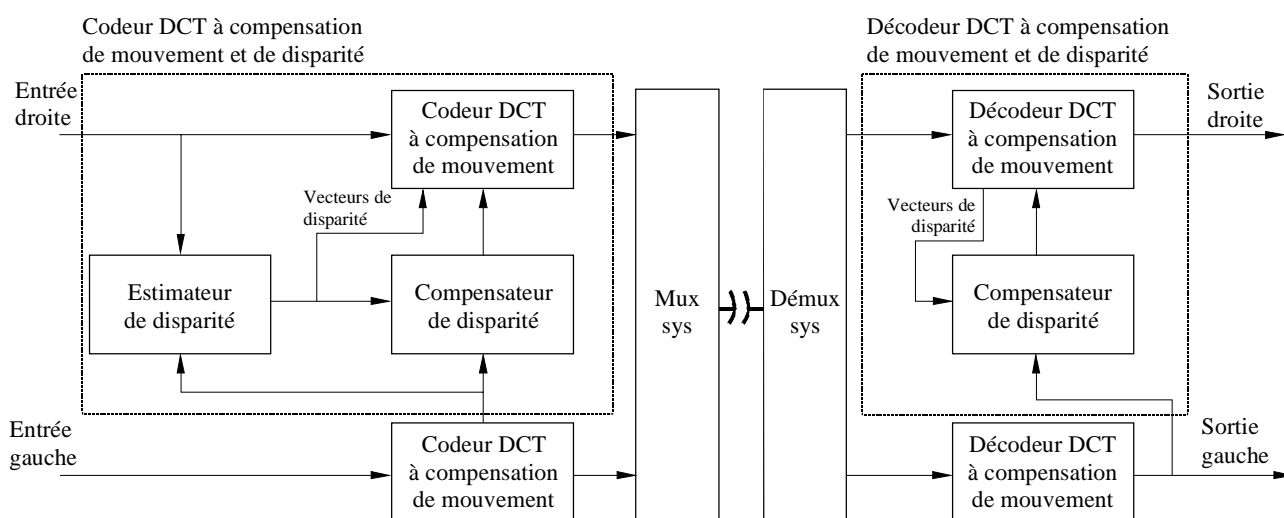
L'extension de la norme vidéo MPEG-2 (Recommandation UIT-T H.262 | ISO/CEI 13818-2: Technologies de l'information – Codage générique des images animées et du son associé: données vidéo) aux applications multivues (par exemple, pour la vidéostéréoscopie) est devenue une norme internationale définitive à la réunion ISO/CEI JTC 1/SC 29/GT 11 de septembre 1996 (Amendement 3, GT 11 N1366). Le profil multivues est conçu comme un profil se prêtant aux applications exigeant des points de vue multiples dans le contexte de la Norme vidéo MPEG-2. Le MVP gère des images stéréoscopiques en tant qu'images source pour toute une gamme de résolutions et de qualités d'image en fonction des applications.

1.1 Système de codage

Le modèle de référence de codec pour le MVP est illustré par le diagramme de la Fig. 1. Ses caractéristiques principales sont d'une part le codage monoscopique dans la couche de base pour assurer la compatibilité, d'autre part la prédiction hybride du mouvement et de la disparité pour assurer l'efficacité de la compression. Des outils d'échelonnement temporel sont utilisés pour coder une couche d'amélioration.

FIGURE 1

Modèle de référence de codec pour le MVP



DCT: transformée en cosinus discrète (*discrete cosine transform*)

Rap 2017-01

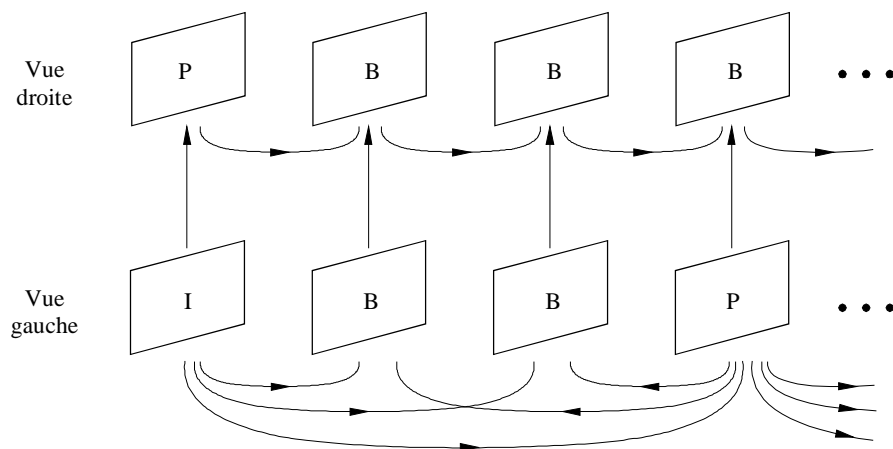
La configuration des modes de prédiction est représentée à la Fig. 2. On applique à la couche de base un codage monoscopique avec les mêmes outils que le profil principal (MP, *main profile*), notamment la Norme ISO/CEI 11172-2. Une couche de base de MVP est affectée à la vision de gauche et une couche d'amélioration est affectée à la vision de droite. On code une couche d'amélioration à l'aide d'outils d'échelonnement temporel et l'on peut utiliser une prédiction hybride du mouvement et de la disparité dans la couche améliorée. On assure une plus forte compression de la vision de droite en vidéostéréoscopie en exploitant la similitude entre la gauche et la droite.

Le MVP, l'un des profils échelonnables en couches à points de vue multiples, présente le même type de caractéristiques de compatibilité; d'autres profils échelonnables présentent cette compatibilité avec le profil MP. Par exemple:

- les décodeurs compatibles avec le MVP à un certain niveau sont capables de décoder des flux binaires compatibles avec le MP au niveau correspondant (compatibilité aval),
- les décodeurs compatibles avec le MP à un certain niveau sont capables de décoder le flux binaire dans la couche de base du MVP (compatibilité amont).

FIGURE 2

Exemple de configuration à prédiction avec codage $M = 3$ d'image bi-trame gauche et droite, codée avec prédiction de disparité par rapport à l'image gauche et avec auto-prédiction du mouvement



Rap 2017-02

1.2 Valeur des paramètres

Le MVP présente un niveau supérieur, un niveau supérieur-1440, un niveau principal et un niveau inférieur. L'échelonnement temporel fait intervenir deux couches, une couche de base et une couche d'amélioration. Ces deux couches ont la même résolution spatiale à la même fréquence image. Les Tableaux 1 à 4 présentent les limites des fréquences d'échantillonnage, des débits de pixels de luminance, des débits binaires et des dimensions du tampon du MVP.

TABLEAU 1

Limites supérieures de la densité d'échantillonnage

Niveau	Couche résolution spatiale		Profil
			Multivues
Supérieur	Amélioration (Image de droite)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	1 920 1152 60
	Inférieure (Image de gauche)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	1 920 1152 60
Supérieur-1440	Amélioration (Image de droite)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	1 440 1152 60
	Inférieure (Image de gauche)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	1 440 1152 60
Principal	Amélioration (Image de droite)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	720 576 30
	Inférieure (Image de gauche)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	720 576 30
Inférieur	Amélioration (Image de droite)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	352 288 30
	Inférieure (Image de gauche)	Echantillon/ligne Lignes/image bi-trame Images bi-trame/s	352 288 30

TABLEAU 2

Limites supérieures de fréquence d'échantillonnage de luminance (pixels/s)

Niveau	Couche résolution spatiale	Profil
		Multivues
Supérieur	Amélioration (Image de droite)	62 668 800
	Inférieure (Image de gauche)	62 668 800
Supérieur-1440	Amélioration (Image de droite)	47 001 600
	Inférieure (Image de gauche)	47 001 600
Principal	Amélioration (Image de droite)	10 368 000
	Inférieure (Image de gauche)	10 368 000
Inférieur	Amélioration (Image de droite)	3 041 280
	Inférieure (Image de gauche)	3 041 280

TABLEAU 3

Limites supérieures de débit binaire (Mbit/s)

Niveau	Profil
	Multivues
Supérieur	130, deux couches
	80, couche de base
Supérieur-1440	100, deux couches
	60, couche de base
Principal	25, deux couches
	15, couche de base
Inférieur	8, deux couches
	4, couche de base

TABLEAU 4

Spécification des dimensions du tampon (bits)

Niveau	Couche	Profil
		Multivues
Supérieur	Amélioration	15 898 480
	Base	9 787 248
Supérieur-1440	Amélioration	12 222 464
	Base	7 340 032
Principal	Amélioration	3 047 424
	Base	1 835 008
Inférieur	Amélioration	950 272
	Base	475 136

1.3 Extension des paramètres de la caméra

On a introduit dans le MVP une extension pour l'information fournie par la caméra. L'extension définit la hauteur de l'unité d'image, la distance focale, l'ouverture de diaphragme, l'angle de champ vertical, la position et l'axe de la caméra, et enfin son orientation dans le plan vertical.

2 Tests d'évaluation du MVP

Les essais de vérification du MVP ont été effectués dans trois laboratoires situés au Japon, en Allemagne et au Canada. Il a été rendu compte de ces essais à la réunion du GT 11 à Chicago (GT 11 N1373, septembre 1996. Test and video subgroup, «Results of MPEG-2 multi-view profile verification test»). Les résultats des différents laboratoires concordent: en général, aux débits binaires testés, les observateurs ont jugé que le système de codage du profil multivues MPEG-2 n'introduisait pas de défaut de codage gênant.

2.1 Méthode d'essai

On a appliqué la méthode à double stimulus utilisant une échelle de dégradation (variante II) présentée dans la Recommandation UIT-R BT.500. Toutefois, au lieu de l'échelle discontinue recommandée par l'UIT-R, on a utilisé une échelle continue afin d'obtenir des évaluations plus précises.

2.2 Conditions d'essai

On a utilisé les séquences d'essai créées pendant l'échange de flux binaires. Le Tableau 5 donne une idée générale des conditions d'essai. Des systèmes de visualisation différents étaient utilisés dans les trois laboratoires.

TABLEAU 5
Conditions d'essai subjectif

Séquences	«Street organ», «Flower pot», «Trapeze» (525/60) «Fun fair» (625/50)
Algorithmes et débits binaires (image de gauche/de droite)	MVP@ML: 6/3 Mbit/s, 9/4 Mbit/s Diffusion jumelée de données MP@ML: 4,5/4,5 Mbit/s, 6,5/6,5 Mbit/s Diffusion jumelée de données MP@ML comme ancrage inférieur: 2,5/2,5 Mbit/s («Street organ», «Fun fair»), 1,5/1,5 Mbit/s («Flower pot», «Trapeze») Original/original comme ancrage supérieur
Méthode d'essai	Méthode à double stimulus utilisant une échelle de dégradation (variante II) décrite dans la Recommandation UIT-R BT.500, avec échelle d'évaluation continue
Système de visualisation stéréoscopique (dimensions de l'image, distance d'observation)	HHI: système à deux miroirs (19 cm × 14 cm, 5 H) CRC: affichage séquentiel temporel et obturateur à LCD (40,6 cm × 30,5 cm, 4 H) NHK: projecteurs de télévision à haute définition (TVHD) avec écran à cristaux liquides et polariseurs (82 cm × 57 cm, 5 H)
Observateurs	HHI: 24 observateurs non experts CRC: 18 observateurs non experts NHK: 19 observateurs non experts (un observateur n'a pas été retenu lors de la sélection faite sur la base de la Recommandation UIT-R BT.500)

HHI: Heinrich-Hertz-Institut für Nachrichtentechnik (Allemagne)

CRC: Communications Research Center (Canada)

NHK: Nippon Hoso Kyokai (Japon)

2.3 Résultats des essais d'évaluation subjective

Les notes moyennes et l'intervalle de confiance à 95% de ces moyennes ont été calculés pour chaque condition d'essai. Les résultats de HHI, CRC et NHK sont reproduits au Tableau 6 et à la Fig. 3. HHI1 et HHI2 correspondent à des essais effectués chez HHI sur deux segments différents de la même séquence. HHI n'avait pas pu essayer de séquence entière par suite de l'insuffisance de la capacité de mémoire de l'unité de visualisation.

TABLEAU 6

Notes moyennes et intervalles de confiance à 95%

a) Séquence: Street organ

	Source	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Ancrage inférieur
NHK	4,71 ±0,17	4,18 ±0,27	4,40 ±0,26	4,06 ±0,39	3,51 ±0,32	1,74 ±0,33
CRC	4,24 ±0,37	4,19 ±0,33	4,33 ±0,29	4,27 ±0,34	4,07 ±0,35	2,19 ±0,35
HHI1	4,89 ±0,12	4,55 ±0,21	4,58 ±0,22	4,23 ±0,26	3,63 ±0,35	1,30 ±0,19
HHI2	4,86 ±0,13	4,68 ±0,19	4,85 ±0,13	4,44 ±0,24	4,24 ±0,32	1,80 ±0,23

b) Séquence: Flower pot

	Source	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Ancrage inférieur
NHK	4,79 ±0,16	4,03 ±0,44	4,28 ±0,25	4,07 ±0,33	4,13 ±0,37	2,28 ±0,32
CRC	4,53 ±0,14	4,57 ±0,20	4,45 ±0,22	4,40 ±0,20	4,40 ±0,21	2,70 ±0,34
HHI1	4,81 ±0,19	4,49 ±0,25	4,52 ±0,26	4,33 ±0,24	4,46 ±0,23	1,96 ±0,25
HHI2	4,83 ±0,14	4,48 ±0,21	4,33 ±0,22	4,08 ±0,26	4,16 ±0,25	1,69 ±0,24

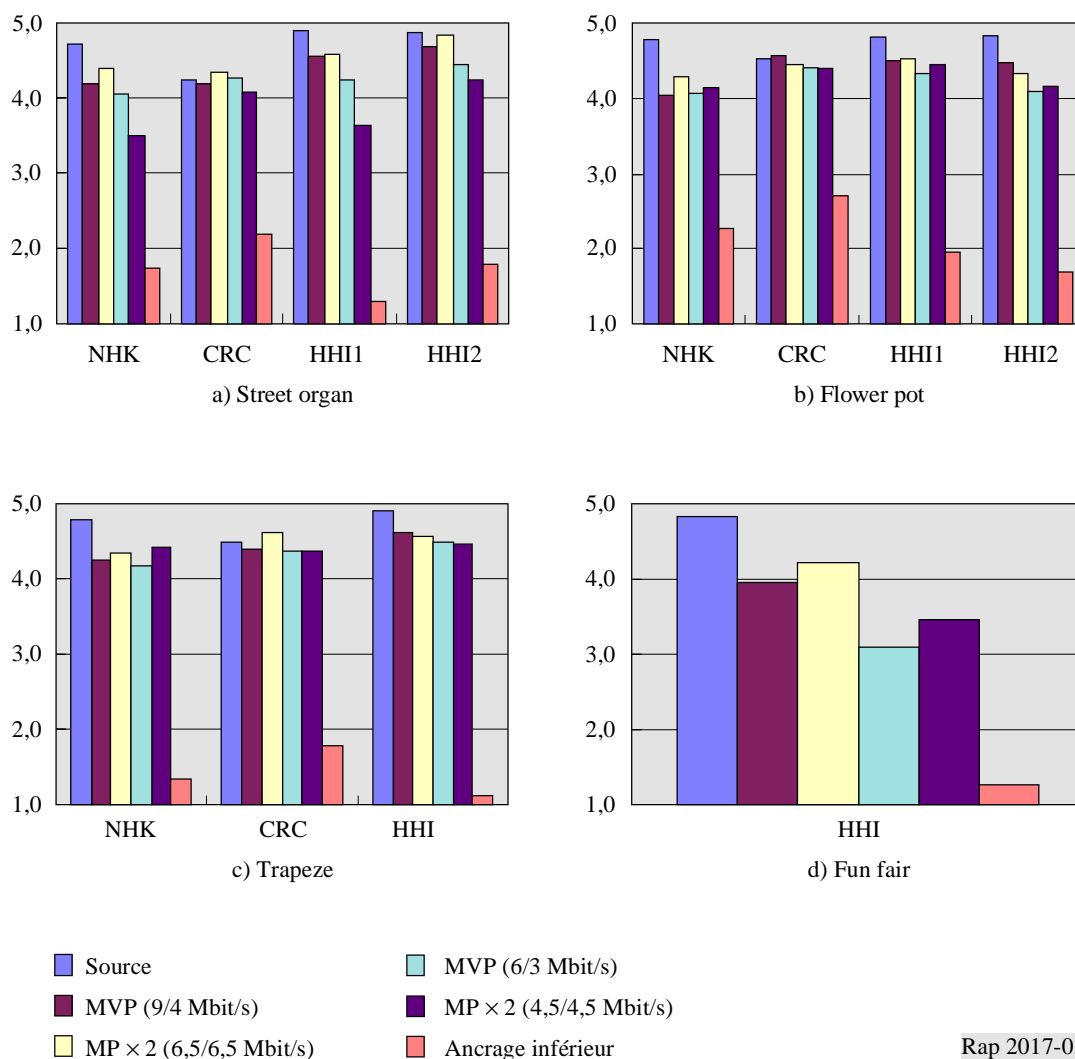
c) Séquence: Trapeze

	Source	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Ancrage inférieur
NHK	4,77 ±0,13	4,24 ±0,25	4,34 ±0,38	4,16 ±0,24	4,41 ±0,23	1,33 ±0,18
CRC	4,48 ±0,22	4,38 ±0,24	4,62 ±0,14	4,37 ±0,23	4,36 ±0,24	1,78 ±0,31
HHI1	4,90 ±0,11	4,60 ±0,19	4,55 ±0,25	4,48 ±0,27	4,46 ±0,28	1,13 ±0,14

d) Séquence: Fun fair

	Source	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Ancrage inférieur
HHI1	4,83 ±0,14	3,96 ±0,32	4,23 ±0,29	3,10 ±0,35	3,46 ±0,27	1,27 ±0,20

FIGURE 3
Notes moyennes d'évaluation subjective



Rap 2017-03

Certains aspects des résultats méritent d'être relevés:

- Dans chacune des quatre séquences, la note moyenne de la séquence MVP à un débit de 9/4 Mbit/s n'est pas sensiblement différente du résultat moyen de la diffusion jumelée de données de MP à un débit de 6,5/6,5 Mbit/s. De même, la note moyenne de la séquence MVP à un débit de 6/3 Mbit/s n'est pas sensiblement différente du résultat moyen de la diffusion jumelée de données de MP à un débit de 4,5/4,5 Mbit/s, sauf pour la paire de la séquence «Street organ». Pour cette dernière, la qualité du MVP est supérieure à celle du MP. Les différences d'évaluation subjective entre MVP et la diffusion jumelée de données de MP sont donc minimales au débit élevé pour les images comportant peu de mouvements («Flower pot» et «Trapeze») et/ou avec différence sensible de luminance entre gauche et droite («Fun fair»).
- «Fun fair» est la séquence où les moyennes présentent l'écart le plus grand. On y a observé la plus grande quantité de mouvement (variation du contenu de l'image d'une image bi-trame à l'autre) par rapport aux autres séquences. Tout particulièrement dans «Fun fair», les objets animés occupent la plus grande partie de l'image.

3 Travaux futurs sur la télévision stéréoscopique

Les progrès réalisés jusqu'à présent montrent que la télévision stéréoscopique est, sur le plan technique, du domaine du possible. Le profil multivues MPEG approuvé récemment constitue une base pour le codage et la compression des séquences vidéostéréoscopiques. Les essais d'évaluation de qualité qui ont été effectués apportent eux aussi des éléments

qui prouvent que, dans les limites des paramètres d'essai choisis, on peut obtenir une qualité d'image subjectivement jugée satisfaisante. Néanmoins, il reste beaucoup de problèmes à résoudre. Certaines des questions qui appellent une étude plus poussée sont les suivantes.

3.1 Critères à respecter

- Il serait souhaitable que tout futur système de télévision stéréoscopique soit compatible avec les systèmes naissants de télévision numérique monoscopique, et que le débit binaire supplémentaire associé soit aussi petit que possible.
- La qualité de l'image principale qui apparaît sur un écran de télévision monoscopique doit être aussi proche que possible de celle d'une image monoscopique utilisant toute la capacité du canal.

3.2 Informations requises pour la télévision à définition normale (TVDN) et pour la TVHD

- Degré de répartition asymétrique de débits binaires qui est possible pour les images de gauche et de droite pour qu'une séquence vidéostéréoscopique présente le minimum de dégradation de qualité pour l'image du niveau de base.
- Effet de la répartition asymétrique de débits binaires entre les images de gauche et de droite sur les défauts de codage et de compression subjectivement perçus, et qualité d'ensemble de la séquence vidéostéréoscopique.
- Facteurs qui pourraient causer une fatigue visuelle pour l'observateur; mesures propres à réduire ou à supprimer cette fatigue.
- Gamme de débits binaires requise pour obtenir une bonne qualité subjective pour l'image stéréoscopique comme pour l'image monoscopique fournie par l'image du niveau de base, définie dans le cadre d'essais additionnels d'évaluation portant sur un grand nombre de séquences représentant une large gamme de programmes et sur une gamme étendue de débits binaires.
- Méthodes d'essai appropriées pour l'évaluation des images stéréoscopiques.
- Algorithmes de codage permettant une compression plus efficace des signaux de télévision stéréoscopique.

Ces études doivent être effectuées en liaison avec le GT 11B et le GTM 10-11Q et avec d'autres groupes de travail et organisations compétents.
