

des vecteurs mouvements : que doit-on vraiment transmettre ?

Guillaume Laroche¹, Joël Jung², Béatrice Pesquet-Popescu³

1. Canon Research Centre France, rue de la touche Lambert
35517 Cesson-Sevigne, France
guillaume.laroche@crf.canon.fr
2. Orange Labs, 38-40 rue du General Leclerc
92794 Issy Les Moulineaux, France
joelb.jung@orange-ftgroup.com
3. Telecom ParisTec, 46 rue Barrault, 75634 Paris, France
beatrice.pesquet@telecom-paristech.fr

RÉSUMÉ. Le standard de compression vidéo H.264/AVC obtient des gains significatifs par rapport à ses prédécesseurs en diminuant fortement l'entropie des résiduels de texture à l'aide, notamment, d'une augmentation du nombre de possibilités de codage. Les travaux pour une future norme ont déjà commencé au sein de chacun des deux groupes à l'origine de la norme H.264/AVC. La méthode de codage de l'information de mouvement par compétition de prédiction de vecteurs a été proposée dans le cadre des travaux exploratoires du groupe VCEG. Cette méthode visant à réduire la quantité d'information de mouvement a été intégrée dans le logiciel JM-KTA visant à préparer le nouveau standard. Dans cet article, une amélioration de cette méthode est proposée, permettant de limiter le débit des indices de prédicteurs. Elle est basée sur les caractéristiques de la fonction de sélection du meilleur prédicteur qui est reproduite au décodeur. Cette méthode, testée dans le JM-KTA, élimine 10 % des indices de prédicteurs et réduit ainsi le débit de l'information de mouvement.

ABSTRACT. The H.264/AVC video coding standard obtains significant compression gains compared to its predecessors. These gains come from a reduction of the texture residual entropy by, notably, increasing the amount of coding possibilities. The work for the future standard have already started in both groups that normalized H.264/AVC. The competition based scheme for motion vectors prediction has been proposed for the exploratory work of VCEG. This scheme dedicated to the reduction of the motion information was also integrated in the JM-KTA software. In this paper, an improvement of the motion vector competition scheme is proposed to reduce the bitrate of the predictors indexes. This scheme is based on the properties of the best predictors selection function which is used at the decoder side. This scheme, tested in the JM-KTA, removes 10% of the predictor indexes and thus reduces the bitrate dedicated to the motion information.

MOTS-CLÉS : HEVC, H.264/AVC, KTA, codage vidéo, critère débit distorsion, méthode de codage par compétition, vecteurs mouvement.

KEYWORDS: HEVC, H.264/AVC, KTA, rate distortion criterion, competitive scheme, motion vectors.

DOI:10.3166/TS.27.421-441 © 2010 Lavoisier, Paris

Extended abstract

The current video coding standard ITU-T SG16-Q6 H.264 of the Video Coding Experts Group (VCEG), also known as ISO/IEC JTC 1/SC 29/WG 11 MPEG-4 AVC [1] for the Moving Pictures Experts Group (MPEG), achieves a high compression gain by the improvement of existing coding tools and the inclusion of new ones. The improvements for the Inter prediction comes from the use of variable block sizes, multiple block-size transforms, $\frac{1}{4}$ pel motion estimation, in-loop deblocking filter and Skip mode prediction. All these tools significantly reduced the entropy of the texture residual. As a consequence, high compression gains were obtained with an increased proportion, in the global bitrate of the side information related to the competition information. This side information takes half of the global bitrate for a sequence encoded with a medium quality and reaches three quarters at low bitrates. Among this side information, the motion information takes the largest proportion. In H.264/AVC, the motion vector for each Inter block partition is predicted. For the classical Inter mode, the motion vector residual, the difference between the motion vector and its predictor, is inserted in the bitstream. For the Skip mode, which is a particular Inter mode where no information is transmitted to the decoder except the mode itself, the motion vector is equal to the predictor. In this standard, the motion vector predictor is a conditional median of the motion vectors used for the encoding of the neighboring blocks. The conditions of this median predictor depend on the reference frames pointed by the neighboring motion vectors and the partition sizes of the neighboring blocks. Today, both VCEG and MPEG groups which standardized H.264/AVC have established a Joint Collaborative Team on Video Coding (JCT-VC) to design its successor: the High Efficiency Video Coding (HEVC). However, VCEG has already investigated some technologies within the Key Technical Area (KTA) software, to which we contributed with the motion vector competition. The motion vector competition can exploit both spatial and temporal correlations in the motion vector field thanks to an optimal selection of the best predictor. In this scheme, several predictors are defined for the current motion vector instead of one in H.264/AVC. The predictor which minimizes the rate distortion criterion is then selected as the predictor for the coding of this current motion vector. In the KTA software, we showed that the best results were obtained with a set including two spatial predictors for the Skip mode and a set with the spatial median and the temporal collocated motion vector for the other Inter modes. The competition between these predictors involves the coding of the predictor selected, except if all the predictors of the set are equal. Indeed, in that case, the decoder is also able to test the equality of the predictors and, consequently, to know if the index needs to be extracted from the bitstream. The scheme proposed in this paper increases the number of predictor indexes which don't need to be transmitted to the decoder by

taking into account the rate distortion criterion and the lossless coding properties of the motion information. The criterion needs to know the predictors and the residual. So, at encoder side, the best predictor is already selected by the rate distortion criterion and the residual computed. At decoder side, the residual is extracted from the bitstream. The two possible reconstructed motion vectors are computed. It means that the predictor number one is added to the residual like the predictor number two. Then, the rate criterion is applied to select the best predictor (among predictor one and predictor two) for each possible reconstructed motion vector. If the best predictor is the same for both possible reconstructed motion vectors, then the encoder and the decoder know that this predictor was selected by the rate distortion criterion. In that case, the index of this predictor doesn't need to be transmitted by the encoder or to be extracted from the bitstream by the decoder. In this paper, we propose two implementations of the proposed scheme at encoder side. In the first one, the criterion is applied just before the writing of the residual. This implementation gives exactly the same selection of predictor as the classical implementation of MVComp. The only difference is the amount of indexes transmitted which is, at maximum, equal to the amount of indexes transmitted with the classical implementation of MVComp. In the second implementation, the proposed scheme is taken into account in the RD criterion. So, for each estimation of the rate dedicated to the motion information, the proposed test is computed. This second implementation can change the mode or motion vector selection for each block. The proposed scheme was implemented in the JM KTA 1.9r1. The results were generated with the Baseline profile according to the VCEG test conditions. Moreover, we added 2 quantization parameters to the four recommended by the group and we generated the results for eighteen other sequences in addition to the twelve sequences of the VCEG conditions in order to use five sequence sizes with various contents. The proposed scheme removes in average 10.6% of the predictor indexes and 13.6% at high bitrates. If we add this reduction to the indexes removed by the equality test, (the equality test is a particular case of the proposed scheme), more than a quarter of the predictor indexes are not transmitted. This reduction reaches 30% for the high bitrates. However, the average bitrate savings is only 0.3% compared to the results of the classical MVComp. Indeed, the predictor indexes represent only 2.6% of the global bitrate. But this reduction is systematic with the first implementation. Yet, the bitrate savings offered by the proposed scheme should increase if a higher number of predictors is used. In this paper, we also notice that the efficiency of the proposed scheme depends on the criterion used to select the best predictor. More precisely, this depends of the growth of the function used to estimate the rate of the residual. We show that if the $\text{floor}(\log(|x|))$ function is used instead of the function used in the KTA software, 40% of the predictors are eliminated with the current criterion. The reduction of information related to the competition, which represents more than half the global bitrate for medium and low bitrates, is an interesting way to increase the efficiency of the next video coding standard. The increase of the decoder knowledge on the encoder choices, as presented in this paper, is a way to reach this goal, even if some of these choices need to be standardized.

1. Introduction

Le standard de compression vidéo de référence, nommé ITU-T SG16-Q6 H.264 par le « Video Coding Experts Group » (VCEG) et ISO/IEC JTC 1/SC 29/WG 11 MPEG-4 AVC (*Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC)* n.d.) par le « Moving Pictures Experts Group » (MPEG), obtient des gains de compression significatifs par rapport à ses prédécesseurs. Ces performances proviennent d'une amélioration du codage des macroblocs Intra et des macroblocs Inter. Pour le codage Intra, trois partitionnements du macrobloc sont mis en compétition (16×16 , 8×8 et 4×4), avec 4 ou 9 prédicteurs pour chacun des blocs. Pour le codage Inter, un partitionnement allant du 16×16 au 4×4 en passant par des partitions rectangulaires est utilisé. Chacun de ces blocs peut être compensé en mouvement sur plusieurs images de référence, avec une précision sous-pixellique allant jusqu'au $\frac{1}{4}$ de pixel. Les blocs prédicteurs proviennent d'images de références qui ont été préalablement filtrées afin de ne pas contenir d'effet de blocs. Tous ces outils permettant de réduire fortement l'entropie des résiduels de texture sont combinés à des outils efficaces pour le codage de ces résiduels, tels que la compétition entre la transformée DCT 4×4 et la transformée DCT 8×8 , le « Coded Block Pattern » (CBP) et le « Context Adaptive Binary Arithmetic Coding » (CABAC). La norme H.264/AVC a ainsi multiplié les possibilités de codage pour réduire l'information de texture. Remarquons que la norme définit uniquement le décodeur, par conséquent, le codage par compétition des ces possibilités de codage n'implique pas nécessairement une recherche exhaustive. Les informations liées à la signalisation de ces différents modes mis en compétition représentent plus de la moitié du débit total d'un flux encodé avec des paramètres de codage correspondant à une qualité moyenne et leur proportion atteint les trois quarts de ce débit total pour un flux encodé avec des paramètres de codage correspondant à une basse qualité. Un codage plus efficace de ces informations est un axe de recherche essentiel pour l'amélioration incrémentale de cette norme.

Aujourd'hui, les deux groupes à l'origine d'H.264/AVC ont commencé le processus de normalisation du successeur de cette norme. Le nouveau groupe « Joint Collaborative Team on Video Coding » (JCT-VC) (VCEG *et al.*, 2010) issu des groupes MPEG et VCEG a d'ores et déjà étudié les 27 réponses à l'appel à propositions (Ohm *et al.*, 2010) pour ce nouveau standard. Un logiciel de référence nommé « Test Model under Consideration » (TMuC) (Davies *et al.*, 2010) est en cours d'implémentation pour cette future norme « High Efficiency Video Coding » (HEVC). Des travaux exploratoires sur de nouvelles technologies pour le codage avaient déjà été menés au sein du groupe VCEG par le biais du logiciel « Key Technical Area » (KTA) (*KTA Software coordination* n.d.). Ce logiciel regroupe des améliorations incrémentales de la norme H.264/AVC. Une grande partie des méthodes concerne l'amélioration de l'interpolation sous-pixellique avec des filtres adaptatifs (Vatis *et al.*, 2005), (Vatis *et al.*, 2006) et les filtres adaptatifs directionnels (Rusanovksyy *et al.*, 2007), (Rusanovksyy *et al.*, 2008). D'autres outils concernent la sélection de la meilleure possibilité de codage comme le « Rate Distortion Optimized Quantization » (RDOQ) (Karczewicz *et al.*, 2008) qui optimise la sélection du pas de quantification (QP) au niveau macrobloc. La méthode « Internal Bit Depth Increase » (IBDI) (Chujoh *et al.*, 2007) augmente la précision des calculs à l'encodeur et au décodeur. Une transformée directionnelle de type KLT a aussi été

intégrée pour le codage Intra « Mode Dependent Directional Transform » (MDDT) (Ye *et al.*, 2007). Tous les outils intégrés dans le KTA concernent l'amélioration du codage de la texture, seuls deux outils proposent de réduire les autres informations de codage. Parmi ces deux outils, les grandes tailles de blocs (32×32 , ou 64×64) (Chenn *et al.*, 2009) intégrés dans les dernières versions du JM KTA, permettent de regrouper efficacement un ensemble d'informations complémentaires (modes, vecteurs mouvement, images de référence, etc.). L'autre outil est notre méthode de codage de vecteurs mouvement par compétition de prédictions (MVComp) (Jung *et al.*, 2006), (Laroche *et al.*, 2008) qui permet de réduire l'entropie de l'information de mouvement à l'aide d'une compétition de prédicteurs spatio-temporels. La méthode MVComp a été intégrée dans la quasi totalité des 27 réponses à l'appel à proposition de HEVC (Ohm *et al.*, 2010).

Dans cet article, nous proposons une méthode qui permet de réduire l'information de compétition liée à la méthode MVComp grâce aux caractéristiques de sélection de la meilleure prédiction du vecteur mouvement courant. Le principe de la méthode est de reproduire, au niveau du décodeur, des choix effectués au codeur. Les méthodes de codage distribué ont pour principe de transférer l'intelligence de l'encodeur vers le décodeur. Ces méthodes proviennent des travaux sur la théorie de l'information de Wyner et Ziv (Wyner *et al.*, 1976). La philosophie des méthodes de codage vidéo distribué est de reconstituer au niveau du décodeur les images prédites sans que l'encodeur n'ait effectué la totalité des opérations. Dans ces méthodes, les informations de mouvement manquantes sont calculées par le décodeur grâce aux bits de parité envoyés par le codeur canal. Dans notre approche, aucune information n'est transmise en plus du flux standard. Les méthodes de codage distribué ont été largement étudiées (B. *et al.*, 2005) avec le développement du codeur PRISM (Puri *et al.*, 2002) et plus récemment le codec développé dans le projet Discover (Artigas *et al.*, 2007). L'objectif, ici, est cependant différent, car il ne s'agit pas de réduire la complexité de l'encodeur ou d'augmenter la résistance aux erreurs, mais de réduire le débit de l'information liée aux choix de codage. Face à l'augmentation de cette information de compétition, ce principe de codage est de plus en plus étudié dans l'état de l'art. Dans (Sugimoto *et al.*, 2004), un algorithme de mise en correspondance des pixels voisins du bloc courant est utilisé pour la construction de nouveaux prédicteurs Intra dans les images Inter. Le décodeur est capable de reproduire le même algorithme de mise en correspondance. Par conséquent, aucune information complémentaire n'est transmise de l'encodeur au décodeur. Selon ce principe, dans (Kamp *et al.*, 2008) et (Laroche, 2009), une estimation de mouvement reproductible au décodeur utilisant les données causales du bloc courant, est appliquée dans le but de supprimer une partie de l'information de mouvement.

La suite de cet article est organisée de la manière suivante : la deuxième section décrit le codage et la sélection de l'information de mouvement dans la norme de référence et la méthode de codage par compétition de prédictions de vecteurs mouvement. La troisième section présente la méthode proposée pour deux prédicteurs et sa généralisation au cas de N prédicteurs. Enfin, la quatrième section donne les résultats expérimentaux, et présente une analyse de l'impact de la méthode sur le codage, puis, de la relation entre les propriétés de la fonction de sélection et le nombre de prédicteurs implicites éliminés.

2. État de l'art

2.1. Sélection et codage de l'information de mouvement dans la norme H.264/AVC

En compression vidéo, afin d'exploiter au mieux les redondances temporelles entre les images successives, on introduit une notion de mouvement pour caractériser les différences entre les images. Dans la norme H.264/AVC, le modèle du mouvement est de type translationnel par bloc. Le coût de l'information de mouvement pour les algorithmes de compression utilisant une compensation de mouvement par bloc, dépend de trois paramètres : la taille des blocs, la résolution sous-pixellique et l'entropie de l'information. Pour réduire le débit lié à l'information de mouvement, la norme H.264/AVC utilise un codage prédictif des vecteurs mouvement. Par conséquent, un résiduel de vecteurs mouvement est transmis au lieu du vecteur mouvement. Pour le codage du vecteur mouvement courant mv , le résiduel ε_{mv} est donné par la formule suivante :

$$\varepsilon_{mv} = mv - p \quad [1]$$

où p est le prédicteur. Ce prédicteur p est un médian spatial pour chacune des composantes horizontale et verticale. Nous notons ce prédicteur $mv_{H.264}$. Les trois vecteurs voisins du vecteur courant mv , utilisés pour le calcul du médian sont mv_a , mv_b et mv_c , représentés dans la figure 1. En fonction de la taille des blocs voisins et du bloc courant, le vecteur mv_c peut être remplacé par le vecteur mv_d . Pour des cas particuliers, dépendant des caractéristiques des blocs voisins, le prédicteur p peut être égal à mv_a ou mv_b ou mv_c ou 0. Ces caractéristiques sont l'appartenance de ces blocs à l'image, la taille du bloc courant, des blocs voisins et les images de référence utilisées pour le codage des blocs voisins. Par exemple, si un seul des vecteurs voisins a la même image de référence que le vecteur courant, la valeur de p est égale à ce vecteur voisin. De même, si l'un des blocs a , b , c , ou d est codé en Intra (modes pour lesquels il n'y a pas de vecteur mouvement) le vecteur mouvement pour ce bloc est égal à 0.

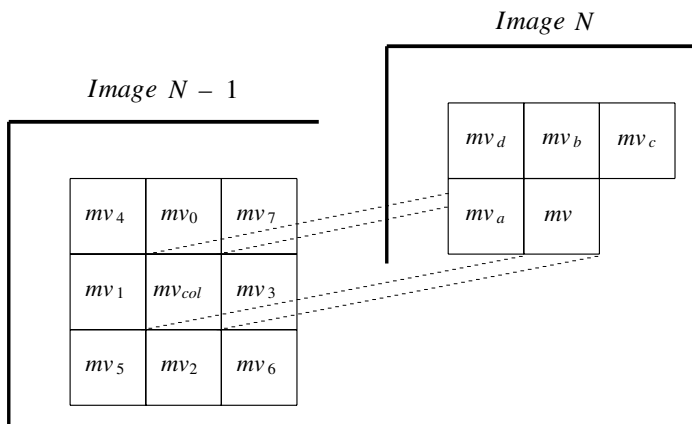


Figure 1. Localisation des vecteurs spatiaux et temporels utilisés dans la méthode MVComp

Le but de la compression vidéo avec pertes est de minimiser le débit pour une qualité maximum. Ce but peut être atteint grâce à la minimisation d'un critère débit distorsion (RD) (Sullivan *et al.*, 1998) donné par la formule suivante :

$$J = D + \lambda R \quad [2]$$

où D est la distorsion, R est le débit et λ est le multiplicateur de Lagrange. Le débit R est l'addition des débits de toutes les composantes dont les résiduels de vecteur mouvement R_{mv} . Chaque composante du résiduel d'un vecteur mouvement est codée à l'aide d'une table de Golomb pour le « Context Adaptive Variable Length Coding » (CAVLC). Pour le CABAC, ces bits sont codés arithmétiquement. Dans le logiciel de référence de la norme (*H.264/AVC software coordination* n.d.), le critère de l'équation (2) est calculé pour l'ensemble des modes de codage (prédictions Intra et Inter et l'ensemble des partitionnements), pour l'ensemble des positions entières de l'estimation de mouvement.

Le codage de l'information de mouvement a été largement étudié dans l'état de l'art. On peut distinguer deux types de méthode : les méthodes de codage avec pertes (Joshi *et al.*, 1995) et les méthodes de codage sans pertes. Les méthodes de codage avec pertes ont pour philosophie de traiter le mouvement en tant qu'information quantifiable, comme l'information de texture. Ce type de méthode a été testé récemment pour améliorer le codage de la norme H.264/AVC (Cagnazzo *et al.*, 2009) avec des résultats encourageants pour des champs de vecteurs mouvement complexes. Pour les méthodes de codage de vecteurs sans perte par prédiction, comme utilisé dans la norme H.264/AVC, l'efficacité dépend de la pertinence du prédicteur. L'hypothèse selon laquelle les prédicteurs spatiaux sont inefficaces dans le cas d'une rupture du mouvement dans le champ de vecteurs a été vérifiée dans (Yeh *et al.*, 1995), avec l'utilisation de prédicteurs temporels. Dans (Chen *et al.*, 1997), les redondances spatiales et temporelles sont exploitées pour améliorer le codage des vecteurs. Le choix entre deux types de prédicteurs dépend des statistiques locales, ce qui ne permet pas un choix optimal entre les deux types de corrélations. Une compétition exhaustive entre des prédicteurs spatiaux et des prédicteurs temporels est utilisée dans (Tourapis *et al.*, 2005) mais uniquement pour le mode Direct des images B. Pour réduire la complexité de l'estimation de mouvement, dans (De Natale *et al.*, 2001), une compétition entre un nombre élevé de prédicteurs est effectuée, et l'indice du prédicteur est transmis au décodeur, en plus de l'erreur de résiduel. Une compétition exhaustive entre trois vecteurs voisins spatiaux est mise en oeuvre dans (Deuk Kim *et al.*, 1999). Dans cette méthode, seule la redondance spatiale est exploitée, ce qui peut être inefficace lors d'une grande disparité dans le champ de vecteurs mouvement. Dans ce cas, une prédiction temporelle devrait être plus efficace. Enfin, nous avons proposé une méthode de codage par compétition de prédicteurs spatiaux et temporels dans (Laroche *et al.*, 2008) que nous décrivons dans la section suivante.

2.2. Codage par compétition des vecteurs mouvement

La compétition de prédictions de vecteurs mouvement permet d'exploiter les corrélations spatiales et temporelles des champs de vecteurs mouvement, grâce à une

sélection optimale du meilleur prédicteur. Parmi les prédicteurs temporels, nous avons défini le prédicteur collocated mv_{col} , qui est le vecteur mouvement utilisé pour le codage du bloc de même position que le bloc courant dans la première image de référence (cf. figure 1). Nous avons aussi défini des combinaisons de prédicteurs, des mises à l'échelle en fonction des distances temporelles parcourues par les prédicteurs et le vecteur courant mv , à la fois pour les images P et les images B. Dans la méthode MVComp, N prédicteurs sont mis en compétition à l'aide du critère RD (cf. équation (2)). Le débit des vecteurs mouvement R_{mv} est remplacé par le débit $R_{mv/mm}$ défini par :

$$R_{mv/mm} = \min\{\zeta(\varepsilon_{mv_i}) + \zeta(i)\}_{i \in \{1, \dots, N\}} \quad [3]$$

où ε_{mv_i} est le résiduel associé au prédicteur p_i ($\varepsilon_{mv_i} = mv - p_i$) et $\zeta(x)$ est le coût de débit réel ou estimé de l'information x . $\zeta(i)$ est donc le coût de l'indice à transmettre. Afin de limiter l'impact de la compétition sur le débit, un test d'égalité des prédicteurs est utilisé. En effet, si tous les prédicteurs sont égaux, l'encodeur ne transmet pas l'indice i au décodeur car ce dernier peut lui aussi calculer tous les prédicteurs et tester leur égalité. La méthode est aussi appliquée au mode Inter Skip qui est un mode particulier de codage puisqu'aucune information, excepté le mode de codage lui-même, n'est transmise au décodeur. Par conséquent, pour ce mode, aucun résiduel de vecteur mouvement n'est inséré dans le flux binaire. Avec la méthode MVComp, pour chaque macrobloc codé en Skip, un indice de prédicteur est transmis, sauf si tous les prédicteurs sont égaux.

Les expériences ont montré (Laroche *et al.*, 2008) que le meilleur résultat moyen est obtenu avec la compétition de deux prédicteurs spatiaux pour le mode Skip et la compétition du médian de la norme H.264 ($mv_{H.264}$) et du prédicteur collocated (mv_{col}) pour les autres modes Inter. Nous utiliserons cette configuration comme référence pour la méthode proposée dans cet article. Dans cette configuration, les indices de prédicteurs de vecteurs mouvement transmis au décodeur représentent 15 % de l'information de mouvement et leur débit atteint près de 2,6 % du débit total. Dans cet article, nous proposons de réduire le débit de cette nouvelle information.

3. Méthode proposée

La méthode présentée ci-dessous permet de « deviner » certains prédicteurs sélectionnés par la compétition de prédictions de vecteurs mouvement. Pour cela, il faut que le décodeur soit capable de reproduire le processus de sélection du prédicteur du vecteur mouvement réalisé par l'encodeur. L'encodeur et le décodeur utilisent donc le même critère pour discerner les prédicteurs à éliminer.

Nous expliquerons tout d'abord le procédé d'élimination de l'indice lorsque deux prédicteurs sont en compétition, puis le problème sera étendu au cas d'ensembles composés de N prédicteurs.

3.1. Élimination d'indices avec deux prédicteurs

Dans la méthode de codage des vecteurs mouvement par compétition de prédictions, pour chaque vecteur, le décodeur calcule les prédicteurs mis en

compétition au codeur. Le décodeur teste l'égalité des prédicteurs. Si les prédicteurs ne sont pas égaux, l'indice du prédicteur utilisé est lu dans le flux et le vecteur mv est obtenu par l'addition du résiduel ε_{mv} et du prédicteur correspondant à l'indice décodé. Considérons maintenant que le décodeur connaît le critère de choix utilisé pour le codage.

Le codeur a déterminé le meilleur prédicteur à l'aide de l'équation (3) de la manière suivante :

Si

$$\min\{\zeta(\varepsilon_{mv_1}) + \zeta(1), \zeta(\varepsilon_{mv_2}) + \zeta(2)\} = \zeta(\varepsilon_{mv_1}) + \zeta(1) \quad [4]$$

alors le prédicteur p_1 est sélectionné, sinon

$$\min\{\zeta(\varepsilon_{mv_1}) + \zeta(1), \zeta(\varepsilon_{mv_2}) + \zeta(2)\} = \zeta(\varepsilon_{mv_2}) + \zeta(2) \quad [5]$$

et par conséquent, p_2 est le prédicteur sélectionné.

Notons $mv_1 = \varepsilon_{mv} + p_1$ le vecteur mouvement décodé avec le prédicteur p_1 , et $mv_2 = \varepsilon_{mv} + p_2$ le vecteur mouvement décodé avec le prédicteur p_2 . Le vecteur original est soit mv_1 soit mv_2 . Si p_1 a été choisi alors :

$$\min\{\zeta(mv_1 - p_1) + \zeta(1), \zeta(mv_1 - p_2) + \zeta(2)\} = \zeta(mv_1 - p_1) + \zeta(1) \quad [6]$$

Si l'on ajoute à cette condition que :

$$\min\{\zeta(mv_2 - p_1) + \zeta(1), \zeta(mv_2 - p_2) + \zeta(2)\} = \zeta(mv_2 - p_1) + \zeta(1) \quad [7]$$

alors il est certain que le prédicteur p_1 est le prédicteur qui a été choisi par le codeur, car il est sélectionné pour les deux vecteurs possibles, mv_1 et mv_2 . Le décodeur peut effectuer le même processus. Dans ce cas, il n'est donc pas nécessaire de transmettre l'indice du prédicteur si l'équation (6) et l'équation (7) sont vérifiées.

3.2. Élimination d'indices avec N prédicteurs

Généralisons notre critère au cas où l'ensemble des prédicteurs de vecteurs mouvement contient N prédicteurs. Soit $mv_i = \varepsilon_{mv} + p_i$ le vecteur mouvement décodé avec le prédicteur p_i . Si l'indice du prédicteur n'est pas transmis, il existe alors N vecteurs possibles. Le prédicteur p_i est le prédicteur utilisé et il n'est pas nécessaire de transmettre son indice i , si et seulement si :

$$\begin{aligned} \min\{\zeta(mv_1 - p_1) + \zeta(1), \dots, \zeta(mv_1 - p_i) \\ + \zeta(i), \dots, \zeta(mv_1 - p_N) + \zeta(N)\} \neq \zeta(mv_1 - p_1) + \zeta(1) \\ \vdots \end{aligned}$$

$$\begin{aligned}
 &\min\{\zeta(mv_{i-1} - p_1) + \zeta(1), \dots, \\
 &\quad \zeta(mv_{i-1} - p_N) + \zeta(N)\} \neq \zeta(mv_{i-1} - p_{i-1}) + \zeta(i-1) \\
 \min\{\zeta(mv_i - p_1) + \zeta(1), \dots, &\zeta(mv_i - p_i) + \\
 &\zeta(i), \dots, \zeta(mv_i - p_N) + \zeta(N)\} = \zeta(mv_i - p_i) + \zeta(i) \\
 &\min\{\zeta(mv_{i+1} - p_1) + \\
 &\zeta(1), \dots, \zeta(mv_{i+1} - p_N) + \zeta(N)\} \neq \zeta(mv_{i+1} - p_{i+1}) + \zeta(i+1) \\
 &\quad \vdots \\
 \min\{\zeta(mv_N - p_1) + \zeta(1), \dots, &\zeta(mv_N \\
 - p_i) + \zeta(i), \dots, \zeta(mv_N - p_N) &+ \zeta(N)\} \neq \zeta(mv_N - p_N) + \zeta(N)
 \end{aligned}$$

[8]

3.3. Implémentation

Au niveau du décodeur, les N prédicteurs sont calculés et le test d'égalité est effectué. Si les prédicteurs ne sont pas égaux, on applique la méthode proposée : on lit le résiduel du vecteur courant dans le flux et on teste l'équation (8). Si l'on obtient un prédicteur implicite, on n'extrait pas du flux l'indice du prédicteur, mais on calcule directement la prédiction inverse avec ce prédicteur implicite.

Au niveau de l'encodeur, deux implémentations sont possibles. Dans la première, on peut appliquer l'algorithme uniquement au moment de l'écriture de l'indice du prédicteur dans le flux. Cette implémentation fournira exactement la même sélection des prédicteurs de vecteurs mouvement et les mêmes modes de codage que ceux obtenus avec la méthode MVComp pour une configuration identique. La seule différence est le nombre d'indices transmis au décodeur, qui est au maximum égal à celui de la méthode MVComp. Cette implémentation ne modifie pas la distorsion. Par conséquent, aucune perte, en termes de débit-distorsion, ne peut être observée par rapport à MVComp. Cette implémentation correspond à un « transcodage » d'un flux utilisant la méthode MVComp. On notera cette implémentation $Trans_{SuppPred}$.

La deuxième implémentation consiste à insérer la méthode dans le critère débit distorsion. C'est à dire qu'elle est appliquée à chaque calcul d'un résiduel de vecteur, notamment pendant la phase de sélection du meilleur vecteur mouvement pour un bloc. Dans ce cas, le critère de suppression influe donc sur les choix de codage. Théoriquement, cette deuxième implémentation que nous noterons $RD_{SuppPred}$ est plus efficace, car pour chaque bloc le meilleur choix de codage est sélectionné.

3.4. Remarques et discussion

La méthode proposée a été implémentée au même moment que la méthode MVComp. Cependant, nous ne l'avons pas soumise au jugement du groupe VCEG

lors de la présentation de MVComp en juillet 2006 (Jung *et al.*, 2006), car la philosophie d'utilisation de choix de codage au décodeur était jugée comme allant à l'encontre de certaines règles imposées par le groupe. En effet, ce type de méthode implique la normalisation de choix de codages engendrant la normalisation d'une partie du codeur (seul le décodeur est normalisé actuellement). Cependant, il faut remarquer qu'il est possible de s'affranchir de cette contrainte en normalisant, au décodeur uniquement, plusieurs fonctions de choix susceptibles d'être utilisées dans une implémentation réaliste et en transmettant, à l'aide d'un indice, la fonction utilisée par le codeur pour chaque séquence. Un de ces indices indiquerait notamment l'utilisation d'une fonction de choix non normative par le codeur, qui impliquerait la non utilisation du principe décrit dans cet article pour le décodage. Ceci permettrait de laisser une totale liberté aux choix d'implémentation du codeur pour ne pas se limiter aux fonctions définies. Pour la méthode proposée, il suffirait de normaliser le critère de sélection du meilleur prédicteur de vecteur mouvement défini dans l'équation (3). Pour cette fonction de choix de codage, basée sur le critère débit distorsion, il serait nécessaire de définir plusieurs fonctions d'estimation du débit $\zeta(x)$. Parmi ces fonctions, on pourrait définir celle utilisée dans le logiciel de référence qui correspond à une table de Golomb et que nous noterons $Golomb(x)$. Pour des implémentations peu complexes, il serait intéressant de normaliser la fonction de valeur absolue de l'information x , $|x|$. De même, pour modéliser le codage CABAC, on pourrait utiliser les fonctions x^2 , e^x , $floor(\log(|x|))$ ou, tout simplement, le calcul du débit réel de l'information x .

Comme nous l'avons vu dans l'introduction, des méthodes basées sur des choix de codage conjoints à l'encodeur et au décodeur ont cependant été proposées (Sugimoto *et al.*, 2004), (Kamp *et al.*, 2008). En février 2009, une méthode utilisant cette philosophie (Kim *et al.*, 2009) a été soumise au groupe VCEG pour le KTA. Cette méthode est une extension de MVComp qui utilise un algorithme de mise en correspondance pour sélectionner automatiquement le meilleur prédicteur de vecteur mouvement. Cependant, cette méthode n'a pas été acceptée parmi les méthodes du KTA. Les méthodes (Sugimoto *et al.*, 2004), (Kamp *et al.*, 2008), (Kim *et al.*, 2009) nécessitent la normalisation d'un critère (mesure de disparité et nombre de pixels utilisés) de mise en correspondance du signal causal de texture pour le décodeur. De plus, il faut aussi définir pour le décodeur l'ensemble des positions à tester ou l'algorithme utilisé pour que l'encodeur et le décodeur testent exactement les mêmes positions. Ces méthodes peuvent se révéler complexe au niveau du décodeur. Par rapport à la méthode de l'article (Kim *et al.*, 2009), notre méthode est moins complexe si l'on considère que le critère de mise en correspondance d'un pixel a la même complexité que le critère d'estimation de débit d'une composante d'un résiduel de vecteurs. En effet, dans notre méthode moins de données sont traitées. De plus, avec notre méthode le bon choix du prédicteur au sens du critère débit distorsion peut être fait pour chaque vecteur mouvement, ce qui n'est possible avec le critère défini dans l'article (Kim *et al.*, 2009).

Une des problématiques soulevées par les méthodes utilisant des informations temporelles pour le codage de l'information de mouvement est la reconstruction exacte de ces informations de mouvement en cas de pertes de données. Notons que la reconstruction exacte de l'information de mouvement ne permet pas un décodage correct des images précédées d'une perte de données. Seule une image totalement

codée en Intra permet un tel décodage. Les méthodes (Sugimoto *et al.*, 2004), (Kamp *et al.*, 2008), (Kim *et al.*, 2009) utilisant l'information de texture pour la construction du champ de vecteurs mouvement, il est impossible de reconstruire exactement le champ de vecteurs mouvement si l'information de texture spatiale et/ou temporelle a été modifiée. Dans le cas de la méthode MVComp, si un prédicteur temporel est utilisé, le champ de vecteurs ne peut pas être reconstruit correctement. De plus, la suppression des indices de prédicteurs par l'égalité ou par la méthode proposée implique un ordonnancement spécifique de ces indices dans le flux. Dans le cas contraire, une mauvaise suppression d'un indice de prédicteur, créée par un prédicteur temporel faux, provoque des erreurs dans la lecture du train binaire de l'image courante. Pour pallier à ceci, nous proposons d'indiquer le nombre d'indices suivi de ces indices en début ou fin de slices. Cette écriture du train binaire permet une lecture de ce train binaire pour MCVomp et pour la méthode proposée en cas de pertes de données. Enfin, la norme H.264/AVC permet un partitionnement des éléments de syntaxe du train binaire en paquets afin de protéger de manière inégale ces paquets. Avec cet outil, il est possible de protéger efficacement l'information de mouvement afin qu'aucune perte ne puisse affecter la perte d'un résiduel de vecteurs. Dans ce cas, la méthode MVComp et la méthode proposée engendrent une reconstruction exacte du champ de vecteurs contrairement aux méthodes (Sugimoto *et al.*, 2004), (Kamp *et al.*, 2008), (Kim *et al.*, 2009) qui ne peuvent pas reconstruire correctement cette information en cas de pertes de l'information de texture.

Les méthodes décrites dans (Sugimoto *et al.*, 2004), (Kamp *et al.*, 2008) ont été proposées en réponses à l'appel à propositions du remplaçant d'H.264/AVC (Ohm *et al.*, 2010). Les méthodes d'estimation de mouvement au décodeur font actuellement l'objet d'une étude d'efficacité (Wien *et al.*, 2010). Les méthodes utilisant des choix de codage pourraient donc être admises dans la future norme de codage vidéo.

4. Résultats expérimentaux

Dans cette section nous discutons nos choix d'implémentation, puis nous présentons les réductions de débit obtenues avec l'implémentation $RD_{SuppPred}$ de l'algorithme, avant d'analyser l'impact de la méthode sur le codage. Enfin, nous étudions les performances de la méthode selon la fonction de choix $\zeta(x)$ utilisée.

4.1. Conditions de tests

Les expériences ont été menées dans le JM KTA 1.9r1 avec le profil Baseline IPP. Nous avons utilisé les conditions de tests recommandées par le groupe VCEG (Tan *et al.*, 2008). De plus, afin de viser les hauts, moyens et bas débits, nous présenterons les résultats avec les QP 17 et 42 en plus des quatre QP 22 27 32 37 recommandés par le groupe. Nous avons ajouté, aux 12 séquences VCEG, 18 autres séquences (4 CIF, 3 4CIF, 7 720p et 4 1080p). Dans cet ensemble de 30 séquences, les résolutions QCIF, CIF, 4CIF, 720p et 1080p sont représentées. Cet ensemble a été défini dans le but d'avoir des contenus variés représentant à la fois des mouvements complexes (objets en mouvement et/ou caméra mobile) et uniformes. Les conditions de tests pour le profil Baseline IPP sont les suivantes : seule la première image est

codée en Intra, le codage entropique utilisé est le CAVLC, les blocs prédicteurs de la compensation de mouvement peuvent provenir de quatre images de référence et une fenêtre de recherche de 32×32 est autorisée pour l'estimation de mouvement. Les gains sur le débit total sont calculés avec la métrique de Bjøntegaard (Jung *et al.*, 2007) adoptée par le groupe VCEG. Pour la méthode MVComp et la méthode proposée, la compétition des vecteurs mouvement pour le mode Skip est activée.

4.2. Réduction de débit des indices de prédicteurs

Le tableau 1 donne la proportion des indices de prédicteurs de vecteurs mouvement qui sont transmis et ceux qui sont éliminés par la méthode proposée et par le test d'égalité des prédicteurs. On remarque que le pourcentage d'indices non transmis représente plus d'un quart des indices, atteignant près de 30 % pour le QP 17. Nous notons que l'égalité des prédicteurs n'est qu'un cas particulier du critère défini dans cet article. La réduction de débit des indices de prédicteurs dépasse 10 % pour les QP recommandés par le groupe VCEG. Cette réduction est plus élevée à haut débit, avec 13,6 % de réduction pour le QP 17 contre 8,8 % pour le QP 42. Le débit des indices de prédicteurs représente seulement 2,6 % du débit total, comme on a pu le constater dans les résultats expérimentaux de la méthode MVComp (Laroche, 2009). Par conséquent, les gains en réduction de débit total restent faibles avec respectivement 0,2 %, 0,3 %, 0,3 % de réduction de débit par rapport à MVComp pour les ensembles de QP représentant respectivement les hauts, moyens et bas débits. Par contre, le débit des indices représente près de 15 % du débit de l'information de mouvement. Par conséquent, un gain de 1,6 % est obtenu sur le débit de l'information de mouvement. L'efficacité de la méthode devrait augmenter avec un nombre plus élevé de prédicteurs car la proportion des indices de prédicteurs dans le débit total et dans le débit de l'information de mouvement augmente avec le nombre de prédicteurs. Par exemple, la proportion du débit des indices de prédicteurs représente 30 % du débit de l'information de mouvement lorsque 4 prédicteurs sont mis en compétition. Remarquons que la méthode MVComp pourrait s'appliquer conjointement avec une méthode de codage de vecteurs mouvement quantifiés comme nous l'avons étudié dans (Cagnazzo *et al.*, 2009). Dans ce contexte, si le pas de quantification des vecteurs mouvement est supérieur à 1, le nombre de prédicteurs égaux devrait augmenter. Par contre, dans le cas où le pas de quantification est inférieur à 1, le nombre de prédicteurs égaux devrait diminuer mais le nombre de prédicteurs implicites devrait, lui, augmenter.

4.3. Analyse

Le tableau 2 donne la proportion de sélection du prédicteur médian et du prédicteur collocated, ainsi que leur égalité en fonction des QP lorsque la méthode de suppression des prédicteurs est activée. Ces résultats sont à comparer avec ceux du tableau 3, qui présente ces pourcentages lorsque seule la méthode MVComp est utilisée. On remarque que la proportion de sélection diffère par rapport à la méthode MVComp. En effet, pour MVComp, le collocated est le prédicteur le plus sélectionné pour les trois derniers QP. Avec l'ajout de cette méthode, le prédicteur temporel est maintenant le plus sélectionné pour les QP représentant les moyens et bas débits.

Tableau 1. Proportion des indices de prédicteurs des vecteurs mouvement transmis et non transmis grâce au test d'égalité et ceux éliminés par le critère proposé, en fonction des six QP utilisés

QP	Indices transmis	Égalité des prédicteurs	Prédicteurs implicites
17	71,3	15,1	13,6
22	73,4	15,1	11,5
27	74,4	15,3	10,3
32	75,1	14,9	10,0
37	74,6	15,6	9,8
42	73,1	18,1	8,8
Moyenne	73,7	15,7	10,6

Cette proportion de sélection du prédicteur collocated augmente pour tous les QP. En moyenne, cette sélection est de 45,4 % contre 43,7 % dans la référence. L'impact de la méthode est relativement faible sur la sélection des modes de codage. De même, la proportion de la sélection des blocs prédicteurs dans les images de référence n'a pas évolué par rapport à la méthode MVComp. On peut considérer que l'augmentation de la proportion de prédicteurs collocated est probablement liée à l'évolution de la sélection du meilleur vecteur mouvement. En effet, cette sélection peut être modifiée par l'introduction de la méthode, car le critère de suppression est calculé pour chaque estimation du meilleur prédicteur, et donc de chaque estimation du débit de l'information de mouvement. On notera que la proportion moyenne d'égalité des prédicteurs a, elle aussi, augmenté, passant de 14,4 % pour la méthode MVComp à 15,7 %. Remarquons que l'augmentation du nombre de prédicteurs égaux augmente en fonction du QP à la fois pour MVComp et pour la méthode proposée. Ceci est lié à l'influence du débit de l'information de mouvement dans le critère débit distorsion car le paramètre de Lagrange augmente l'importance de ce débit dans le critère débit distorsion à bas débit.

Il est aussi intéressant de constater que l'efficacité de l'algorithme est corrélée avec les images de référence. Le tableau 4 donne le pourcentage d'élimination des indices supprimés par le test d'égalité et ceux supprimés par la méthode proposée, en fonction des images de référence. On remarque que le critère d'élimination des prédicteurs implicites évince une plus grande proportion de prédicteurs lorsque le prédicteur de bloc est sélectionné sur la première image de référence. La proportion de suppression ne cesse de diminuer lorsque la distance temporelle entre le bloc courant et le bloc prédicteur augmente. Pour la première image de référence, 30,2 % des indices ont été éliminés contre 17,1 % pour l'image de référence Ref_3 .

Les pourcentages d'élimination d'indices de prédicteurs obtenus avec la méthode sont naturellement liés au critère de choix du meilleur prédicteur défini dans l'équation (3). En effet, si on remplace dans cette équation la fonction d'estimation

Tableau 2. Proportion de sélection des prédicteurs $mv_{Scol_{R_0}}$ et $mv_{H.264}$, et de leur égalité lorsque la méthode proposée est activée, en moyenne pour l'ensemble des séquences et des six QP utilisés

QP	Prédicteur spatial $mv_{H.264}$	Prédicteur temporel $mv_{Scol_{R_0}}$	Égalité des prédicteurs (même valeur)
17	44,7	40,2	15,1
22	44,1	40,8	15,1
27	42,2	42,5	15,3
32	39,1	46,0	14,9
37	34,4	50,0	15,6
42	28,6	53,2	18,1
Moyenne	38,9	45,4	15,7

Tableau 3. Proportion de sélection des prédicteurs $mv_{Scol_{R_0}}$ et $mv_{H.264}$, et de leur égalité lorsque la méthode MVComp est activée, en moyenne pour l'ensemble des séquences et des six QP utilisés

QP	Prédicteur spatial $mv_{H.264}$	Prédicteur temporel $mv_{Scol_{R_0}}$	Égalité des prédicteurs (même valeur)
17	48,0	38,4	13,6
22	47,0	39,2	13,8
27	45,1	40,8	14,1
32	42,0	44,2	13,8
37	37,5	48,0	14,4
42	31,6	51,5	17,0
Moyenne	41,9	43,7	14,4

de coût $\zeta(x)$ qui est la fonction $Golomb(x)$ pour les résiduels de vecteurs mouvement dans le logiciel de référence, ces pourcentages d'élimination évoluent. Nous avons donc remplacé la fonction $\zeta(x)$ par les fonctions listées dans la section 3.4 pour les résiduels de vecteurs mouvement.

Les pourcentages d'élimination engendrés par chacune de ces fonctions, pour la séquence *Foreman* QCIF 15Hz codée avec le profil Baseline à QP 32, sont reportés

Tableau 4. Proportion de suppression des indices des prédicteurs avec l'égalité des prédicteurs et la méthode proposée et le pourcentage total d'indices non transmis en fonction des images de référence, en moyenne pour l'ensemble des séquences et des six QP utilisés

Images de référence	Égalité des prédicteurs (même valeur)	Prédicteurs implicites	Total indices non transmis
Ref_0	17,8	12,4	30,2
Ref_1	11,8	7,9	19,7
Ref_2	11,2	7,3	18,5
Ref_3	10,8	6,3	17,1

dans le tableau 5. L'ensemble des fonctions testées ont un pourcentage de suppression d'indices de prédicteurs supérieur à celui obtenu avec la fonction $Golomb(x)$ de référence. Pour les fonctions $|x|$, x^2 et e^x , la raison de cette supériorité est liée à leurs croissances. La figure 2 est la représentation graphique de l'ensemble des fonctions que nous avons testées. Si l'on exclut le cas de la fonction $floor(\log(|x|))$, les pourcentages de suppression sont ordonnés selon la rapidité de croissance de ces fonctions. Cependant, la fonction qui obtient la plus forte proportion de suppression de prédicteurs est la fonction $floor(\log(|x|))$ qui a justement la croissance la plus faible. Avec cette fonction, les coûts des deux résiduels ε_{mv_1} et ε_{mv_2} sont très souvent égaux, donc le prédicteur numéro 1 est souvent sélectionné. De même, les coûts de $mv_1 - p_1$ et $mv_1 - p_2$ ou $mv_2 - p_1$ et $mv_2 - p_2$ ont la même valeur dans de nombreux cas. Par conséquent, les équations (6) et (7) sont plus souvent vérifiées que si l'on utilise la fonction $Golomb(x)$.

Tableau 5. Proportion de suppression des indices des prédicteurs obtenue avec la méthode proposée, en fonction de plusieurs fonctions de sélection du meilleur prédicteur, pour la séquence Foreman QCIF 15Hz codée avec le profil Baseline à QP 32

Fonction de choix	Prédicteurs implicites
$Golomb(x)$	12,7
$ x $	14,7
x^2	16,9
e^x	30,8
$floor(\log(x))$	40,1

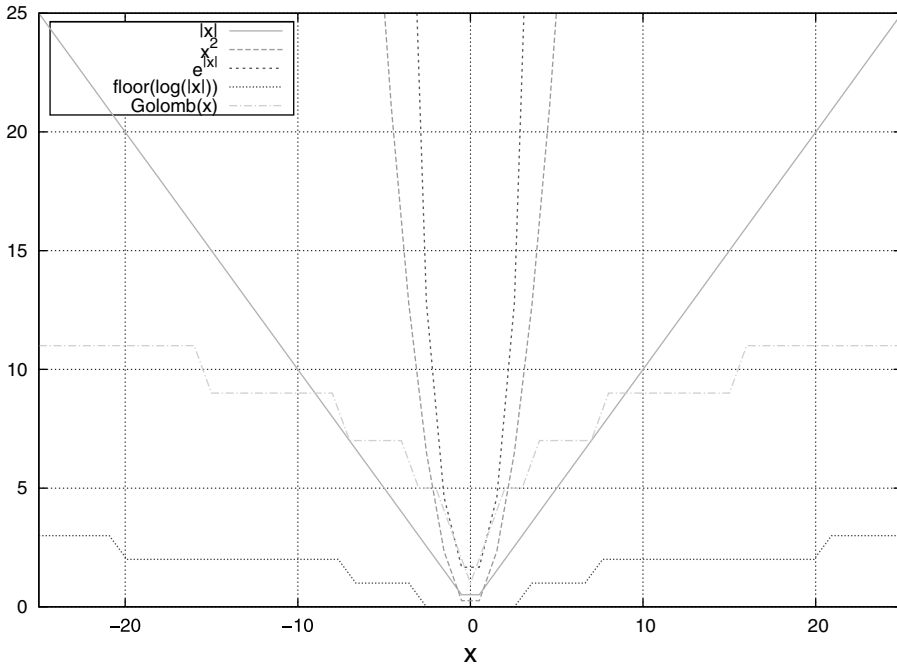


Figure 2. Représentation graphique des fonctions $\text{Golomb}(x)$, $|x|$, x^2 , e^x , $\text{floor}(\log(|x|))$

Bien que toutes les fonctions testées engendrent une proportion plus élevée d'élimination d'indices de prédicteurs, elles ne fournissent pas un choix optimal, au sens du critère débit-distorsion, du meilleur prédicteur du vecteur mouvement. Dès lors, la réduction de débit obtenue par la baisse du débit des indices de prédicteurs ne comble pas, en moyenne, le choix sous optimal du prédicteur du vecteur mouvement de chacune de ces fonctions.

5. Conclusion

L'information liée à la compétition des prédictions de vecteurs de la méthode MVComp proposée dans (Laroche *et al.*, 2008) représente en moyenne 15 % du débit de l'information de mouvement. Dans cet article, nous avons proposé de réduire cette information de compétition en utilisant une méthode de suppression d'indice de prédicteurs. Dans cette méthode, le critère de sélection du meilleur prédicteur calculé par l'encodeur, est utilisé au décodeur afin de déterminer des prédicteurs implicites grâce aux caractéristiques du codage sans perte de l'information de mouvement. La méthode permet de supprimer plus de 10 % des indices de prédicteurs. Si l'on ajoute à ce pourcentage les indices supprimés par le critère d'égalité des prédicteurs, plus d'un quart de l'information de compétition des prédicteurs n'est pas transmise au décodeur. Cette information a cependant une faible proportion dans le débit total

lorsque seulement deux prédicteurs sont en compétition. Par conséquent, l'impact de cette modification est faible sur l'efficacité globale de codage. Cependant, l'utilisation d'un nombre plus élevé de prédicteurs de vecteurs mouvement devrait augmenter l'efficacité de la méthode.

La réduction de l'information de compétition est une voie de recherche essentielle pour l'amélioration incrémentale de la norme de référence H.264/AVC, car cette information représente plus de la moitié du débit total pour les moyens et bas débits. L'augmentation de l'intelligence du décodeur, par connaissance des choix de codage, tel que présenté dans cet article, est une manière d'y parvenir.

Bibliographie

Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC) n.d.

H.264/AVC software coordination n.d.

KTA Software coordination n.d.

Artigas X., Ascenso J., Dalai M., Klomp S., Kubasov D., Ouaret M. (2007), « The Discover Codec : Architecture, Techniques And Evaluation », *Picture Coding Symposium (PCS)*, Lisbonne, Portugal, Novembre.

B.G., Aaron A., Rane S., Rebollo-Monedero D. (2005), « Distributed Video Coding », *IEEE, Special ISSUE on Advances in Video Coding and Delivery*, vol. 93, n° 1, p. 71-83, Janvier.

Cagnazzo M., Agostini M., Antonini M., Laroche G., Jung J. (2009), « Motion vector quantization for efficient low bit-rate video coding », *Proc. SPIE Conf. on Visual Communications and Image Processing (VCIP)*, San jose, Californie, USA, Janvier.

Chen C., Willson A. (1997), « A spatial and temporal motion vector coding algorithm for low-bit-rate video coding », *Image Processing, ICIP, IEEE International Conference on*, vol. 2, Santa Barbara, Californie, USA, p. 791-794, Octobre.

Chenn P., Ye Y., Karczewicz M. (2009), Video Coding Using Extended Block Sizes, Proposition n° VCEG-COM16-C123, ITU-T VCEG, Genève, Suisse, Février.

Chujoh T., Noda R. (2007), Internal bit depth increase for coding efficiency, Proposition n° VCEG-AE13, ITU-T VCEG, Marrakech, Maroc, Janvier.

Davies T., Andersson K., Sjöberg R., Wiegand T., Marpe D., Ugur K., Ridge J., Karczewicz M., Chen P., Martin-Cocher G., McCann K., Han W., Bjøntegaard G., Fuldseth A. (2010), Suggestion for a Test Model, Proposition n° JCTVC-A033, JCT-VC, Dresden, Allemagne, Avril.

De Natale F., Granelli F. (2001), « Low-complexity context-based motion compensation for VLBR video encoding », *Multimedia & Expo, ICME, IEEE International Conference on*, Tokyo, Japon, p. 439-442, Août.

Deuk Kim S., Beom Ra J. (1999), « An efficient motion vector coding scheme based on minimum bitrate prediction », *IEEE Trans.on Image Processing*, vol. 8, n° 8, p. 1117-1120.

- Joshi R., Fischer T., Bamberger R. (1995), « Lossy encoding of motion vectors using entropy-constrained vector quantization », vol. 3, p. 109-112, Octobre.
- Jung J., Laroche G. (2006), Competition-Based Scheme for Motion Vector Selection and Coding, Information n° VCEG-AC06, ITU-T VCEG, Klagenfurt, Autriche, Juillet.
- Jung J., Pateux S. (2007), An excel add-in for computing Bjøntegaard metric and its evolution, Information n° VCEG-AE07, ITU-T VCEG, Marrakech, Maroc, Janvier.
- Kamp S., Evertz M., Wien M. (2008), « Decoder side motion vector derivation for inter frame video coding », *Image Processing, ICIP, IEEE International Conference on*, vol. 1, San Diego, USA, p. 1120-1123, Octobre.
- Karczewicz M., Ye Y., Chong I. (2008), Rate Distortion Optimized Quantization, Proposition n° VCEG-AH21, ITU-T VCEG, Antalya, Turquie, Janvier.
- Kim S., Yang J., Won K., Jeon B., (2009) Improved Motion Vector Coding with Optimal PMV Selection, Proposition n° VCEG-COM16-C122, ITU-T VCEG, Genève, Suisse, Février.
- Laroche G. (2009), Modules de codage par compétition et suppression de l'information de compétition pour le codage de séquences vidéo, PhD thesis, École Nationale Supérieure des Télécommunications, France, Mai.
- Laroche G., Jung J., Pesquet-Popescu B. (2008), « RD Optimized Coding for Motion Vector Predictor Selection », *IEEE Trans.on Circuits and System for Video Technology*, vol. 18, n° 9, p. 1247-1257, Septembre.
- Ohm J., Sullivan G. (2010), Joint Call for Proposals on Video Compression Technology, Call for Proposals n° VCEG-AM91, ITU-T VCEG and ISO/IEC MPEG, Kyoto, Japon, Janvier.
- Puri R., Ramchandran K. (2002), « PRISM : A New Robust Video Coding Architecture Based on Distributed Compression Principles », *Proc.Of 40th Allerton Conference on Communication, Control, and Computing*, Allerton, USA, Octobre.
- Rusanovksyy D., Ugur K., Hallapuro A., Lainema J. (2008), Implementation of Directional Interpolation Filter with 16-bit Integer Arithmetic, Proposition n° VCEG-AH18, ITU-T VCEG, Antalya, Turquie, Janvier.
- Rusanovksyy D., Ugur K., Lainema J. (2007), Adaptive Interpolation with Directional Filters, Proposition n° VCEG-AG21, ITU-T VCEG, Shenzhen, Chine, Octobre.
- Sugimoto K., Kobayashi M., Suzuki Y., Kato S., Boon C.S. (2004), « Inter frame coding with template matching spatio-temporal prediction », *Image Processing, ICIP, IEEE International Conference on*, vol. 1, Singapour, p. 465-468, Octobre.
- Sullivan G., Wiegand T. (1998), « Rate-distortion optimization for video compression », *IEEE Signal Processing Magazine*, vol. 15, n° 6, p. 74-90, Novembre.
- Tan T., Sullivan G., Wedi T. (2008), Recommended Simulation Common Conditions for Coding Efficiency Experiments Revision 2, Input /Discussion n° VCEG-AH10r3, ITU-T VCEG, Antalya, Turquie, Janvier.
- Tourapis A., Wu F., Li S. (2005), « Direct mode coding for bipredictive slices in the H.264 standard », *IEEE Trans.on Circuits and System for Video Technology*, vol. 15, n° 1, p. 119-126, Janvier.
- Vatis Y., Edler B., Wassermann I., Nguyen D., Ostermann J. (2005), Two-dimensional non-separable Adaptive Wiener Interpolation Filter for H.264/AVC, Proposition n° VCEG-Z17, ITU-T VCEG, Busan, Corée du Sud, Avril.

- Vatis Y., Ostermann J. (2006), Prediction of P-and B-Frames Using a Two-dimensional Non-separable Adaptive Wiener Interpolation Filter for H.264/AVC, Proposition n° VCEG-AD08, ITU-T VCEG, Hangzhou, Chine, Octobre.
- VCEG I.-T., MPEG I. (2010), Terms of Reference of the Joint Collaborative Team on Video Coding Standard Development, Terms of reference, Kyoto, Japon, Janvier.
- Wien M., Chiu Y. (2010), Tool Experiment 1 : Decoder-Side Motion Vector Derivation, Tool Experiment Description n° JCTVC-A031, JCT-VC, Dresden, Allemagne, Avril.
- Wyner A., Ziv J. (1976), « The rate-distortion function for source coding with side information at the decoder », *IEEE Trans.on Infomation Theory*, vol. IT-22, n° 1, p. 1-10, Janvier.
- Ye Y., Karczewicz M. (2007), Improved Intra Coding, Proposition n° VCEG-AG11, ITU-T VCEG, Shenzhen, Chine, Octobre.
- Yeh J., Vetterli M., Khansari M. (1995), «Motion compensation of motion vectors », *Image Processing, ICIP, IEEE International Conference on*, vol. 1, Washington, District de Columbia, USA, p. 574-577, Octobre.

Article reçu le 11/06/2010
Accepté le 18/10/2010



Guillaume Laroche a reçu le titre de docteur de Télécom ParisTech, spécialité Signal et Images en 2009. Cette thèse a été menée en partenariat avec Orange Labs et a abouti à l'introduction d'une technologie dans le « draft » du prochain standard de compression video HEVC. Ces recherches portent donc sur la compression vidéo. Il est actuellement Ingénieur de Recherche chez Canon Research centre France dans ce domaine.



Joël Jung est diplômé d'une thèse de Doctorat reçue de l'Université de Nice-Sophia Antipolis, Nice, France, en 2000. De 1996 à 2000, il a contribué à l'amélioration de la qualité du décodage vidéo, en traitant les artefacts de compression, au sein du Laboratoire I3S du CNRS à Sophia Antipolis. Il a rejoint Philips Research France en 2000 en tant qu'Ingénieur de Recherche, travaillant sur le codage vidéo, les post-traitements, les métriques de qualité et le codage basse consommation. Il est actuellement Ingénieur de Recherche chez Orange Labs Paris. Il est un participant actif à la normalisation ITU-T V ISO MPEG, avec plus de 40 contributions aux différents standards. Il participe à la normalisation de HEVC, et coordonne des « Core Experiments ». Il est l'auteur de 30 articles publiés dans des revues et conférences internationales et détient 35 brevets dans le domaine de la compression d'image et de vidéo.



Béatrice Pesquet-Popescu est Professeur à Télécom ParisTech, Responsable du Groupe Multimédia au sein de Telecom ParisTech, et Directrice Scientifique du Laboratoire Commun avec Alcatel-Lucent, Ubimedia. Ses activités de recherche portent sur la compression vidéo scalable robuste et les représentations adaptées au contenu. Elle est auteur de plus de 180 articles publiés dans des revues et conférences internationales et de plus d'une vingtaine de brevets dans le domaine. Elle a également la responsabilité du Groupe de Travail « Médias du Futur » de l'Institut Télécom, et au niveau international est membre de l'Administrative Committee de l'EURASIP, rédacteur associé de plusieurs revues internationales, telles que IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology, Image Communication, International Journal of Digital Multimedia Broadcasting, ainsi que membre des comités techniques IEEE SPS Image, Video and Multidimensional Signal Processing (IVMSP) et membre associé de Multimedia Signal Processing (MMSP TC). Par ailleurs, elle a été ou sera amenée à assumer le rôle de General Chair de plusieurs manifestations internationales telles que IEEE Multimedia Signal Processing (MMSP) 2010, European Signal Processing Conference (EUSIPCO) 2012, IEEE International Conference on Image Processing (ICIP) 2014.

