

Évaluation de la qualité vidéo pour les nouveaux outils de codage émergents

Naty SIDATY¹, Wassim HAMIDOUCHE¹, Pierrick PHILIPPE², Olivier DEFORGES¹

¹Laboratoire IETR/INSA

20 Avenue des Buttes de Coësmes, 35043 Rennes Cédex, France

²btcom & Orange

1219 Avenue des Champs Blancs, 35510 Cesson-Sévigné, France

Naty.Sidaty@insa-rennes.fr

Wassim.Hamidouche@insa-rennes.fr, Olivier.Deforges@insa-rennes.fr

Résumé – Avec l’augmentation drastique des applications multimédias et de la consommation importante des contenus vidéo, la compression et l’évaluation de la qualité du contenu multimédia sont devenus un sujet de recherche très actif. Récemment, un nouvel outil de codage, appelé Joint Exploration Model (JEM), a été développé avec l’objectif principal de fournir un gain en débit considérable par rapport au standard de référence HEVC. Dans cet article, nous présentons une comparaison de performance de codage entre cet outil et celui de la référence HEVC (HM). A cette fin, des mesures objectives et des évaluations subjectives de la qualité ont été effectuées. Un ensemble de séquences vidéo, dans deux résolutions spatiales: Haute Définition (HD) et Ultra-Haute Définition (UHD), ont été utilisées dans cette étude. Ces vidéos sont encodées en utilisant le JEM.3 et le HM.17 à différents débits binaires. Les résultats ont montré que l’encodeur JEM permet, subjectivement, une amélioration de la qualité jusqu’à 40% à des débits binaires similaires. Objectivement, cette amélioration varie de 35 à 37% en fonction de la résolution spatiale utilisée.

Abstract – With the drastic increasing of multimedia applications and video coarse consumption, video compression and content quality evaluation have become an exciting and challenging topic. Recently, a new coding tool has been developed under the Joint Exploration Model (JEM) software with the main goal to provide a high bit rate saving compared to the HEVC standard. In this paper we present a performance-based comparison between the JEM and HEVC reference software (HM) through an objective measurement and a subjective quality assessment. A set of video sequences, in two spatial resolutions High Definition (HD) and Ultra-High Definition (UHD), have been used in this study. These videos are encoded using both JEM and HM software at different bitrates. Results have shown that the JEM codec enables, subjectively, a quality enhancement up to 40% at similar low bit rates. Objectively, this quality improvement is ranging from 35% to 37% depending on the spatial resolution.

1 Introduction

Le dernier standard de codage vidéo High Efficiency Video Coding (HEVC) [1], développé conjointement par le *Motion Picture Experts Group* (MPEG) et le *Video Coding Expert Groupe* (VCEG) sous le *Joint Collaborative Team on Video Coding* (JCTVC), permet une réduction de débit jusqu’à 60% [2], pour la même qualité subjective de la vidéo, par rapport à la norme MPEG4/AVC [3]. Les applications vidéo émergentes, telles que la réalité virtuelle (*Virtual Reality*, VR 360) combinées à la haute gamme dynamique (*High Dynamic Range*, HDR), offrent un rendu immersif et plus proche de la réalité. Cependant, ces nouveaux services nécessitent une résolution plus importante (4K, 8K) pour garantir une meilleure qualité de service, requise par les utilisateurs. En outre, l’utilisation des services vidéo augmente, de façon exponentielle, avec la prolifération des appareils mobiles et l’utilisation accrue des applications vidéo : les réseaux sociaux, le streaming vidéo, la TVIP, les jeux vidéo, etc. Pour faire face à ces nouveaux défis,

le groupe de travail (JVET) a récemment été créé par MPEG et l’UIT pour explorer une technologie de codage vidéo future *future video coding technology with a compression capability that significantly exceeds that of the current HEVC standard* [4]. Par conséquent, de nouveaux outils de codage ont été intégrés au standard de référence HEVC (HM) conduisant à une plate-forme de codage vidéo, appelée Joint Exploration Model (JEM). L’objectif principal de cette plate-forme est de fournir un gain de débit entre 25-30% [5, 6] par rapport au standard de référence HEVC. Ce gain de codage est rendu possible au détriment d’une complexité de codage supplémentaire estimée à x10 du côté du codeur/décodeur par rapport à la référence HEVC [1].

Le but principal de cette étude est d’évaluer, subjectivement, les améliorations, en terme de qualité, apportées par le JEM par rapport au standard HEVC dans deux résolutions HD et UHD (4K). En effet, des études objectives (basées signal) ont été réalisées mais aucune analyse subjective n’a été effectuée. Le reste du papier est organisé comme suit : une brève descrip-

TABLE 1 – Principaux outils inclus dans le JEM

Acronym	Tool name	Gain
AMT	Adaptive Multiple Transforms	5%
PMMVD	Pattern matched motion vector derivation	5%
QTBT	Quad-tree plus Binary Tree	4%
ALF	Adaptive Loop Filters	4%
BIO	Bidirectional Optical Flow	2%
NSST	Secondary transforms	2%
AMVR	Adaptive Motion Vector Resolution	2%
OBMC	Overlap Block Motion Compensation	2%
ATMVP	Advanced Temporal level motion vector prediction	2%
CCLM	Cross-component linear model prediction	1%
PDPC	Position dependent intra prediction combination	1%
AFFINE	Affine Motion Compensation	1%
RSAF	Reference Sample Adaptive Filtering	1%
Expected total BD-rate gain		32%

tion des outils de codage les plus importants introduits dans le JEM est fournie dans la section 2. La section 3 décrit la méthodologie et les procédures d'évaluation de la qualité utilisée dans cette expérimentation. Les résultats obtenus sont analysés dans la section 4. Enfin, la section 5 conclut ce document.

2 Description des outils de codage JEM

Le schéma de codage JEM est considéré comme une extension de la structure de codage HEVC. En effet, le codec JEM est construit en fonction du codage de référence HEVC (appelé HM). Ainsi, tous les outils JEM peuvent être individuellement désactivés en fonction du HM [8], lorsque tous les outils sont désactivés. La majorité des gains dans JEM sont fournis par une liste d'environ 15 outils, listés dans le tableau 1 avec leurs améliorations de codage individuelles. Le gain de codage pour chaque outil est estimé à l'aide de la métrique Bjøntegaard (BD)-rate [7] lorsque celui-ci est activé individuellement. La performance de chaque outil de codage a été mesurée dans la configuration *Random Access* comme spécifié dans les conditions de test de référence [9], d'une manière similaire à [5, 6]. Ces outils permettent des gains de codage allant jusqu'à 5%. En total, les contributions individuelles résultent en un gain hypothétique, en termes de réduction de débit, allant jusqu'à 32%. Pour plus d'informations sur ces outils, une description détaillée est donnée dans [10].

Afin d'évaluer l'efficacité du codage JEM par rapport au standard de référence HM, nous avons réalisé une campagne de tests subjectifs, impliquant des observateurs humains qui évaluent la qualité perçue d'un ensemble de vidéos HD et UHD. Nous présentons ci-dessous la méthodologie et la procédure d'évaluation suivies pour réaliser une cette expérimentation.

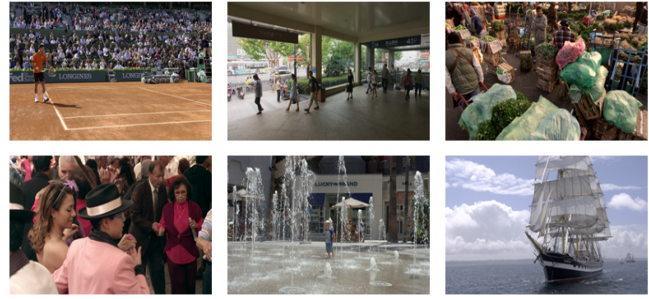


FIGURE 1 – Images extraites de la base d'étude.

3 Évaluation subjective de la qualité

3.1 Environnement global de l'expérience

Les tests subjectifs se sont déroulés dans la salle psychovisuelle du laboratoire IETR/INSA, aménagée conformément à la recommandation ITU-R BT.500-13 [11]. Un écran professionnel UHD rétro-éclairé à LED de 65 pouces (x651UHD) a été utilisé pour visualiser les séquences vidéos. Dix-sept observateurs, 10 hommes et 7 femmes âgés de 18 à 44 ans, ont participé à cette campagne.

3.2 Séquences de test

Dans cette campagne de tests, un ensemble de séquences vidéo, de contenu visuel varié, a été utilisé. Deux catégories de résolutions spatiales sont employées : HD (1920 × 1080) et UHD (3840 × 2160), comme indiqué dans le tableau 2. Les vidéos sont principalement sélectionnées à partir des bases de données MPEG et des vidéos du projet 4EVER¹. La Figure 1 illustre un exemple d'images extraites des différentes vidéos de la base d'étude. Quatre débits pour les contenus HD et 4 pour les contenus UHD (4K) ont été considérés pour créer l'ensemble de données des tests subjectifs. Les débits binaires ont été sélectionnés en se basant sur les plages de fonctionnement d'un scénario de diffusion TV (broadcast). Les séquences ont été encodées dans la plage de 500 kbps-8 Mbps en HD et de 2 à 24 Mbps pour les contenus UHD. Par conséquent, pour chaque contenu et résolution vidéo, 4 séquences sont générées à l'aide du codec de référence HEVC (HM16.7) et 4 séquences en utilisant le JEM3.0. Enfin, un total de 96 séquences vidéos ont été utilisées dans cette étude (48 HD et 48 UHD).

TABLE 2 – Description de la base de vidéos utilisée

Séquence	HD	UHD	Fps
Djoko	1920 x 1080	3840 x 2160	50
FoodMarket2	1920 x 1080	3840 x 2160	60
Metro	1920 x 1080	3840 x 2160	60
Tango	1920 x 1080	3840 x 2160	60
ToddlerFountain	1920 x 1080	3840 x 2160	60
Voiles	1920 x 1080	3840 x 2160	50

1. Le projet Enhanced Video ExpeRience 2 (4EVER-2) sur www.4ever-2.com

3.3 Procédure d'évaluation

Dans cette évaluation subjective de la qualité, la méthode de classification par catégorie de dégradation *Degradation Category Rating (DCR)* a été utilisée, conformément à la Recommandation UIT-T P.910 [11]. Chaque vidéo compressée a été présentée aux participants accompagnée de sa version de référence (originale). Les participants ont été invités à quantifier numériquement la qualité de la vidéo compressée. Autrement dit, chaque participant doit attribuer un score de qualité à chacune des 96 vidéos de tests, selon une échelle de notation allant de 1 (mauvaise qualité) à 10 (qualité excellente), comme l'a recommandé l'UIT P.910. Enfin, les scores des observateurs sont collectés et analysés afin de calculer le score moyen : Mean Opinion Score (MOS).

4 Résultats et discussions

Tout d'abord, des métriques objectives bien connues, comme les mesures PSNR et BD-BR (Bjøntegaard-Delta Bit Bit Rate) [7], sont utilisées pour la comparaisons objective des résultats. La Figure 2 montre les résultats obtenus en se basant sur le wPSNR pour l'ensemble des vidéos utilisées dans la campagne des tests. Comme nous pouvons le constater, le JEM obtient des valeurs de wPSNR plus élevées par rapport au standard de référence HM (HEVC) et, *a priori*, une meilleure qualité vidéo. Ce comportement est constamment observé pour différentes séquences vidéo à différents débits binaires, pour les deux résolutions de test (HD) et (UHD).

Cependant, ces résultats sont principalement basés "signal" et ne tiennent pas en compte les propriétés du système visuel humain (SVH).

Le tableau 3 présente les résultats obtenus par la mesure Bjøntegaard (BD-BR) pour les contenus HD et UHD. En moyenne, le codec JEM permet, en terme de débit binaire, des économies d'environ 35% et 37 % pour les séquences vidéo HD et UHD (4K), respectivement.

Subjectivement, la Figures 3 illustre les résultats obtenus en terme de MOS, associé à leurs intervalles de confiance, pour les deux résolutions spatiales utilisées. Le codec JEM permet d'améliorer, à débit équivalent, la qualité subjective pour les différentes vidéos par rapport au codec de référence HM. Les participants ont jugé les séquences encodées avec JEM avec une meilleure qualité par rapport à celles compressées à l'aide du HM, souvent avec des intervalles de confiance très faibles. Ces améliorations dépendent légèrement de la séquence considérée. Pour la séquence vidéo *Djoko*, l'amélioration apportée par le JEM est inférieure à d'autres séquences, en particulier pour les séquences *Metro* et *Tango*.

Néanmoins, pour des contenus particuliers, comme la séquence *ToddlerFountain* en résolution HD, ni le HM ni le JEM n'a permis de coder à une bonne qualité avec les débits considérés. En effet, cette vidéo a un contenu bruité (gouttes d'eau), qui est difficile à encoder. Enfin, une analyse statistique, en utilisant

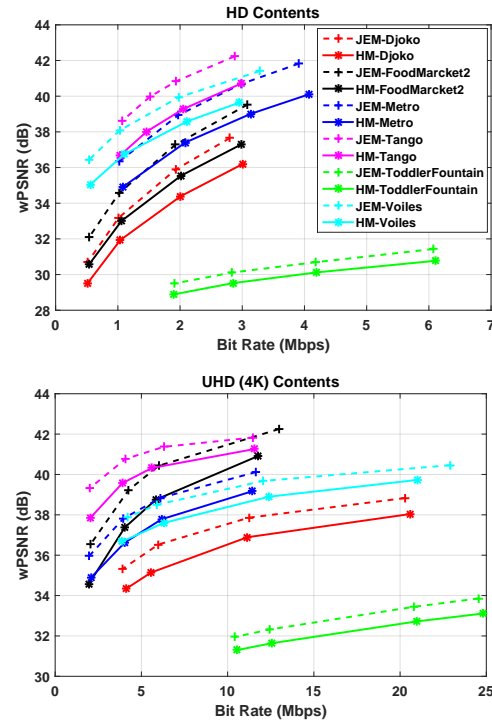


FIGURE 2 – Les performances de compression basées wPSNR pour les résolutions (haut) HD et (bas) UHD (4K).

Analyse de Variance Multivariée (MANOVA), a été réalisée sur le MOS collectés en considérant le débit, le contenu, la résolution et le codec comme des variables indépendantes. Les résultats montrent que seul le paramètre codec (JEM ou HM) avait une influence significative sur les scores subjectifs, avec $p\text{-value} < 0,0001^2$.

TABLE 3 – BD-BR between JEM and HM video codecs

Sequence	BD-BR (HD)	BD-BR (UHD)
Djoko	-30.94%	-37.17%
FoodMarket2	-35.52%	-36.31%
Metro	-35.59%	-37.21%
Tango	-38.28%	-42.18%
ToddlerFountain	-31.34%	-28.45%
Voiles	-42.34%	-40.83%
Mean	-35.66%	-37.02%

5 Conclusion

Dans cet article, nous avons étudié les performances du codage JEM par rapport à celui de référence HEVC (HM) sur des vidéos de différentes résolutions. Grâce à un ensemble de tests subjectifs, nous avons pu évaluer la qualité des vidéos encodées avec les deux encodeurs ainsi que les gains en termes de débit binaire. Les résultats montrent que l'encodeur JEM permet une

2. un facteur est considéré influant si $p\text{-value} < 0,05$

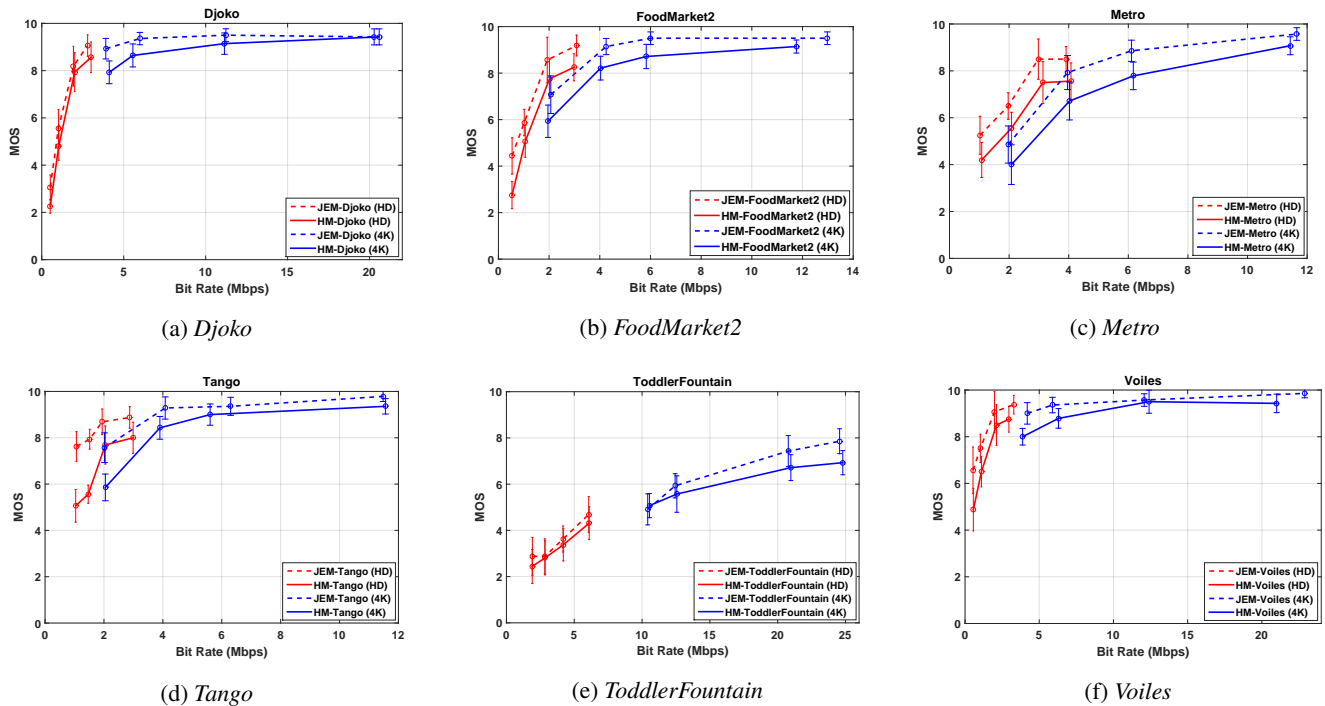


FIGURE 3 – MOS-based comparison, with associated 95% confidence intervals, for the used video dataset.

amélioration significative de la qualité vidéo subjective, pour les mêmes débits binaires, que le standard de référence HM. En outre, à des débits binaires élevés, le codec HM permet une bonne qualité vidéo et atteint la qualité de la vidéo codée avec le JEM. Enfin, pour une séquence vidéo particulière (*ToddlerFountain*) en résolutions HD, le HM et JEM ne permettent pas d'atteindre une bonne qualité aux débits considérés. Il est donc impératif d'imaginer une solution de codage séquentielle par laquelle les régions saillantes ont plus d'allocation de débit par rapport au reste de la vidéo (background).

Références

- [1] J-S. Gary, J-R. Ohm, W-J. Han, and T. Wiegand, *Overview of the High Efficiency Video Coding (HEVC) Standard*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
- [2] T-K. Tan, R. Weerakkody, M. Mrak, N. Ramzan, V. Baroncini, J-R. Ohm, and J-S. Gary, *Video Quality Evaluation Methodology and Verification Testing of HEVC Compression Performance*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 76–90, Jan. 2016.
- [3] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, and A. Luthra, *Overview of the H.264/AVC video coding standard*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560–576, July 2003.
- [4] *Jvet terms of reference*, in http://wftp3.itu.int/av-arch/jvetsite/jvet_tor.htm.
- [5] H. Schwarz, C. Rudat, M. Siekmann, B. Bross, D. Marpe, and T. Wiegand, *Coding Efficiency / Complexity Analysis of JEM 1.0 coding tools for the Random Access Configuration*, in Document JVET-B0044 3rd 2nd JVET Meeting : San Diego, CA, USA, February 2016.
- [6] E. Alshina, A. Alshin, K. Choi, and M. Park, *Performance of JEM 1 tools analysis*, in Document JVET-B0044 3rd 2nd JVET Meeting : San Diego, CA, USA, February 2016.
- [7] G. Bjøntegaard, *Calculation of Average PSNR Differences Between RD-curves*, in VCEG-M33 ITU-T Q6/16, Austin, TX, USA, April 2001.
- [8] C. Rosewarne, B. Bross, M. Naccari, K. Sharman, and G. Sullivan, *High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Improved Encoder Description*, in Document JCTVC-X1002 JCTVC Meeting : Geneva, CH, May 2016.
- [9] K. Suehring and X. Li, *JVET common test conditions and software reference configurations*, in Document JVET-B0044 3rd 2nd JVET Meeting : San Diego, CA, USA, February 2016.
- [10] J. Chen, E. Alshina, G. J. Sullivan, J-R. Ohm, and J. Boyce, *Algorithm description of Joint Exploration Test Model 3 (JEM3)*, in Document JVET-C1001 3rd JVET Meeting, Geneva, CH, May 2016.
- [11] ITU-R BT.500-13 Recommendation, *Methodology for the subjective assessment of the quality of television picture*, Geneva, 2012.