

UN CODEUR HYBRIDE H264 - MAILLAGE

A. Buisson¹

N. Laurent¹

¹ France Telecom R&D DIH/HDM/CIM

4 Rue du Clos Courtel 35512 Cesson-Sevigne FRANCE

{alexandre.buisson, nathalie.laurent}@rd.francetelecom.com

Résumé

Afin de surmonter le problème de la surcharge des réseaux, il convient de fournir des solutions de codage de plus en plus efficaces de l'information pour tous les média et en particulier pour la vidéo. La plupart des codecs actuels s'articulent autour d'un schéma basé blocs (DCT+BMA). Cependant, au cours des 20 dernières années de nombreuses techniques ont été introduites afin d'améliorer la qualité des vidéos codées et de proposer des solutions à bas débits. Parmi celles-ci, les maillages constituent des outils intéressants pour le codage et la représentation du mouvement. Dans cet article, nous présentons donc une hybridation du codeur vidéo H264 avec des outils de codage par maillages afin d'améliorer le codage des images P et B en qualité et/ou débit. Nous avons pu valider cette approche par une série de tests sur diverses séquences.

Mots clefs

Codage vidéo, H264, Maillages, Bas débits.

1 Introduction

De nombreuses études antérieures [1, 2] ont montrées que les techniques basées maillages offrent de bonnes performances ainsi que des fonctionnalités intéressantes. Cependant, les outils dont nous disposons ne sont pas encore assez matures pour permettre la mise en œuvre d'un schéma de codage uniquement basé maillages. Une voie envisageable est donc d'hybrider ces techniques avec un codec vidéo existant. En effet, le codec hybride ainsi construit, nous permettra d'évaluer et de démontrer les potentiels des maillages, pour améliorer la qualité des vidéos codées et pour atteindre de très bas débits. Nous avons déjà réalisé dans [3] une première version d'hybridation, mais l'efficacité de la technique proposée restait limitée en raison de la décorrélation entre les modèles basés blocs et maillages. Afin de résoudre ce problème, nous réalisons une hybridation plus complète des maillages avec le codec le plus récent et le plus performant, c'est à dire H264. La section 2 présente brièvement le codeur vidéo H264 et le concept général de son hybridation avec les maillages. Nous décrivons en détail dans la section 2.1 l'amélioration du codage des images P, puis nous présentons dans la sec-

tion 2.2 la mise en œuvre d'une amélioration du codage des images B. Les performances du codec développé et les perspectives de cette étude sont présentées dans la section 3.

2 Description

A l'heure actuelle un nouveau codec appelé H264, détaillé dans [4], est développé par l'ITU-T. Celui-ci se montre plus performant que toutes les autres solutions existantes comme l'ont montré des tests subjectifs réalisés en 2001 à la demande du groupe de travail MPEG-Visual. La base de ce codec repose sur un mécanisme d'optimisation débit-distortion (RD) décrit dans [5] qui permet de réaliser des choix de codage très efficaces. De plus, afin d'incrémenter ses performances H264 utilise 7 types de macroblocs et un buffer de 5 images de référence. Il exploite celles-ci au cours de sa procédure d'optimisation RD afin de déterminer pour chaque macrobloc à coder le mode de codage optimal (forme, mouvement, référence).

Afin d'hybrider les maillages avec H264, nous avons inséré dans ce buffer une image supplémentaire qui correspond à une approximation grossière de l'image à coder obtenue par une estimation-compensation de mouvement basée maillages. Ce procédé repose sur des maillages hiérarchiques emboîtés et sur une estimation de mouvement hiérarchique et multirésolution décrite dans [1, 2]. L'algorithme 1 résume les différentes étapes qui composent la brique basée maillages. On réalise une estimation de mouvement avant et une estimation arrière entre l'image courante et la dernière image codée. On réalise alors un choix entre la prédiction avant $I(\vec{d}_F)$ et la prédiction arrière $I(\vec{d}_B)$ en se basant sur le PSNR des deux images compensées obtenues. Aucune information n'est nécessaire pour indiquer la direction d'estimation choisie, car nous avons réalisé une hypothèse sur la continuité et la bijectivité du champ de mouvement. Ainsi, \vec{d}_B peut-être remplacé par $-\vec{d}_B$. Cette hypothèse, bien que non effective, fonctionne fort bien et nous a permis d'améliorer de manière considérable les coûts des informations de mouvement et également les PSNR des images maillages.

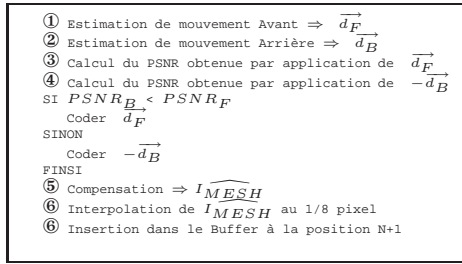


Figure 1 – Algorithme de la procédure basée maillages

Les mouvements obtenus pour les nœuds de la hiérarchie de maillages sont codés en différentiel et une technique de prédiction des vecteurs mouvements du niveau de base a été développée. Celle-ci exploite la corrélation (en amplitude et en direction), des paramètres de mouvement du niveau 0 de la hiérarchie qui modélisent un déplacement global, d’une estimation à l’autre.

La procédure basée maillages est immédiatement suivie du codage H264 hybride d’images P qui exploite notre image de référence supplémentaire. Mais cette hybridation résumée par la figure 2 a impliqué la modification d’un certain nombre d’algorithmes du codec H264 afin de permettre au cours du processus de décision de choisir pour un macrobloc de forme i de le coder soit avec le module H264, soit avec le module basé maillages, soit avec une solution hybride combinant les maillages et H264.

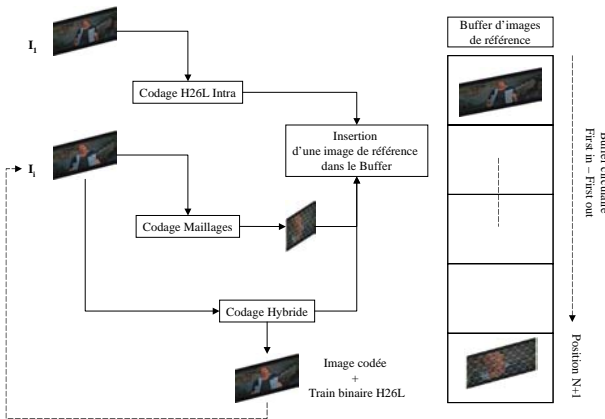


Figure 2 – Synopsis de la mise en œuvre du codeur hybride par le biais des maillages.

2.1 Hybridation pour les images P

Le principe de l’optimisation RD H264 reste inchangé, mais certains algorithmes ont été modifiés et de nouvelles fonctionnalités ont été ajoutées. Ainsi, on dispose maintenant d’un mode $COPY_{MESH}$ qui fait référence à l’image basée maillages. De plus, les modes Inter H26L

($16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$) ont la possibilité de prendre notre image N+1 comme référence, ce qui correspond à un raffinement de la prédiction basée maillages par le biais du codec H264.

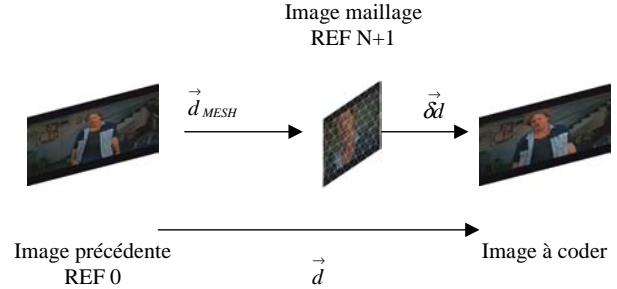


Figure 3 – Types de vecteurs utilisés dans notre approche

Toutefois, afin de favoriser le choix de la référence N+1 dans la procédure d’optimisation RD, nous avons fait en sorte de la confondre avec la référence 0. Cela nous a conduit à modifier la procédure de prédiction des vecteurs de mouvement. En effet, celle-ci doit tenir compte des 3 types de déplacements présentés par la figure 3. Les vecteurs $\vec{\delta d}$ sont des vecteurs de raffinement de la prédiction basée maillages. Les vecteurs \vec{d} sont des vecteurs H264 classiques ayant pour référence l’image 0. Et les déplacements d_{MESH} correspondent aux vecteurs de déplacement moyen issu du champ dense généré par les maillages. Ainsi, si la prédiction des vecteurs mouvement reste toujours basée sur la méthode décrite dans [4], divers cas présentés dans la figure 4 peuvent se produire et la gestion de ceux-ci nécessite de classer les blocs que nous utilisons :

- Un *bloc maillage* sera un bloc Inter H264 qui fera référence à l’image $N + 1$.
- Un *bloc H26L* sera un bloc Inter H264 qui fera référence à une de N premières images du buffer et dont le mode sera soit INTER, soit INTRA.
- Enfin, les *blocs COPY* ou $COPY_{MESH}$ feront référence à l’images 0 ou N+1 du buffer et ne nécessiteront pas de prédiction.

Ainsi, dans le cas des prédictions par un des blocs A,B,C ou D présentées en 4(a) la prédiction peut-être donnée par :

- $\vec{d}_{MESH} + \vec{\delta d}$ si E est H264 et le prédicteur Maillages.
- $\vec{d} - d_{MESH}$ si E est Maillage et le prédicteur H264.
- $\vec{\delta d}$ si E et le prédicteur sont maillages.

De plus, dans le cas où le médian est utilisé, il faut prendre en compte les configurations décrites en figure 4(b) et les prédictions seront définies par :

- le médian des $\vec{\delta d}$ des blocs maillages voisins et des $\vec{d} - d_{MESH}$ des blocs H26L voisins si E est Maillages.
- le médian des \vec{d} des blocs voisins H26L et des

$\vec{\delta d} + \vec{d}_{MESH}$ des blocs voisins maillages si E est H264.

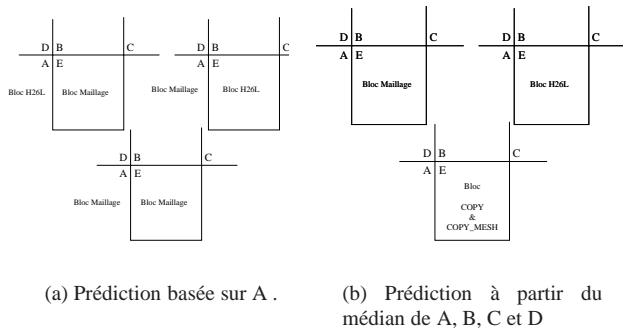


Figure 4 – Modification des prédictions par un vecteur ou par le médian dans plusieurs cas spécifiques.

2.2 Hybridation pour les images B

Dans le cas d'un modèle de codage IBP, les maillages constituent un outil très puissant. En effet, [3] nous a permis de montrer que champ de mouvement basé maillages estimé entre 2 images P pouvait permettre le codage d'images B pour quelques bits. Cette fonctionnalité nous permet d'améliorer le processus de codage des images B, car en disposant d'un champ basé maillages (entre la P précédente et la P suivante), notre codec a la possibilité d'utiliser deux techniques :

- coder des images B uniquement basées maillages,
- coder des images B de type H264, en exploitant tout de même un certain nombre d'informations provenant des maillages.

Comme le montre la figure 5, le codage des images B est désormais composé d'une procédure d'interpolation temporelle de trame basée maillages ③, qui peut-être suivie ou non d'un codage H264 hybride ⑤ en fonction du critère mesuré dans la procédure de choix ④.

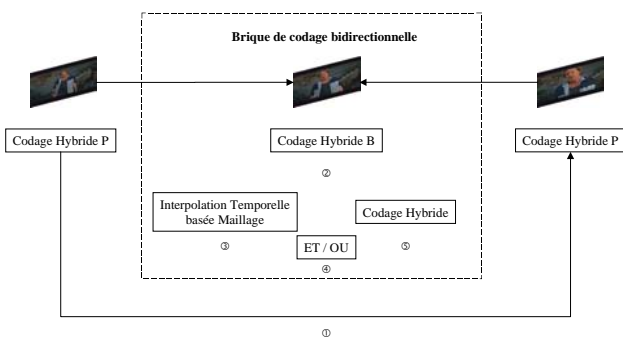


Figure 5 – Schéma de l'hybridation des maillages avec H264 pour le codage des images B

Cette méthode de choix permet, si la qualité des prédictions basées maillages est suffisante, de représenter une ou plusieurs images B avec un coût de quelques bits. Dans le cas contraire, un codage H264 hybride est réalisé. Nous avons en effet modifié l'algorithme de détermination des mouvements MV_F et MV_B du mode DIRECT, mais également ajouté un mode $COPY_{MESH}$ qui fera référence l'image notée I_{B_MESH} obtenue par le biais du mécanisme ③.

L'étape de détermination des mouvements MV_F et MV_B du mode DIRECT est modifiée si le macrobloc de forme i dans l'image P suivante a été codé à partir de l'image basée maillages $N + 1$ du buffer. On peut donc résumer la nouvelle procédure par l'algorithme 6 dans lequel MV est le mouvement estimé par H264, MV_{MESH} le déplacement moyen déterminé à partir du champ dense et TR_B, TR_D les coefficients d'interpolation temporelle.

```

SI le macrobloc courant a été codé pour
l'image P suivante en référence à l'image
basée maillages N+1
ALORS
   $MV_F = (TR_B \times (MV + MV_{MESH}))/TR_D$ 
   $MV_B = (TR_B - TR_D) \times (MV + MV_{MESH})/TR_D$ 
SINON
   $MV_F = (TR_B \times MV)/TR_D$ 
   $MV_B = (TR_B - TR_D) \times MV/TR_D$ 
FINSI
  
```

Figure 6 – Algorithme de calcul des vecteurs de mouvement avant et arrière du mode DIRECT

3 Résultats et conclusions

On présente ici les résultats des comparaisons en terme de débit et de PSNR du codec H264 et du codec hybride. On étudie en effet les performances de ces 2 codecs à qualité équivalente et à même débit. Tous les tests ont été réalisés avec le codec H26L TM8.5 et sa version hybride. Nous avons effectué des tests de codage de la séquence "FOREMAN" au format CIF avec différents pas de quantification. Dans le tableau 1, on peut voir les résultats de ces codages à qualité quasi équivalente en terme de PSNR. On observe que notre codec hybride génère toujours un débit inférieur et que le gain augmente lorsque l'on se rapproche des très bas débits ce qui confirme l'intérêt de notre approche hybride et plus généralement celui des maillages pour le codage à très bas débits. Toutefois, il faut constater également qu'à tous les pas de quantification, les séquences codées avec le codec hybride présentent un PSNR légèrement inférieur à celui des séquences H264 par définition de toute optimisation RD Lagrangienne.

Afin de valider notre approche, nous avons réalisé des tests sur différents types de séquences, films et dessins animés. Pour ce faire, nous avons déterminé quels sont les quantificateurs à utiliser avec notre codec pour obtenir un débit équivalent à celui du codec H264. Le codage de la séquence "GORILLAZ" montre qu'afin d'obtenir un débit

	Codeur H26L		Codeur hybride		Gains (Kbits)	Ratio (%)
	PSNR	Débit (bits)	PSNR	Débits (bits)		
Q25	31.88	287088	31.71	269680	18	7
Q26	31.31	256184	31.07	233840	23	9
Q27	30.60	226296	30.37	202224	24	11
Q28	30.03	198360	29.63	171528	27	14
Q29	29.43	178488	28.90	146448	32	18
Q30	28.53	152416	27.98	118944	37	23

Tableau 1 – Comparaison des débits générés par le codec H26L et par notre codec hybride sur le codage des 50 premières images de la séquence FOREMAN CIF à différents pas de quantification. On utilise un modèle de codage IBP avec le même pas de quantification pour toutes les images.

égal à celui du codage H264 avec des pas de quantification de 30, nous pouvons avec notre codec utiliser des pas de quantification de 28. Sur cet exemple, notre technique hybride améliore en moyenne de 2dB la qualité de la vidéo codée comme le montre la figure 7. On peut également observer une réduction des défauts des vidéos codées comme par exemple sur les images présentées dans le tableau 2.

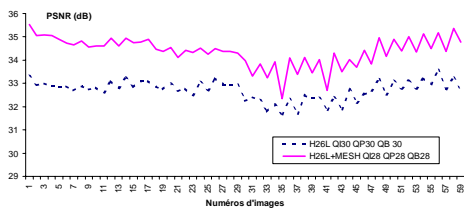


Figure 7 – Performances du codec H264 et du codec hybride (H26L+Maillages) comparées à débit équivalent sur la séquence GORILLAZ au format CIF. Le débit généré est de 37Kbits/s à une fréquence de 25Hz. Notre codec réussit à générer une séquence avec une qualité moyenne de 34,38 dB, soit un gain d'environ 2dB vis à vis H264.

De plus, les performances obtenues sur la séquence SUSIE présentées par la figure 8 proviennent en partie du codage hybride d'images B. En effet, sur cette séquence, 9 images B ont été codées avec 1bit en utilisant l'interpolateur d'images basé maillages. Ce codec hybride développé en détail dans [6] constitue donc une solution efficace pour atteindre de très bas débits ou pour fournir une amélioration de la qualité visuelle. Il faut également noter que de nombreuses optimisations sont encore envisageables sur la brique basée maillages, en terme de qualité et de vitesse de calculs, et sur l'hybridation efficace des modèles dans un codeur vidéo utilisant une optimisation RD.

Références

[1] P. Lechat, N. Laurent, et H. Sanson. Représentation d'images et estimation de mouvement basées maillage.



H264

Codec Hybride

Tableau 2 – Zoom sur les détails et sur la réduction des artefacts sur les séquences FOREMAN et GORILLAZ CIF

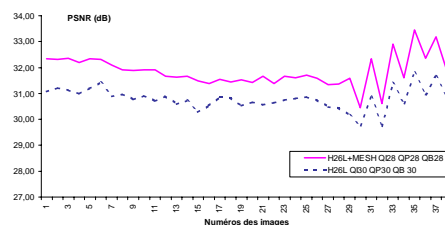


Figure 8 – Performances à débit équivalent du codec H264 et du codec hybride sur la séquence SUSIE CIF. Le débit généré est de 40Kbits/s à une fréquence de 25Hz. Le débit du codec hybride se compose de 57672 bits pour H264, de 3976 bits de mouvements maillages et de 256 bits provenant du codage des niveaux de hiérarchie utilisés pour chaque image.

application à un codeur tout-maillage. Actes de la conférence CORESA'98, 1998.

[2] N. Laurent. Hierarchical mesh-based global motion estimation, including occlusion areas detection. Dans *Proceedings of ICIP'00*, Vancouver, 2000.

[3] A. Buisson et N. Laurent. Interpolation temporelle de trames et codage bidirectionnel pour les séquences vidéos. Dans *Actes de la conférence CORESA'01*, Poitiers, 2001.

[4] T. Wiegang. *H26L test model long term Number 9*. ADOC VCEG-N83 d1, H26L Web Site, 2001.

[5] G. Sullivan et T. Wiegang. Rate-distortion optimization for video compression. *IEEE Signal Processing Magazine*, 15 :74–98, 1998.

[6] A. Buisson. *Implémentation efficace d'un codeur vidéo hiérarchique granulaire sur une architecture à processeurs multimédias*. Thèse de doctorat, Université de Rennes 1, 2002.