

Compensation *spatio*-temporelle Globale pour le Codage Vidéo par Ondelettes 3D

B. Le Guen^{1,2}

S. Pateux¹

J. Weiss²

¹ France Télécom R&D, 4, rue du Clos Courtel, 35512 Cesson-Sévigné

{benjamin.leguen, stephane.pateux}@orange-ft.com

² Supélec-SCEE/IETR-AC, avenue de la Boulaie, 35511 Cesson-Sévigné

jacques.weiss@supelec.fr

Résumé

Ce papier propose un schéma de codage vidéo par analyse-synthèse prenant en compte les corrélations à la fois temporelles et spatiales. Le principe général est de déformer un groupe d'images (GOF) pour l'adapter aux directions de filtrage d'une ondelette séparable 3D. Ici, géométrie et mouvement sont représentés par un même modèle (maillage actif) et estimés par des algorithmes similaires. Le schéma est conçu pour limiter le coût additionnel de la géométrie puisque une seule géométrie doit être transmise pour l'ensemble d'un GOF. Les résultats de compression obtenus montrent des gains visuels au niveau des contours des images décodées par rapport à une analyse-synthèse seulement temporelle.

Mots clefs

Codage, Vidéo, Mouvement, Géométrie, Maillage.

1 Introduction

Pour diminuer le coût de codage d'un signal, il faut prendre en compte les corrélations qu'il contient. Une vidéo est un signal 3D (2D+t) possédant des corrélations temporelles mais également spatiales. Les corrélations temporelles sont essentiellement liées au *mouvement* des objets au cours du temps. Les corrélations spatiales sont liées notamment aux contours des objets ou aux motifs des textures, nous parlerons plus généralement de *géométrie*.

Les codeurs vidéo de l'état de l'art [1], sont des codeurs prédictifs reposant sur une boucle de compensation en mouvement et un encodage 2D des résidus de prédiction par Transformée en Cosinus Discret (DCT). Pour décorrélérer le mouvement sur plus de 2 images, un encodage par transformée (Motion Compensated Temporal Filtering MCTF [2]) peut aussi être appliqué le long des trajectoires de mouvement. Pour pouvoir reconstruire les coefficients de texture, le décodeur doit alors connaître parfaitement les trajectoires utilisées à l'encodage. Ainsi, une légère altération de l'information de mouvement due à des pertes réseau ou au tronçage du flux peut conduire à de fortes chutes de qualité. On notera d'autre part que la

DCT ou les modèles de mouvement par blocs utilisés dans les standards ne sont pas des outils multi-résolution, propriété qui facilite la création d'un flux scalable. De plus, la DCT ne permet pas de prendre en compte les orientations spatiales corrélées des images. Des adaptations de DCT orientée [3] ont été réalisées pour l'encodage des images intra mais n'apportent un léger gain qu'à très haut débit.

Parallèlement, des techniques par analyse-synthèse [4, 5] ont été proposées. L'idée est d'aligner toutes les images d'un groupe d'images (GOF) sur un instant de projection avant de réaliser une décomposition multi-résolution en ondelettes 3D. En d'autres termes, par opposition aux techniques précédentes, les directions de filtrage sont ici fixes et c'est la vidéo qui est déformée pour s'y adapter. Ainsi, le décodage des coefficients de texture ne nécessite aucune information de filtrage. Une altération dans le flux de mouvement transmis va certes nuire à la déformation inverse des images lors de la reconstruction finale, mais la qualité de la vidéo ne chute pas significativement. Un tel codeur est présenté dans [5]. Ce codeur utilise un maillage actif comme modèle de mouvement, ce qui permet de définir des correspondances bijectives et *globales* entre les images déformées et les images originales. Les auteurs montrent notamment qu'un tel modèle autorise un décodage avec perte de l'information de mouvement.

Nos travaux se situent dans la continuité du schéma précédent. D'une manière générale, nous observons que les codeurs par ondelettes 3D ne prennent pas en compte les corrélations spatiales qui ne sont ni horizontales ni verticales. Ceci nuit à la parcimonie de la représentation et se traduit par des rebonds d'ondelettes au niveau des contours. Notre objectif est de proposer une solution à ce problème en utilisant une ondelette adaptative. Un grand nombre d'ondelettes adaptatives existent. La plupart (e.g. [6, 7]) adoptent un modèle par blocs, où, à chaque bloc est associée une orientation ou une courbe de filtrage. Par analogie avec la classification utilisée pour le mouvement, certaines méthodes [7] nécessitent de connaître parfaitement les orientations au décodage, d'autres [6] peuvent être assimilées à des techniques d'analyse-synthèse car elles pro-

posent de déformer chaque bloc avant d’appliquer l’ondelette 2D classique. Dans [8], nous avons proposé un schéma de codage d’images fixes par analyse-synthèse. Ce schéma est basé sur un maillage actif permettant une déformation globale de l’image. La déformation est calculée de façon à minimiser le coût de codage de l’image déformée dans une base d’ondelettes classiques. Le travail présenté ci-dessous montre comment combiner les schémas proposés dans [5] et [8] de manière à déformer un groupe d’images à la fois selon ses corrélations temporelles et spatiales. La problématique principale est de limiter le surcoût de codage dû à la transmission de la géométrie. Nous montrons que le schéma proposé permet de ne transmettre qu’une seule géométrie par GOF.

L’article est organisé comme suit. Dans la seconde section, nous établissons un parallèle entre estimation-compensation en mouvement et en géométrie. Dans la troisième partie, nous décrivons le schéma proposé. La dernière partie présente des résultats de codage sur des séquences usuelles.

2 Similarités Géométrie-Mouvement

De nombreuses similarités existent entre mouvement et géométrie : chaque caractéristique définit des trajectoires régulières à exploiter, temporelles ou spatiales. Dans nos travaux, géométrie et mouvement sont représentés par le même modèle : le maillage déformable. Les paramètres de ce modèle sont calculés par des procédures équivalentes.

2.1 Représentation par maillage déformable

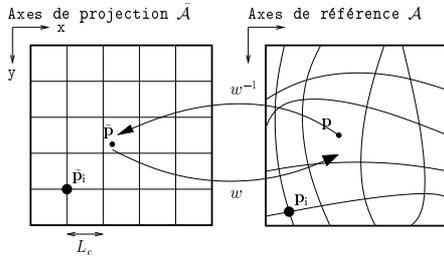


Figure 1 – Correspondance bijective par maillage déformable.

Considérons 2 systèmes d’axes placés arbitrairement dans le plan de la caméra à 2 instants, distincts ou non : un système d’axes de projection $\tilde{\mathcal{A}}$ et un système d’axes de référence \mathcal{A} . Nous définissons alors un mapping de coordonnées w faisant correspondre une position $\tilde{\mathbf{p}}$ dans $\tilde{\mathcal{A}}$ à une position \mathbf{p} dans \mathcal{A} . Dans ce papier, tout mapping sera représenté par un maillage quadrangulaire régulier. Conformément à la Fig. 1, nous plaçons une grille uniforme fixe dans $\tilde{\mathcal{A}}$. Le mapping est alors défini totalement par les paramètres suivants : d’une part la longueur L_e d’une arête de la grille uniforme dans $\tilde{\mathcal{A}}$, et d’autre part les correspondants des noeuds de cette grille dans \mathcal{A} . Le correspondant d’une position $\tilde{\mathbf{p}}$ quelconque dans $\tilde{\mathcal{A}}$ est obtenu en interpolant les correspondants des noeuds qui l’en-

tourent. Toutes les expériences présentées dans la suite ont été réalisées à l’aide d’une interpolation bilinéaire. Après avoir estimé les paramètres d’un tel mapping, toute image f peut alors être projetée de \mathcal{A} vers $\tilde{\mathcal{A}}$ pour donner une image déformée \tilde{f} :

$$\tilde{f}(\tilde{\mathbf{p}}) = f(w(\tilde{\mathbf{p}})) \quad (1)$$

Un maillage régulier permet en outre une déformation *globale* et continue de l’image, de telle sorte que toute zone peut être reconstruite.

2.2 Estimation de mouvement

Dans le cas du mouvement, les 2 systèmes d’axes $\tilde{\mathcal{A}}$ et \mathcal{A} sont placés dans le plan de la caméra à 2 instants distincts t_0 et t_1 . Soit f_{t_0} et f_{t_1} les images de la séquence vidéo à ces 2 instants. Le mapping de coordonnées recherché est celui qui permet l’alignement optimal, au sens de l’erreur aux moindres carrés, entre $f_{t_0}(\tilde{\mathbf{p}})$ et $f_{t_1}(w(\tilde{\mathbf{p}}))$. Nous notons ce mapping $w_{t_1}^m$:

$$w_{t_1}^m = \arg \min_w \sum_{\tilde{\mathbf{p}}} [f_{t_0}(\tilde{\mathbf{p}}) - f_{t_1}(w(\tilde{\mathbf{p}}))]^2 \quad (2)$$

Dans notre implémentation, le problème exprimé dans l’Eq. 2 est résolu par une technique de descente en gradient. Cette technique est itérative. Les positions des noeuds dans \mathcal{A} sont initialisées sur la grille uniforme de $\tilde{\mathcal{A}}$ (i.e. $w = Id$). A chaque itération, ses positions sont raffinées en annulant la dérivée de l’erreur et en résolvant un système linéaire éparé. Cette technique peut être améliorée de différentes façons mais ce n’est pas le sujet de cet article.

A l’issue de l’estimation, une compensation en mouvement globale permet de calculer l’image déformée $f_{t_1}(w_{t_1}^m(\tilde{\mathbf{p}}))$. Nous la notons $\tilde{f}_{t_1}^m$.

2.3 Estimation de géométrie

Dans le cas de la géométrie, les 2 systèmes d’axes $\tilde{\mathcal{A}}$ et \mathcal{A} sont placés dans le plan de la caméra au même instant. Considérons une image fixe f à un instant quelconque. Cette image a un certain coût de codage dans une base d’ondelettes séparable classique. Il est connu que cette base n’est pas optimale pour représenter notamment les contours courbes, qui ne sont ni horizontaux, ni verticaux. De plus, dans une direction donnée, une discontinuité brusque (bruitée) est représentée par des coefficients sur une large gamme de fréquences. Admettons donc qu’il existe une version compensée en géométrie \tilde{f}^g de f dont le coût de codage dans cette base soit, lui, optimal. Pour estimer un mapping w^g permettant de tendre vers cet optimal, nous avons proposé dans [8] une technique basée sur l’expression et la minimisation du coût de codage de \tilde{f}^g en fonction de w^g . Le résultat principal de ce papier est que le mapping recherché est celui qui permet l’alignement optimal, au sens de l’erreur aux moindres carrés, entre f et une image dite image cible f_{cible}^g :

$$w^g = \arg \min_w \sum_{\tilde{\mathbf{p}}} [\tilde{f}_{cible}^g(\tilde{\mathbf{p}}) - f(w(\tilde{\mathbf{p}}))]^2, \quad (3)$$

où \tilde{f}_{cible}^g est une combinaison linéaire des approximations basses fréquences de \tilde{f}^g (voir [8] pour plus de détails). En comparant les équations (2) et (3), nous observons que **notre formulation du problème d'estimation géométrique est très proche du problème d'estimation de mouvement**. La différence est que l'image sur laquelle on souhaite s'aligner dans le cas géométrique est inconnue. Pour cette raison, nous avons proposé dans [8] une procédure itérative permettant d'optimiser le couple (w^g, \tilde{f}^g) de façon conjointe. L'unique différence par rapport à une estimation de mouvement est que, à chaque itération de la descente en gradient, l'image compensée \tilde{f}^g est raffinée sachant la nouvelle observation de w^g .

2.4 Compensation Spatio-Temporelle

Les algorithmes présentés plus haut ont pour but d'aligner une image soit sur une image à un instant différent, soit sur une image cible à moindre coût de codage spatial. Ces 2 types de compensation peuvent être combinés en un seul opérateur. Par exemple, une estimation de géométrie peut être réalisée sur l'image compensée en mouvement $\tilde{f}_{t_1}^m$, donnant un mapping $w_{t_1}^g$. Le mapping combiné $w_{t_1}^g \circ w_{t_1}^m$, que nous notons $w_{t_1}^{mg}$, permet de calculer en une seule opération une version $\tilde{f}_{t_1}^{mg}$ de f_{t_1} compensée en mouvement et en géométrie :

$$\tilde{f}_{t_1}^{mg}(\mathbf{p}) = f_{t_1}(w_{t_1}^{mg}(\mathbf{p})) \quad \forall \mathbf{p} \quad (4)$$

Les compensations en mouvement et en géométrie peuvent aussi être combinées dans l'ordre opposé. Cependant, dans nos travaux, nous avons fait le choix d'une analyse t+2D d'un GOF qui conduit à combiner les mappings dans cet ordre.

3 Analyse-Synthèse t+2D

3.1 Analyse

Le schéma de codage proposé opère sur des groupes de N images (GOF). L'objectif premier est d'obtenir à partir de chaque GOF un groupe d'images déformées adapté à une décomposition dans une base d'ondelettes 3D séparable. Pour y parvenir, nous proposons d'aligner chaque image du GOF à la fois temporellement et spatialement au cours d'une étape d'*analyse*. Ce groupe d'images compensé en mouvement et en géométrie sera appelé groupe de textures (GOT). Le terme de texture est ici utilisé en référence au monde de la 3D où un objet est souvent représenté par un maillage 3D (sa géométrie dans l'espace) et une texture venant se plaquer sur le maillage. Dans notre cas, chaque image originale est de même représentée à la fin de l'analyse par une texture et un mapping de coordonnées.

L'analyse se déroule en 3 temps. Comme il serait trop coûteux, à la fois en débit et en complexité, de calculer une géométrie pour chaque image du GOF, nous avons élaboré l'analyse t+2D suivante. Elle permet de ne considérer qu'une seule géométrie pour l'ensemble du GOF.

Alignement temporel :

Nous plaçons les axes de projection $\tilde{\mathcal{A}}$ dans le plan de la caméra à l'instant t_0 . Pour chaque instant $t \in \{1, \dots, N\}$, une estimation de mouvement entre t et t_0 donne un mapping w_t^m et une prédiction \tilde{f}_t^m de f_{t_0} . L'ensemble des prédictions donne un volume compensé en mouvement.

Alignement spatial :

En supposant que l'alignement temporel a été efficace, toutes les images \tilde{f}_t^m contiennent approximativement la même structure géométrique. De ce fait, nous décidons d'estimer la géométrie sur une basse fréquence temporelle f_{LF} du volume compensé en mouvement :

$$f_{LF}(\mathbf{p}) = \frac{1}{N} \sum_{t=1..N} \tilde{f}_t^m(\mathbf{p}) \quad (5)$$

L'estimation de géométrie sur f_{LF} donne la géométrie w_{LF}^g .

Génération du GOT :

Le GOT compensé en mouvement et géométrie $\{\tilde{f}_0^{mg}, \dots, \tilde{f}_N^{mg}\}$ est obtenu en combinant w_{LF}^g avec chaque mapping de mouvement w_t^m et en appliquant l'opérateur de compensation sur chaque image d'origine :

$$\tilde{f}_t^{mg}(\mathbf{p}) = f_t(w_{LF}^g \circ w_t^m(\mathbf{p})) \quad \forall t, \forall \mathbf{p} \quad (6)$$

Les compensations en mouvement et géométrie sont donc réalisées en **une seule passe**.

3.2 Encodage

A l'issue de l'étape d'analyse, le GOT compensé est encodé par un simple codeur ondelettes 3D séparable¹. La qualité multi-résolution de l'ondelette peut également permettre de générer un flux scalable dans le temps et l'espace. Au cours de l'encodage, nous observons qu'aucune compensation n'est requise pour mapper plusieurs pixels corrélés, puisque les trajectoires corrélées ont été alignées. Les coefficients de texture du GOT peuvent de ce fait être décodés sans information de mouvement ou de géométrie. Les maillages représentant le mapping de géométrie w_{LF}^g et l'ensemble des mappings de mouvement $\{w_t^m\}_{t \neq t_0}$ sont encodés indépendamment.

Les flux générés sont envoyés au récepteur.

3.3 Synthèse

Le bloc de synthèse prend comme entrée le GOT décodé, noté $\{\hat{f}_0^{mg}, \dots, \hat{f}_N^{mg}\}$, la géométrie décodée \hat{w}_{LF}^g et l'ensemble des mouvements décodés $\{\hat{w}_t^m\}_{t \neq t_0}$. Puisque tous les mappings sont bijectifs et donc réversibles, les images originales peuvent être reconstruites en utilisant la relation suivante :

$$\hat{f}_t(\mathbf{p}) = \hat{f}_t((\hat{w}_{LF}^g \circ \hat{w}_t^m)^{-1}(\mathbf{p})) \quad \forall \mathbf{p} \quad (7)$$

¹NB : La compensation géométrique peut être optimisée pour d'autres noyaux que l'ondelette 2D séparable, par exemple la DCT. Ainsi, d'autres codeurs volumiques (3D-DCT) pourrait être utilisés pour encoder le GOT.

Les compensations inverses en mouvement et géométrie sont donc réalisées en **une seule passe**.

Notons que le décodage sans perte des maillages n'est pas requis pour pouvoir reconstruire des images de bonne qualité. En outre, une perte sur le maillage permet d'affecter plus de débits au décodage du GOT, ce qui peut permettre au final de reconstruire des images de meilleure qualité subjective.

4 Résultats expérimentaux

4.1 Résultats d'analyse

Dans ce paragraphe, nous illustrons l'étape d'alignement spatial proposée dans 3.1 car c'est l'apport majeur de notre travail par rapport à [5]. Le résultat présenté Fig. 2 a été obtenu en lançant le codeur sur la vidéo Foreman CIF 30 Hz avec des GOF de taille 8. L'instant de projection est situé en milieu de GOF ($t_0 = 3$). Le mouvement est modélisé par un maillage actif avec $L_e = 16$ et 15 itérations de la descente en gradient sont réalisées pour calculer les positions de ces noeuds à chaque instant. Chaque image du GOF est ensuite projetée sur t_0 et la basse fréquence temporelle f_{LF} est calculée. L'estimation de géométrie est alors réalisée sur f_{LF} . A droite sur la Fig. 2 est affichée la f_{LF} du GOF n°1. La géométrie calculée est superposée à l'image. Elle est modélisée avec des mailles de taille $L_e = 8$ et estimée avec 15 itérations de l'optimisation conjointe proposée.

Nous observons que les mailles se sont resserrées autour des contours saillants de l'image au cours des itérations. Si nous compensons f_{LF} globalement en utilisant le mapping défini par le maillage (comme représenté sur la Fig. 1), nous obtenons l'image compensée affichée sur la gauche. Dans cette image compensée, on observe que certains contours ont été étirés et on voit un effet escalier apparaître dans certains endroits : l'estimation de géométrie a tendance à lisser et ré-orienter les contours sur l'horizontale ou la verticale. Ceci explique pourquoi l'image compensée est représentée plus efficacement par l'ondelette séparable que l'image d'origine. Notons qu'en réalité les compensations en mouvement et géométrie ne sont pas réalisées sur f_{LF} mais sur chacune des images d'origine du GOF. De plus, ces images compensées ne sont pas destinées à être affichées, ce sont juste des images intermédiaires dans le process.

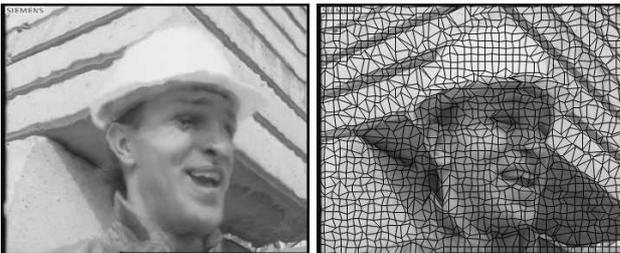


Figure 2 – [Droite] Maillage superposé à f_{LF} , [Gauche] Compensation géométrique globale de f_{LF} . GOF n°1.

4.2 Résultats de compression

Ce paragraphe compare 3 types de codeur. Premièrement, le schéma par Analyse-Synthèse t+2D proposé (**A-S t+2D**) avec les paramètres d'analyse donnés dans le paragraphe précédent. Les positions des noeuds des maillages sont quantifiées avec une précision au 1/2 pixel pour le mouvement et au pixel pour la géométrie. Les symboles sont encodés à l'aide d'un codeur arithmétique. Le GOT est décorrélé temporellement à l'aide de l'ondelette 5/3. Les images de résidus temporels sont enfin envoyées à JPEG2000 VM 8.0 pour décomposition spatiale et encodage. Le second codeur implémente le schéma d'analyse-synthèse temporelle (**AS t**) proposé dans [5], avec les mêmes paramètres et procédures d'encodage que précédemment. Le troisième codeur est un codeur prédictif IPP simple (**Prédictif**) qui prend la première image de chaque GOF comme image intra. Le mouvement est modélisé avec des blocs de taille 16×16 et estimé avec une recherche exhaustive. Le mouvement est quantifié au 1/2 pixel et encodé de la même façon que pour les codeurs A-S. Pour évaluer justement le potentiel des schéma A-S, nous choisissons d'encoder les résidus de prédictions du codeur prédictif avec JPEG2000.

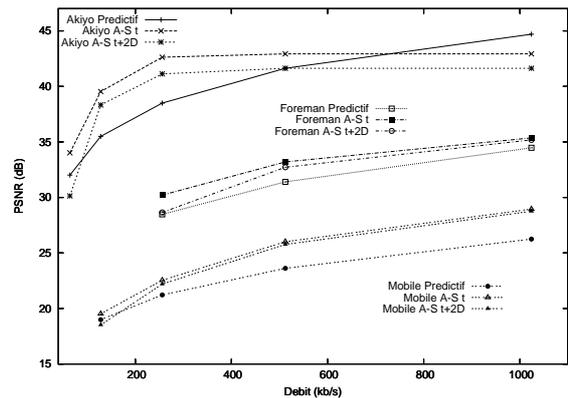


Figure 3 – Courbes de PSNR. Séquence Akiyo, Foreman et Mobile CIF 30 Hz.

La Fig. 3 donne les courbes de PSNR obtenues pour Akiyo, Foreman et Mobile CIF 30 Hz. Les tailles de GOF choisies sont respectivement 16, 8 et 8. Le premier constat est que le schéma A-S t est le plus performant dans la majorité des cas. Ensuite, la courbe du codeur prédictif croise et dépasse celles des schémas A-S à partir d'un certain débit dans le cas d'Akiyo, mais cette tendance doit être une règle générale quelque soit la séquence traitée. L'explication est que les codeurs AS ne permettent pas la reconstruction parfaite du fait des pertes introduites lors des ré-échantillonnages successifs. Enfin, on voit que la qualité objective obtenue pour le schéma AS t+2D est légèrement en-dessous de celle du schéma AS t. D'une manière générale, cette observation est confirmée par la qualité subjective globale des images reconstruites. La raison principale est la part importante du

Tableau 1 – Coût des paramètres de compensation (kb/s).

video / codec	Prédicatif	A-S t	A-S t+2D
Akiyo	16.78	18.13	40.86
Foreman	76.21	94.20	113.31
Mobile	50.46	52.61	78.40

débit occupée par la géométrie. Le tableau 1 donne en kb/s le coût des paramètres de compensation (modèles blocs, ou maillages) dans les 3 cas. Nous voyons que la prise en compte de la géométrie augmente ce coût d'environ 20 kb/s. Ce coût additionnel n'est donc pas suffisamment compensé par la diminution du coût de codage du GOT. Néanmoins, les résultats obtenus restent positifs : d'une part des gains visuels sont observés dans la synthèse des contours, comme le montre la Fig. 4 ; d'autre part, certaines améliorations peuvent être envisagées pour réduire le coût de la géométrie (taille de maille adaptative, codage contextuel des positions...).

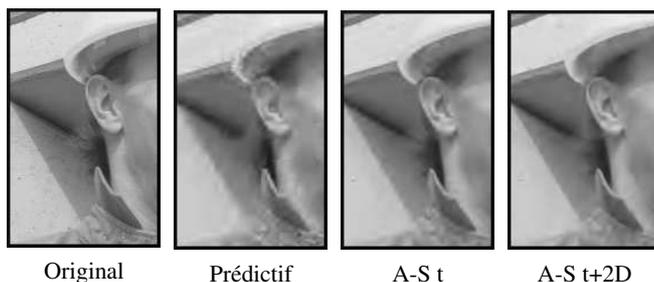


Figure 4 – Résultats visuels sur Foreman à 256kb/s.

Pour finir, nous illustrons sur la Fig. 5 la possibilité de décoder les paramètres de compensation avec perte dans le cas des schémas A-S par rapport au schémas prédicatifs. Nous encodons ainsi la séquence Akiyo à l'aide du codeur prédicatif et du codeur AS t+2D en utilisant un codage du mouvement en plans de bits. Au décodage, nous ne reconstruisons pas le dernier plan de bits de mouvement. La Fig. 5 montre une image reconstruite à très haut débit. A gauche (codeur prédicatif), cette légère perte introduit une distorsion visuelle importante. Si une MCTF était ajoutée au schéma, cette distorsion pourrait provoquer une dérive désastreuse lors de la reconstruction des images adjacentes. A droite (codeur AS t+2D), la perte est peu visible et l'image reste de bonne qualité. Il est ainsi possible d'obtenir des images de meilleure qualité en décodant le mouvement et/ou la géométrie avec perte et en reportant le débit gagné sur le décodage des textures.



Figure 5 – Perte de mouvement. A gauche, codeur prédicatif. A droite, codeur AS t+2D. Image 63, 1024 kb/s.

5 Conclusion

Nous avons présenté un schéma par Analyse-Synthèse prenant en compte à la fois la géométrie et le mouvement d'une séquence vidéo. Dans ce schéma, géométrie et mouvement sont modélisés par un maillage actif et estimés par des algorithmes équivalents. Le schéma a été conçu de manière à limiter le coût additionnel de la géométrie. Même si ce coût reste encore trop important, des gains ont été obtenus au niveau des contours saillants des images. Dans la suite, nous chercherons à réduire le coût de la géométrie et à tester d'autres modèles que le maillage actif.

Références

- [1] Advanced video coding for generic audio-visual services. Draft ITU-T Recommendation and Final Draft International Standard, JVT-G050r1, 2003.
- [2] L. Luo, J.Li, S.Li, Z.Zhuang, et Y-Q.Zhang. Motion-compensated lifting wavelet and its application in video coding. Dans *Proc. IEEE International Conference on Multimedia and Expo*, Tokyo, Japan, Août 2001.
- [3] A. Robert, I. Amonou, et B. Pesquet-Popescu. Amélioration de codeurs DCT par orientations des blocs de la transformée. Dans *Actes de la Conférence CORESA*, Caen, France, Novembre 2006.
- [4] D. Taubman et A. Zakhor. Orientation adaptive sub-band coding of images. *IEEE Transaction on Image Processing*, 3 :421–436, Juillet 1994.
- [5] N. Cammas et S. Pateux. Fine grain scalable video coding using 3D wavelets and active meshes. Dans *Proc. SPIE Visual Communication on Image Processing*, Santa Clara, California, USA, Janvier 2003.
- [6] E. Le Pennec et S. Mallat. Sparse Geometric Image Representations with Bandelets. *IEEE Transaction on Image Processing*, 14(4) :423–438, Avril 2005.
- [7] W. Ding, F. Wu, et S. Li. Lifting-based Wavelet Transform with Directionally Spatial Prediction. Dans *Proc. Picture Coding Symposium*, San Francisco, USA, Décembre 2004.
- [8] B. Le Guen, S. Pateux, et J. Weiss. Non-Geometric Energy Formulation for Adaptive Image Compression. Dans *Proc. IEEE International Conference on Image Processing*, San Antonio, USA, Octobre 2007.