



HAL
open science

Compression de contenu vidéo Super Multi-Vue avec parallaxe horizontale et verticale

Antoine Dricot, Joël Jung, Marco Cagnazzo, Beatrice Pesquet-Popescu, Frederic
Dufaux

► **To cite this version:**

Antoine Dricot, Joël Jung, Marco Cagnazzo, Beatrice Pesquet-Popescu, Frederic Dufaux. Compression de contenu vidéo Super Multi-Vue avec parallaxe horizontale et verticale. Colloque COMpression et REprésentation des Signaux Audiovisuels (CORESA'2014), Nov 2014, Reims, France. <hal-01436168>

HAL Id: hal-01436168

<https://hal.science/hal-01436168v1>

Submitted on 10 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Compression de contenu vidéo Super Multi-Vue avec parallaxe horizontale et verticale

A. Dricot^{1,2}, J. Jung¹, M. Cagnazzo², B. Pesquet², et F. Dufaux²

¹Orange Labs

²Institut Mines-Télécom;
Télécom ParisTech; CNRS LTCI

Résumé

La vidéo Super Multi-Vue (SMV) est une technologie clé pour mettre en place les futurs services de vidéo 3D. Le SMV permet une visualisation sans lunette et élimine beaucoup des causes d'inconfort présentes dans les technologies de vidéo 3D actuelles. Le contenu vidéo SMV est composé de dizaines ou de centaines de vues d'une scène, qui peuvent être alignées soit uniquement dans la direction horizontale, soit dans les directions horizontale et verticale. Cet article compare plusieurs schémas de codage, puis propose une structure de codage qui exploite les corrélations inter-vues dans les deux directions, permettant de réduire le débit (pour une qualité donnée) de 29.1% par rapport à une structure de référence basique. De plus, une amélioration des outils de codage Neighboring Block Disparity Vector (NBDV) et Inter-View Motion Prediction (IVMP) est proposée afin d'exploiter efficacement les structures de codage en deux dimensions, avec une réduction de débit allant jusqu'à 4.2% par rapport à l'encodeur référence 3D-HEVC.

Super Multi-View video (SMV) is a key enabler for future 3D services. SMV allows a glasses-free viewing and eliminate many causes of discomfort existing in current 3D video technologies. SMV content consists of tens or hundreds of views of a scene, that can be aligned along the horizontal direction or both along horizontal and vertical directions. This paper compares several coding schemes, and proposes a coding structure that exploits inter-view correlations in horizontal and also vertical directions. This new structure provides a rate reduction (for the same quality) up to 29.1% when compared to a basic anchor structure. Neighboring Block Disparity Vector (NBDV) and Inter-View Motion Prediction (IVMP) coding tools are further improved to efficiently exploit coding structures in two dimensions, with rate reduction up to 4.2% with respect to the reference 3D-HEVC encoder.

Mots clé : Compression vidéo 3D, multi-vue, parallaxe de mouvement

1. Introduction

Le développement des technologies liées à la vidéo 3D tend à créer des expériences de visualisation de plus en plus immersives. Cependant, les technologies vidéo 3D actuellement disponibles sur le marché présentent plusieurs limitations [DPC13]. Avec la stéréoscopie 3D, le manque de confort dû au port des lunettes est combiné à des stimuli de perception qui ne sont pas naturels, comme le conflit entre vergence et accommodation, et qui peuvent causer des douleurs oculaires ainsi que des migraines. Avec les systèmes auto-stéréoscopiques sans lunette, le nombre réduit de vues ne permet pas d'avoir une parallaxe de mouvement fluide (c'est à dire que la visualisation n'est pas continue quand on bouge devant l'écran) et restreint la taille de la zone de visualisation, ce qui altère particulièrement la qualité et le confort de l'expérience de visualisation.

Une étude de la vidéo Super Multi-Vue (SMV) a été initiée

pendant le meeting MPEG FTV d'Octobre 2013 [TSO* 13a]. Le SMV utilise des dizaines ou des centaines de vues afin de créer une représentation du *light-field* d'une scène. Le *light-field* représente en principe tous les rayons de lumière dans une scène en 3D. C'est donc une fonction de deux angles (direction du rayon) et de trois coordonnées spatiales. Cette fonction en 5 dimensions est la fonction plénoptique [AB91]. Beaucoup des artefacts existants dans les technologies 3D actuelles peuvent être éliminés avec une représentation en *light-field*, en particulier le conflit vergence-accommodation. Elle permet une visualisation réaliste, sans lunette, et avec une parallaxe de mouvement (qui est un élément clé dans la perception du relief) fluide dans la direction horizontale et potentiellement dans la direction verticale. Plusieurs entreprises ont déjà démontré de l'intérêt pour le SMV en travaillant sur des écrans et des systèmes d'affichage dits *light-field*. La visio-conférence immersive est présentée comme un cas d'utilisation cible typique, ainsi que la diffusion live d'événements sportifs en 3D, comme les Jeux Olympiques de 2020 au Japon, qui pourrait être filmés

par des ensembles de caméras et projetés sur des écran SMV géants dans les lieux publics de plusieurs grandes villes dans le monde [TSO* 13c]. Il existe donc une demande et un besoin pour des nouvelles technologies de codage efficaces qui peuvent traiter la grande quantité de données nécessaire pour le SMV [TSO* 13b].

Les extensions multi-vues des encodeurs standards peuvent permettre d'encoder du contenu SMV avec parallaxe horizontale. Des modifications de ces encodeurs sont proposées dans la littérature scientifique pour encoder du contenu avec parallaxe horizontale et verticale. Les méthodes de l'état de l'art présentent cependant des limitations dans leur utilisation des deux dimensions pour la prédiction inter-vues. On propose ici un schéma de prédiction inter-vues efficace pour exploiter les dimensions horizontale et verticale au niveau de la structure de codage. On propose ensuite des améliorations d'outils de codage inter-vues pour exploiter les structures en deux dimensions également au niveau pixellique.

Le reste de cet article est organisé comme suit. Des méthodes d'acquisition d'un ensemble de vues avec parallaxe horizontale et verticale sont décrites dans la Section 2. La Section 3 décrit des méthodes de l'état de l'art pour le codage de ce contenu SMV avec parallaxe dans les deux dimensions. Dans la Section 4, on décrit le schéma de prédiction inter-vues proposé et on montre des résultats expérimentaux contre les schémas de l'état de l'art. Les outils de codage inter-vues améliorés et adaptés à la parallaxe en deux dimensions sont proposés dans la section 5, qui contient également des résultats expérimentaux. La Section 6 conclut finalement cet article.

2. Acquisition et représentation du light-field

Une représentation du light-field d'une scène peut être obtenue à partir de plusieurs images captées depuis différents points/angles de vue. On considère ici deux technologies qui permettent d'obtenir du contenu avec parallaxe dans les directions horizontale et verticale : l'acquisition multi-caméras et l'imagerie intégrale.

Un contenu SMV peut être filmé avec une matrice de caméra comme le montre la Figure 1. Ces caméras peuvent être alignées horizontalement (donnant une parallaxe horizontale uniquement) ou horizontalement et verticalement dans le cas du contenu avec parallaxe en deux dimensions. Ce système de caméra peut être arrangé en linéaire, en arc ou même de manière aléatoire. Chaque caméra capte la scène d'un point de vue différent et le contenu résultant est un ensemble de vues avec des disparités horizontales et verticales.

L'imagerie intégrale (ou holoscopie) est une technique basée sur la photographie plénoptique [Lip08]. L'acquisition plénoptique est basée sur l'utilisation d'un panneau lenticulaire placé devant une caméra. Ce dispositif est schématisé dans la Figure 1. Le panneau lenticulaire est composé d'un grand nombre de microlentilles, pouvant avoir une forme ronde ou carrée, et pouvant être alignées en grille ou en quinconce. L'image holoscopique résultant de cette captation est une matrice de Micro-Images (MIs). Chaque microlentille produit une MI, et chaque MI contient de l'information sur la scène provenant de différents angles de vues. La connexion entre une image intégrale et un ensemble de vue d'une scène peut être tracée par l'extraction de vues, comme le montre

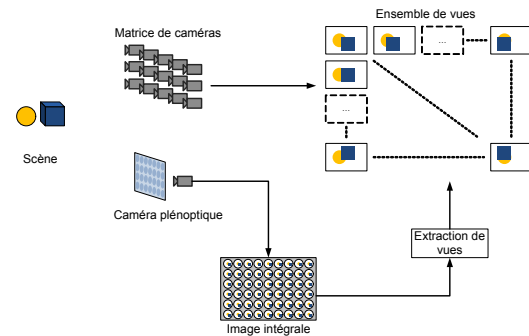


Figure 1: Deux méthodes d'acquisition d'un light-field

la Figure 1. Plusieurs méthodes sont proposées dans l'état de l'art pour extraire des vues à partir d'une image intégrale [GL10] [Lin13].

L'imagerie intégrale et l'acquisition SMV peuvent donner toutes les deux une représentation light-field car elles font un échantillonnage du light-field d'une scène en captant des images de cette scène selon plusieurs angles de vues. Cependant, cet échantillonnage est fait de deux façons distinctes, impliquant les compromis suivants. Un ensemble de caméras permet d'obtenir une *baseline* plus large (avec par exemple une distance de plusieurs mètres entre les deux extrémités du système de caméras) qu'une caméra holoscopique dont la *baseline* est limitée par les dimensions du panneau lenticulaire. Avec une caméra plénoptique, la résolution des vues extraites est limitée car un seul capteur est partagé par toutes les vues captées, alors qu'avec un ensemble de caméra, la résolution totale de chaque caméra est disponible pour chaque vue. De plus, les caméras plénoptiques permettent un échantillonnage plus dense du light-field, car la distance entre les caméras d'un ensemble de caméras est limitée par leurs tailles.

3. État de l'art

3.1. Extensions multi-vues des encodeurs vidéo standards

SMV définit du contenu vidéo 3D multi-vues avec des dizaines ou des centaines de vues, avec une parallaxe soit uniquement horizontale, soit dans les deux dimensions. Le nombre massif de vues augmente la quantité de données à traiter par rapport aux technologies vidéo 3D actuelles. La corrélation inter-vues est également augmentée. Les encodeurs multi-vues standards actuels sont conçus et implémentés pour un contenu avec parallaxe horizontale uniquement et avec un nombre limité de vues. MVC et MV-HEVC sont respectivement les extensions multi-vues des standards H.264/AVC et HEVC [Ohm13]. Ces extensions introduisent de la syntaxe haut-niveau permettant la prédiction inter-vues. L'extension 3D-HEVC fournit des outils de codage liés aux cartes de profondeur, des outils de prédictions inter-composantes (c'est à dire entre texture et carte de profondeur) et de nouveaux outils de codage pour les vues aux niveaux des *Coding Unit* (dans HEVC les CUs remplacent les *macroblocks* de H.264/AVC).

Dans la version du logiciel de référence utilisée pour nos

expériences (HTM7.0), la définition suivante s'applique. *Neighboring Block Disparity Vector* (NBDV) [ZCM12] et *Inter-View Motion Prediction* (IVMP) [TKCY12] sont des outils spécifiques à l'extension 3D-HEVC, conçus pour l'encodage multi-vues horizontal classique. Pour le CU courant, NBDV cherche un vecteur de disparité (DV) dans des CUs voisins (temporaux et spatiaux) déjà codés. Le DV dérivé par NBDV est utilisé par IVMP pour créer l'*Inter-view Predicted Motion Candidate* (IPMC). IPMC correspond aux paramètres de mouvement (vecteurs de mouvement et images de référence temporelle) du CU pointé par le DV dans la vue de référence. IPMC est inséré à la première place dans la liste de candidats du mode *Merge* [HOB*12]. Finalement, le DV lui-même est également inséré dans la liste du *Merge* en tant que candidat *Disparity Motion Vector* (DMV).

3.2. Améliorations pour les configurations avec parallaxe en deux dimensions

La première approche considérée pour encoder du contenu SMV avec parallaxe en deux dimensions est l'utilisation d'un encodeur multi-vue standard, avec une adaptation au niveau de la structure des références inter-vues. Dans [SGM11], les vues sont d'abord scannées en spirale, comme illustré dans la Fig. 2 (a), puis réalignées horizontalement. L'arrangement horizontal est ensuite encodé par un encodeur MVC en utilisant une structure de prédiction IBP (b). La Figure 2 (c) montre le schéma résultant de la représentation sur deux dimensions de cette structure IBP. Le principal inconvénient de cette approche est l'introduction de prédictions incohérentes et inefficaces.

Dans [MSMW07], il est proposé d'appliquer des structures horizontales en IPP ou IBP (Fig. 5(e) and (f)) à chaque ligne de la matrices de vues, et d'ajouter une prédiction inter-vue verticale uniquement sur la première colonne de vues ou sur la centrale, comme illustré dans les Fig. 3 (a),(b) and (c). Le nombre de prédictions inter-vues verticales utilisées est limité dans de telles structures.

Dans [CSK08], [CJSK09] et [CJSK10], une autre structure est proposée, illustrée dans la Fig. 3 (d). Chaque ligne de vues utilise une structure horizontale IBP et des prédictions inter-vues verticales sont ajoutées, donnant des vues de types : B1 qui ont uniquement deux références horizontales ou deux références verticales, B2 avec une référence horizontale et deux références verticales, et B3 avec deux ré-

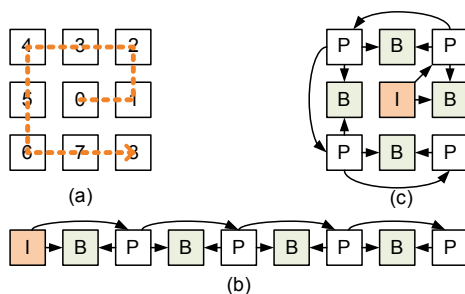


Figure 2: Méthode de l'état de l'art [SGM11] pour 9 vues (a) scan en spirale, (b) structure de prédiction IBP, (c) schéma IBP équivalent en 2 dimensions

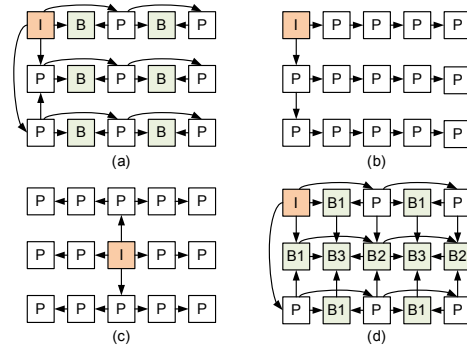


Figure 3: Structures de l'état de l'art (a),(b) and (c) proposées dans [MSMW07], et (d) proposée dans [CSK08, CJSK09, CJSK10].

férences dans chaque direction. Les principales limites sont le nombre réduit de vues qui utilise une combinaison de références horizontales et verticales (moins de la moitié des vues sont de types B2 ou B3) et la distance entre la vue courante/à coder et sa vue de référence.

Une seconde approche au niveau CU est considérée dans [ADCL*12] et dans [CSK08], [CJSK09] et [CJSK10]. Des méthodes similaires sont proposées, basées sur la prédiction d'un DV pour la vue courante par interpolation des DVs des vues voisines.

4. Proposition d'une nouvelle structure de prédiction inter-vues

4.1. Schémas de référence et schéma proposé

On propose une structure pour les images de référence inter-vues, nommée *Central2D* et illustrée dans la Fig. 4 (b), qui permet d'exploiter efficacement un alignement de vues en deux dimensions. Pour une configuration avec $N \times M$ vues, le schéma *Central2D* est construit comme suit. La vue centrale est d'abord codée sans référence inter-vues. Les $N - 1$ (respectivement $M - 1$) vues qui sont sur le même axe horizontal (resp. vertical) que la vue centrale sont ensuite codées avec une référence inter-vue, étant la vue la plus proche dans la direction du centre. Toutes les autres vues sont codées en utilisant une référence inter-vues horizontale et une verticale, étant les vues les plus proches dans la direction du centre. Le schéma permet donc d'utiliser une combinaison de références horizontale et verticale pour un grand nombre de vues (seulement $M + N - 1$ vues n'utilisent pas une référence dans chaque dimension). De plus, cette méthode minimise la distance entre la vue courante/à coder et ses images

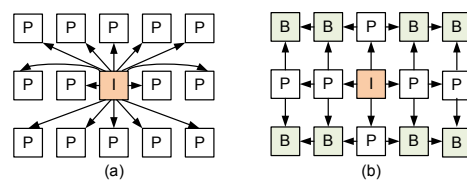


Figure 4: (a) ancrage basique, (b) schéma proposé *Central2D*

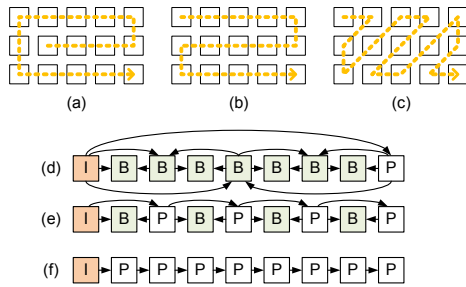


Figure 5: Ordre de scan : (a) spirale, (b) perpendiculaire, (c) diagonale et structures horizontales d'images de référence inter-vue : (d) hiérarchique, (e) IBP, (f) IPP

de référence inter-vues, et n'utilise pas de référence diagonale.

Dans la section suivante, le schéma proposé est comparé à une structure de référence basique (voir Fig. 4 (a)) avec seulement la vue centrale comme image de référence inter-vues pour toutes les autres vues. Cette comparaison permet de déterminer le bénéfice d'une prédiction inter-vues dans les deux directions et avec une distance réduite entre la vue à coder et sa vue de référence. Les structures tirées de l'état de l'art sont aussi testées dans nos expérimentations : [MSMW07] et [CSK08] correspondent aux schémas illustrés dans la Fig. 3 (c) and (d). [SGM11] correspond au scan en spirale avec une structure IBP (Fig. 2). Dans un but de comparaison, on propose également d'étendre la méthode [SGM11] en faisant varier l'ordre de scan et la structure comme c'est illustré dans la Fig. 5.

4.2. Résultats expérimentaux

Dans cette section, on teste les schémas de l'état de l'art et le schéma proposé avec MV-HEVC. La structure de prédiction temporelle reste telle que décrite dans les *Common Test Conditions* (CTC) [RMV12]. Les expérimentations sont effectuées avec le logiciel de référence de MV-HEVC dans sa version 7.0 (HTM7.0, avec la macro QC_MVHEVC activée). Deux séquences sont testées : *CoastalGuard* (50 frames, *computer generated*, de résolution 768×384) et *Akko&Kayo* (290 frames, filmée, de résolution 640×480). Des configurations de 3×3 vues et de 11×5 vues sont testées. Les résultats sont mesurés en utilisant la métrique dite Bjøntegaard Delta rate (BD-rate) [Bjø01], sur les QPs (*Quantization Parameter*) 22-27-32-37. Cette métrique mesure la variation de débit à qualité égale, par rapport à une technique de référence. Dans nos expériences, la référence est le schéma d'ancrage basique (Fig. 4 (a)). Les valeurs négatives représentent un gain, c'est-à-dire, une réduction de débit par rapport à la référence.

Le Tableau 1 montre que pour les deux séquences, avec la configuration 3×3 vues, le schéma *Central2D*, la méthode [MSMW07] et la structure IPP associée aux scans perpendiculaire et spirale sont plus efficaces que les autres méthodes. Ces schémas n'utilisent pas de référence inter-vues en diagonale et minimisent la distance entre la vue à coder et la référence inter-vues. Le gain supplémentaire pour *Central2D* est dû à l'utilisation de références inter-vues dans les deux di-

Coast 3×3			
	spirale	perpendiculaire	diagonale
IPP	-1.2%	-2.2%	5.1%
IBP	9.1%	7.1%	11.4%
Hiérarchique	3.0%	4.4%	8.4%
Méthode [CSK08]	2.1%		
Méthode [MSMW07]	-6.8%		
CENTRAL2D	-7.1%		
Akko 3×3			
	spirale	perpendiculaire	diagonale
IPP	-4.9%	-5.5%	8.8%
IBP	2.7%	-4.0%	-1.9%
Hiérarchique	1.9%	2.4%	4.0%
Méthode [CSK08]	7.8%		
Méthode [MSMW07]	-7.7%		
CENTRAL2D	-8.2%		

Table 1: Variations du BD-rate pour les structures de l'état de l'art et proposée comparées à l'ancrage basique - avec 3×3 vues

Coast 11×5			
	spirale	perpendiculaire	diagonale
IPP	-20.5%	-19.6%	16.1%
IBP	-15.9%	-14.9%	-13.9%
Hiérarchique	-8.4%	-9.3%	-13.0%
Méthode [CSK08]	-19.5%		
Méthode [MSMW07]	-24.4%		
CENTRAL2D	-29.1%		
Akko 11×5			
	spirale	perpendiculaire	diagonale
IPP	-22.9%	-24.8%	-6.5%
IBP	-20.0%	-23.4%	-2.4%
Hiérarchique	-14.9%	-20.2%	-3.7%
Méthode [CSK08]	-24.2%		
Méthode [MSMW07]	-25.9%		
CENTRAL2D	-27.6%		

Table 2: Variations du BD-rate pour les structures de l'état de l'art et proposée comparées à l'ancrage basique - avec 11×5 vues

rections (horizontale et verticale). Le Tableau 2 montre que le schéma *Central2D* reste la structure la plus cohérente et efficace avec un plus grand nombre de vues.

Le gain final en BD-rate apporté par la structure *Central2D* sur l'ancrage basique monte jusque 8.2% et 29.1% dans les configurations de 3×3 et 11×5 vues respectivement.

5. Adaptation et amélioration d'outils de codage inter-vues

5.1. Amélioration de la liste de candidats du mode Merge

On propose dans cette section une modification normative des outils de codage NBDV et IVMP. NBDV et IVMP sont des outils de codage spécifiques implémentés de manière à fonctionner dans les conditions de tests standards (*Common Test Conditions* - CTC [RMV12]), c'est à dire avec une seule image de référence inter-vue horizontale, étant la vue centrale (*baseview*, avec l'indice 0). On adapte ici ces outils

en permettant l'utilisation de plusieurs images de références inter-vues, avec des indices différents de 0, et pouvant être horizontales ou verticales.

En plus de cette adaptation, l'amélioration suivante est proposée. Lorsqu'une vue B utilisant une référence inter-vue horizontale et une verticale est encodée, la version modifiée de NBDV cherche deux DVs (un pour chaque image de référence inter-vues). La recherche du second DV seule n'apporte aucun gain en elle-même, mais va être utilisée pour les candidats IPMC et DMV. Le nouveau second DV est utilisé pour insérer un second IPMC à la seconde place de la liste des candidats du mode Merge. Pour le candidat DMV, le couple de DVs est utilisé, permettant de faire une prédiction bidirectionnelle (ou bi-prédiction).

5.2. Dérivation inter-vues du second DV

On propose d'augmenter les chances de trouver un second DV avec NBDV, afin d'augmenter l'efficacité des candidats IPMC et DMV modifiés. Les étapes sont illustrées dans la Fig. 6. Pour la vue courante, NBDV doit d'abord trouver un premier DV horizontal, pointant sur un CU de référence dans la vue de référence. Si cette vue de référence horizontale a une vue de référence verticale, et si le CU de référence est codé par prédiction inter-vues, le DV vertical utilisé pour la prédiction est hérité/dérivé (par simple copie) en tant que second DV pour le CU courant, et est ensuite utilisé pour les candidats IPMC et DMV (de la manière décrite dans la section précédente). On note que cette méthode peut être utilisée pour les vues de type B ayant une référence inter-vues horizontale et une verticale, ce qui rend la structure *Central2D* précédemment proposée la plus adéquate pour ces outils de codage modifiés.

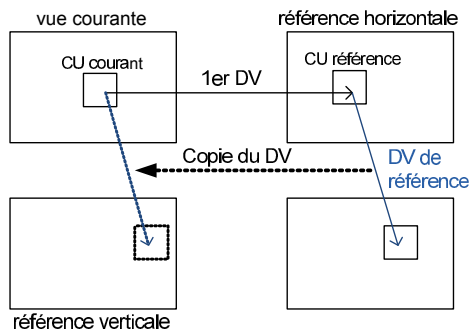


Figure 6: Dérivation inter-vues du second DV

5.3. Résultats expérimentaux

Dans cette section, on teste les modifications proposées des outils de codage NBDV et IVMP. Les expérimentations sont effectuées avec le logiciel de référence de 3D-HEVC dans sa version 7.0 (HTM7.0). Les conditions de test sont les mêmes que dans la Section 4.2 (permettant l'utilisation de structures de codage en deux dimensions). La structure *Central2D* précédemment proposée est utilisée pour tous les tests dans la suite. La référence est 3D-HEVC (HTM7.0 sans modification algorithmique).

Référence : 3D-HEVC (HTM7.0 sans modification algorithmique)				
	3 × 3 vues		11 × 5 vues	
	Coast	Akko	Coast	Akko
Adaptation seule	-1.1%	-2.3%	-2.4%	-3.3%
BiDMV	-1.2%	-2.4%	-2.7%	-3.7%
2 IPMC	-1.1%	-2.3%	-2.8%	-3.5%
Combinaison	-1.3%	-2.5%	-3.1%	-3.9%

Table 3: Variations du BD-rate pour les outils NBDV et IVMP améliorés, utilisant un DV pour chaque image de référence inter-vues

Référence : 3D-HEVC (HTM7.0 sans modification algorithmique)				
	3 × 3 vues		11 × 5 vues	
	Coast	Akko	Coast	Akko
BiDMV + dérivation	-1.9%	-2.9%	-3.4%	-3.9%
2 IPMC + dérivation	-1.3%	-2.4%	-2.8%	-3.5%
Combinaison + dérivation	-2.0%	-2.9%	-3.9%	-4.2%

Table 4: Variations du BD-rate pour les outils NBDV et IVMP améliorés, avec la dérivation inter-vues du second DV

Le Tableau 3 montre que l'adaptation de NBDV et IVMP aux structures bidimensionnelles apporte un gain (BD-rate) jusque 3.3%, ce qui confirme l'impact de l'utilisation des deux dimensions horizontale et verticale au niveau de la structure des images de référence inter-vues. L'insertion d'un second IPMC dans la liste des candidats du mode Merge et la bi-prédiction pour le candidat DMV apporte séparément des gains jusque 2.4% pour la configuration 3 × 3 vues et 3.7% pour la configuration 11 × 5 vues. La combinaison des deux améliorations apporte des gains jusque 2.5% and 3.9% respectivement avec 3 × 3 et avec 11 × 5 vues. Le résultat pour la combinaison des deux outils est légèrement supérieur à la somme des deux pris séparément car la bi-prédiction permet à NBDV de trouver plus souvent un second DV, et donc augmente les chances d'avoir un second IPMC efficace.

Le Tableau 4 montre que la dérivation inter-vues proposées pour le second DV est efficace et améliore la performance du codage de la méthode proposée complète (incluant l'adaptation de NBDV et IVMP aux structures bidimensionnelles, les deux IPMC, la bi-prédiction pour DMV et la dérivation inter-vue du second DV) jusque 2.9% et 4.2% pour la séquence *Akko&Kayo* respectivement avec 3 × 3 et avec 11 × 5 vues.

6. Conclusion

On propose dans cet article une structure pour les images de référence inter-vues adaptée au contenu vidéo 3D *light-field* avec une parallaxe de mouvement en horizontale et en verticale (c'est à dire avec des vues alignées dans les deux directions). La principale caractéristique de cette structure est la distance réduite entre la vue à coder et sa vue de référence, ainsi que l'utilisation de références inter-vues horizontales et verticales. Le schéma proposé surpasse l'ancrage basique avec un gain allant jusque 29.1% (en BD-rate), montrant l'impact de l'utilisation efficace des deux directions horizontale et verticale dans le schéma des images de référence inter-vues. On propose également d'améliorer les outils de

codage NBDV et IVMP (dans 3D-HEVC) afin d'exploiter les directions horizontale et verticale, avec un gain allant jusque 4.2%. Les résultats des méthodes proposées montre qu'exploiter efficacement les deux dimensions horizontale et verticale d'un contenu SMV avec parallaxe dans les deux dimensions au niveau de la structure de codage et au niveau des outils de codage permet d'améliorer de manière significative la performance de compression.

Références

- [AB91] ADELSON E. H., BERGEN J. R. : The plenoptic function and the elements of early vision. *Computational models of visual processing*. Vol. 1, Num. 2 (1991).
- [ADCL*12] AVCI A., DE COCK J., LAMBERT P., BEERNAERT R., DE SMET J., BOGAERT L., MEURET Y., THIENPONT H., DE SMET H. : Efficient disparity vector prediction schemes with modified P frame for 2D camera arrays. *Journal of Visual Communication and Image Representation*. Vol. 23, Num. 2 (February 2012), 287–292.
- [Bjø01] BJØNTEGAARD G. : Calculation of average PSNR differences between RD-curves. In *VCEG Meeting* (Austin, USA, April 2001).
- [CJSK09] CHUNG T.-Y., JUNG I.-L., SONG K., KIM C.-S. : Virtual view interpolation and prediction structure for full parallax multi-view video. In *Advances in Multimedia Information Processing - PCM*, vol. 5879. Springer, 2009, pp. 543–550.
- [CJSK10] CHUNG T.-Y., JUNG I.-L., SONG K., KIM C.-S. : Multi-view video coding with view interpolation prediction for 2D camera arrays. *Journal of Visual Communication and Image Representation*. Vol. 21, Num. 5 (July-August 2010), 474–486.
- [CSK08] CHUNG T., SONG K., KIM C.-S. : Compression of 2-D wide multi-view video sequences using view interpolation. In *15th IEEE International Conference on Image Processing (ICIP)* (San Diego, CA, USA, October 2008), IEEE, pp. 2440–2443.
- [DPC13] DUFAUX F., PESQUET-POPESCU B., CAGNAZZO M. : *Emerging technologies for 3D video : content creation, coding, transmission and rendering*. Wiley Eds, 2013.
- [GL10] GEORGIEV T., LUMSDAINE A. : Focused plenoptic camera and rendering. *Journal of Electronic Imaging*. Vol. 19, Num. 2 (2010), 021106.
- [HOB*12] HELLE P., OUDIN S., BROSS B., MARPE D., BICI M. O., UGUR K., JUNG J., CLARE G., WIEGAND T. : Block merging for quadtree-based partitioning in hevc. *IEEE Transactions on Circuits and Systems for Video Technology*. Vol. 22, Num. 12 (December 2012), 1720–1731.
- [Lin13] LINO J. F. O. : 2D image rendering for 3D holo-scopic content using disparity-assisted patch blending. *Thesis to obtain the Master of Science Degree* (October 2013).
- [Lip08] LIPPMANN G. : Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*. Vol. 7, Num. 1 (1908), 821–825.
- [MSMW07] MERKLE P., SMOLIC A., MULLER K., WIEGAND T. : Efficient prediction structures for multiview video coding. *IEEE Transactions on Circuits and Systems for Video Technology*. Vol. 17, Num. 11 (November 2007), 1461–1473.
- [Ohm13] OHM J.-R. : Overview of 3D video coding standardization. In *International Conference on 3D Systems and Applications* (Osaka, Japan, June 2013).
- [RMV12] RUSANOVSKY D., MULLER K., VETRO A. : Common test conditions of 3DV core experiments. In *International Organisation For Standardisation* (October 2012), ITU-T SG 16 WP 3 & ISO/IEC JTC1/SC29/WG11 JCT3V-B11000.
- [SGM11] SHI S., GIOIA P., MADEC G. : Efficient compression method for integral images using multi-view video coding. In *18th IEEE International Conference on Image Processing (ICIP)* (Brussels, Belgium, September 2011), IEEE, pp. 137–140.
- [TKCY12] TECH G., K.WEGNER, CHEN Y., YEA S. : 3D-HEVC test model 2. In *International Organisation For Standardisation* (October 2012), ITU-T SG 16 WP 3 & ISO/IEC JTC1/SC29/WG11 JCT3V-B1005.
- [TSO*13a] TEHRANI M. P., SENOH T., OKUI M., YAMAMOTO K., INOUE N., FUJII T. : [m31095][FTV AHG] Use cases and application scenarios for super multiview video and free-navigation. In *International Organisation For Standardisation* (October 2013), ISO/IEC JTC1/SC29/WG11.
- [TSO*13b] TEHRANI M. P., SENOH T., OKUI M., YAMAMOTO K., INOUE N., FUJII T. : [m31103][FTV AHG] Introduction of super multiview video systems for requirement discussion. In *International Organisation For Standardisation* (October 2013), ISO/IEC JTC1/SC29/WG11.
- [TSO*13c] TEHRANI M. P., SENOH T., OKUI M., YAMAMOTO K., INOUE N., FUJII T. : [m31261][FTV AHG] Multiple aspects. In *International Organisation For Standardisation* (October 2013), ISO/IEC JTC1/SC29/WG11.
- [ZCM12] ZHANG L., CHEN Y., M.KARCZEWICZ : 3D-CE5.h related : Disparity vector derivation for multiview video and 3DV. In *International Organisation For Standardisation* (July 2012), ISO/IEC JTC1/SC29/WG11 MPEG2012/m24937.