

Etude de maillage rectangulaire déformable pour l'estimation de mouvement dans des séquences vidéo

Vianney Muñoz-Jimenez^{+,*}, Anissa Zergainoh^{+,*}

^{*}L2TI, Institut Galilée, Université Paris 13
99, Avenue Jean Baptiste Clément, 93 430 Villetaneuse, France

⁺LSS/CNRS, Supelec
Plateau de Moulon, 91 192 Gif sur Yvette, France

Email : {munoz, anissa.zergainoh}@lss.supelec.fr

Résumé

L'estimation de mouvement adoptée par les standards vidéo actuels est basée sur l'algorithme d'appariement de blocs. Cet algorithme ne détermine que les mouvements de translation des objets dans la scène. A bas débit, ces mouvements se traduisent par des effets visuels d'artefacts sur les images prédites. Pour mieux décrire les différents types de mouvement d'objets dans la scène, nous étudions l'estimation de mouvement à partir de la déformation des mailles rectangulaires des grilles associées aux images de référence.

Mots clefs

Interpolation, spline, maillage rectangulaire, déformation, estimation de mouvement.

1 Introduction

Une séquence vidéo est caractérisée par un volume de données important qu'elle doit véhiculer à travers un réseau de transmission ou occuper sur un support de stockage. Elle est représentée par un ensemble d'images consécutives fortement corrélées. Celles-ci défilent les unes après les autres, à une cadence temporelle régulière et prédéfinie, pour donner une illusion de mouvement. Afin de diminuer le coût de codage, les redondances temporelles inter-images sont généralement exploitées. En effet, les images de la séquence vidéo ne sont pas codées de manière indépendante. A ce jour, l'estimation et la compensation de mouvement sont des techniques largement développées et utilisées par les standards vidéo dans le but de réduire les redondances temporelles ([1], [2], [3], [4], [5], [6]). Généralement, la prédiction d'une image courante de la séquence vidéo est obtenue à partir d'une ou de plusieurs images préalablement reconstruites. Cette prédiction inclut une opération de compensation de mouvement de l'image courante. Le mouvement apparent 2D dans une séquence vidéo peut avoir plusieurs origines : le mouvement d'objets 3D dans la scène, les mouvements de la caméra ou les variations d'illuminations. Dans les

normes actuelles de vidéo, le champ de mouvement des objets dans la scène est très souvent représenté à partir d'un modèle de translation basé sur une approche d'appariement de blocs de l'image courante avec l'image de référence (ou vis versa). Chaque bloc de l'image courante est obtenu par translation d'un bloc de l'image de référence qui semble se rapprocher au mieux (selon une mesure prédéfinie) du bloc de l'image de référence. Le principal intérêt de cette approche est sa simplicité. Néanmoins les blocs compensés ne forment pas toujours une partition complète de l'image de référence. Il peut apparaître dans l'image prédite des zones qui sont soit découvertes ou recouvertes puisque aucune contrainte n'est imposée aux blocs adjacents dans l'image courante. Afin de remédier aux inconvénients de cette approche, de nouvelles méthodes basées sur la construction d'un maillage déformable ont été développées ([7]).

Des maillages uniformes à base de mailles rectangulaires, triangulaires et hexagonales ont été proposés ([8], [9], [10], [11], [12], [13]). Des méthodes proposent d'utiliser un seul modèle optimal de déformation adapté à toute la séquence vidéo considérée ([14], [15]). Certaines méthodes combinent également l'appariement de bloc avec le maillage uniforme. Tandis que d'autres méthodes emploient des décompositions hiérarchiques (sous-bandes, quad-tree) pour réduire le temps de calcul nécessaire à l'estimation de mouvement ([16], [17], [18], [19], [20], [21], [22], [23]). Certains algorithmes de maillage sont adaptés au contenu de l'image. Bien que performants, certains algorithmes souffrent de complexité de calculs.

Dans cet article nous étudions et comparons les performances de l'estimation de mouvement dans une séquence vidéo à partir de la déformation des mailles rectangulaires des grilles associées aux images. Nous présentons trois types de découpages.

Cet article est organisé comme suit. La section 2 introduit le concept général de la représentation des images animées par maillage rectangulaire. La section 3 concerne l'estimation de mouvement locale obtenue à partir de l'optimisation des déplacements des nœuds des différentes mailles. La section 4 analyse les résultats expérimentaux obtenus.

2 Présentation de l'approche basée maillage

Une séquence vidéo est représentée par un ensemble d'images successives fortement corrélées. Afin d'obtenir de bonnes performances de codage, les standards vidéo actuels classifient les images en trois catégories selon la pertinence de l'information contenue dans les images : (i) les images de type I (codées en Intra), (ii) les images prédites de type P (codées en Inter) et (iii) les images bidirectionnelles de type B (codées en Inter). Dans le cadre de notre premier travail, nous ciblons les applications de type vidéoconférence (cadre simple). Dans cet article, nous utilisons une méthode d'estimation de mouvement basée maillage de type « forward ». Introduisons, tout d'abord, dans le paragraphe ci-dessous la construction de la grille de référence nécessaire pour l'estimation de mouvement.

2.1 Construction de la grille de référence

Considérons la $k-1$ ième image de type I dans la séquence vidéo. Notons I_{k-1} l'intensité de cette image de référence. Une grille de référence, notée G_{k-1} , est associée à l'image I_{k-1} . Cette grille G_{k-1} est construite progressivement à partir d'un ensemble de mailles rectangulaires, notées $\{\mathcal{M}_i^{k-1}, i \in N\}$.

Une maille rectangulaire \mathcal{M}_i^{k-1} est définie à partir de 4 nœuds choisis sur l'image à mailler (voir Figure 3 (a)). Les mailles rectangulaires sont juxtaposées les unes à côté des autres afin de former une partition continue de la grille de référence.

Dans cet article nous comparons les performances de trois types de maillage rectangulaire de la grille de référence G_{k-1} : (i) le maillage uniforme (ii) le maillage semi-régulier et adapté (iii) le maillage hiérarchique et adapté. Les deux derniers types de maillage sont basés sur le gradient de l'image de référence ∇I_{k-1} . Cette stratégie est motivée par le fait que le mouvement des objets dans une scène peut être déduit par uniquement par le suivi du déplacement des contours de ces objets dans l'image suivante. Nous verrons, ci-dessous, que la grille de référence G_{k-1} est construite itérativement. Commençons par présenter le maillage hiérarchique et adapté.

Maillage hiérarchique adapté

Dans le cadre de maillage hiérarchique, nous commençons par un pavage rectangulaire uniforme et grossier de l'image de référence I_{k-1} . Ce découpage s'apparente au découpage quad-tree. La différence réside dans le fait que la grille est initialement découpée en maille rectangulaire de taille identique. Le pavage initial est ensuite affiné en fonction du gradient appliqué à l'image de référence ∇I_{k-1} . Chaque maille est à son tour subdivisée en maille rectangulaire de façon à ce que les contours soient isolés dans des mailles de taille plus petites. Un exemple de pavage est donné par la Figure 1, où les pointillés correspondent aux contours de l'image de référence. Les nœuds des mailles sont marqués par des points.

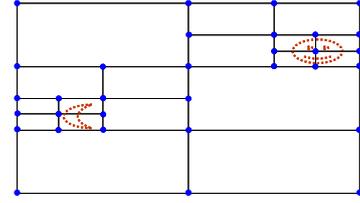


Figure 1 – Maillage adapté hiérarchique

Maillage semi régulier adapté

Dans ce type de maillage, nous procédons tout d'abord à un découpage horizontal (grossier) de l'image de référence I_{k-1} . Celui-ci est ensuite affiné horizontalement selon le gradient de l'image de référence. Pour l'exemple de la Figure 2, sept découpages horizontaux ont été appliqués sur l'image de gradient. Le processus est ensuite répété verticalement. Cinq découpages horizontaux sont réalisés pour construire le maillage de l'exemple de la Figure 2.

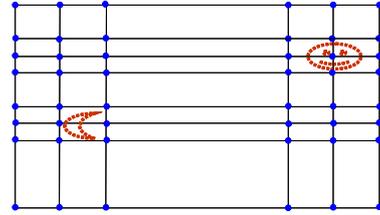


Figure 2 – Maillage adapté et semi régulier

Quelque soit le type de maillage adopté, les mailles forment une partition continue de l'image de référence.

3 Estimation de mouvement par maillage rectangulaire déformable

L'estimation de mouvement présentée dans cet article est basée sur le principe de la déformation de la grille maillée G_{k-1} construite au paragraphe précédent.

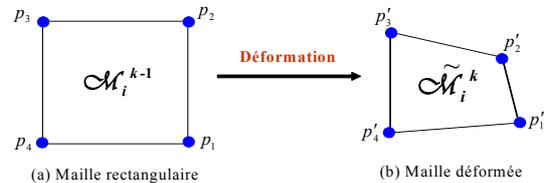


Figure 3 – Mailles : (a) rectangulaire, (b) déformée

Cette déformation se traduit par le déplacement des nœuds de chaque maille tout en préservant la continuité des mailles associées à l'image courante I_k . La Figure 3 présente une maille rectangulaire \mathcal{M}_i^{k-1} décrite par ses 4 nœuds p_1, p_2, p_3, p_4 . Cette maille est déformée en une maille quadrilatère, notée $\tilde{\mathcal{M}}_i^k$, décrite par ses 4 nœuds p'_1, p'_2, p'_3, p'_4 . Les déplacements optimaux des nœuds p_1, p_2, p_3, p_4 sont choisis de façon à minimiser l'erreur quadratique moyenne d'appariement des mailles. Nous

commençons par introduire sommairement la méthode de reconstruction basée sur la fonction spline bilinéaire.

3.1 Fonction spline bilinéaire

Ce paragraphe s'intéresse à la prédiction des pixels à l'intérieur d'une maille rectangulaire quelconque $\mathcal{M}_i^{k-l} \in G_{k-1}$ décrite par ses 4 positions nodales notées p_1, p_2, p_3, p_4 (voir Figure 3). Pour des commodités d'explications, notons les coordonnées de ces nœuds comme suit :

$$p_1 = (s_{i+1}, t_{i+1}), p_2 = (s_i, t_{i+1}), p_3 = (s_i, t_i)$$

et $p_4 = (s_{i+1}, t_i)$ avec $s_i < s_{i+1}$ et $t_i < t_{i+1}$.

De plus, les intensités des pixels en ces nœuds sont connues. Soit $f(s, t)$ la fonction d'interpolation spline bilinéaire définie sur le domaine rectangulaire suivant :

$$[s_i, s_{i+1}] \otimes [t_i, t_{i+1}]$$

où \otimes représente le produit tensoriel.

La fonction spline bilinéaire $f(s, t)$ est donnée par l'expression suivante ([24]) :

$$f(s, t) = \sum_{l=0}^1 \sum_{m=0}^1 a_{l,m} B_{l,[s_i, s_{i+1}]}^1(s) B_{m,[t_i, t_{i+1}]}^1(t)$$

où l'ensemble des coefficients $\{a_{l,m}\}$ représentent les coefficients de la surface spline.

L'ensemble des éléments

$$\{B_{l,[s_i, s_{i+1}]}^1(s) B_{m,[t_i, t_{i+1}]}^1(t)\}$$

constitue la base de l'espace spline. La l ième fonction B-spline non-uniforme $B_{l,[s_i, s_{i+1}]}^1(s)$ correspond à un polynôme de degré un. Donnons ci-dessous les expressions respectives des fonctions B-splines de degré un :

$$B_{0,[s_i, s_{i+1}]}^1(s) = (s - s_i) / (s_{i+1} - s_i) ;$$

$$B_{1,[s_i, s_{i+1}]}^1(s) = (s_{i+1} - s) / (s_{i+1} - s_i) ;$$

$$B_{0,[t_i, t_{i+1}]}^1(t) = (t - t_i) / (t_{i+1} - t_i) ;$$

$$B_{1,[t_i, t_{i+1}]}^1(t) = (t_{i+1} - t) / (t_{i+1} - t_i) ;$$

avec $s_i \leq s \leq s_{i+1}$ et $t_i \leq t \leq t_{i+1}$

Le calcul des coefficients de la surface spline est déterminé facilement ([25]) :

$$a_{0,0} = I_{k-1}(s_{i+1}, t_{i+1}), a_{1,0} = I_{k-1}(s_i, t_i), a_{0,1} = I_{k-1}(s_i, t_{i+1})$$

et $a_{1,1} = I_{k-1}(s_{i+1}, t_i)$

3.2 Prédiction de l'image courante à partir de la grille déformée

Ce paragraphe concerne la prédiction de l'image courante I_k à partir des nœuds appartenant à la grille déformée G_k . Supposons que les positions des nœuds d'une maille déformée soient connues. Les coordonnées de ces nœuds sont notées comme suit :

$$p'_1 = (x_i, y_i), p'_2 = (x_{i+1}, y_{i+1}), p'_3 = (x_{i+2}, y_{i+2}) \text{ et } p'_4 = (x_{i+3}, y_{i+3}).$$

Rappelons que la méthode de reconstruction décrite ci-dessus est valide dans le cas d'une maille rectangulaire puisque les éléments de la base spline sont construits sur

un domaine rectangulaire de type $[s_i, s_{i+1}] \otimes [t_i, t_{i+1}]$. L'approche utilisée dans cet article s'inspire des résultats développés dans le domaine des éléments finis ([26], [27]). Les déplacements des nœuds d'une maille rectangulaire \mathcal{M}_i^{k-l} sont modélisés par une fonction de déformation (warping) notée $w(s, t)$.

Dans cet article, la fonction de déformation est choisie comme étant une fonction d'interpolation bilinéaire ([28], [29]):

$$w(s, t) = \begin{bmatrix} \xi(s, t) \\ \eta(s, t) \end{bmatrix} = \sum_{i=0}^1 \sum_{j=0}^1 \underline{c}_{i,j} B_{i,[s_i, s_{i+1}]}^1(s) B_{j,[t_i, t_{i+1}]}^1(t)$$

où les coefficients $\underline{c}_{i,j}$ sont donnés par :

$$\underline{c}_{0,0} = [x_i, y_i], \underline{c}_{0,1} = [x_{i+1}, y_{i+1}], \underline{c}_{1,0} = [x_{i+2}, y_{i+2}], \\ \underline{c}_{1,1} = [x_{i+3}, y_{i+3}].$$

La fonction de déformation inverse $w^{-1}(x, y)$ permet de transformer la maille déformée (quadrilatère) en une maille rectangulaire (fonction bijective). De ce fait, la prédiction des pixels dans la maille déformée, est ramenée au calcul de la prédiction des pixels dans le référentiel de la maille rectangulaire par le biais de la fonction d'interpolation spline décrite précédemment. La Figure 4 résume les différentes étapes à suivre pour prédire l'image courante.

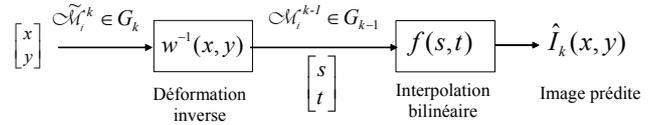


Figure 4 – Prédiction de l'image courante à partir de la grille déformée

3.3 Déformation de la grille de référence

Nous nous intéressons à la construction de la grille maillée de l'image courante G_k . La grille de l'image de référence G_{k-1} est superposée sur le gradient de l'image courante ∇I_k . La grille déformée G_k est construite progressivement. Les nœuds des mailles de la grille G_{k-1} sont déplacés sur le gradient de l'image ∇I_k de façon à minimiser l'erreur quadratique moyenne de prédiction. Nous accordons une importance particulière à cette phase d'optimisation puisque c'est l'erreur de prédiction (information résiduelle) qui est en général codée pour la compensation de mouvement.

L'erreur quadratique moyenne de prédiction des pixels dans la maille \mathcal{M}_i^k est calculée entre l'image courante prédite et l'image courante originale :

$$\varepsilon_p(\mathcal{M}_i^k) = \frac{1}{n(\mathcal{M}_i^k)} \sum_i (I_k(x_i, y_i) - \hat{I}_k(x_i, y_i))^2$$

où $\hat{I}_k(x_i, y_i)$ correspond au pixel prédit de coordonnée (x_i, y_i) et $n(\mathcal{M}_i^k)$ représente le nombre de pixels prédits dans la maille déformée \mathcal{M}_i^k .

On remarque dans ce contexte particulier que l'erreur de quadratique moyenne de prédiction est égale à l'erreur

quadratique moyenne d'appariement, notée $\varepsilon_A(\mathcal{M}_i^k)$, puisque en effet :

$$\varepsilon_A(\mathcal{M}_i^k) = \frac{1}{n(\mathcal{M}_i^k)} \sum_{i \in \mathcal{M}_i^k} (I_{k-1}(x_i - \Delta x_i, y_i - \Delta y_i) - I_k(x_i, y_i))^2$$

où $\hat{I}_k(x_i, y_i) = I_{k-1}(x_i - \Delta x_i, y_i - \Delta y_i)$ et $(\Delta x_i, \Delta y_i)$ représente le déplacement d'un noeud par rapport à sa position initiale (x_i, y_i) .

Pour réduire le temps de calcul nécessaire à la recherche optimale du déplacement d'un noeud, nous proposons de vérifier un certain nombre de points avant de prédire toutes les valeurs des pixels dans chacune des mailles $\{\mathcal{M}_i^k\}$ affectées par le déplacement du noeud. La démarche consiste à vérifier si le noeud que nous envisageons de déplacer : (i) appartient à un contour de l'image courante ($\nabla I_k(p_j) \neq 0$); (ii) n'engendre pas de discontinuité dans la structure de la grille G_k , et (iii) préserve la structure quadrilatère (pas de forme dégénérée). La dernière vérification se traduit par le calcul du Jacobien donné ci-dessous. Celui-ci doit être positif pour tous les points de la maille considérée ([28]):

$$J(s, t) = \begin{vmatrix} \frac{\partial \xi(s, t)}{\partial s} & \frac{\partial \eta(s, t)}{\partial s} \\ \frac{\partial \xi(s, t)}{\partial t} & \frac{\partial \eta(s, t)}{\partial t} \end{vmatrix} > 0$$

4 Analyse des résultats expérimentaux

Dans cette section, nous présentons les résultats obtenus sur la séquence vidéo de test « Miss America » au format QCIF (4:2:0). La résolution spatiale de chaque image est de 144×176 et la résolution temporelle est de 12 images par seconde. Nous présentons les résultats d'estimation de mouvement obtenus sur les 11 premières images de la séquence vidéo numérotées de 0 à 10.

La première image « 0 » est choisie comme image de référence (de type I). Celle-ci est représentée par les trois composantes; I_0 (luminance), Cb_0 (chrominance bleue) et Cr_0 (chrominance rouge). La phase initiale de la méthode d'estimation de mouvement consiste à construire la grille maillée de référence G_0 associée à l'image d'intensité I_0 . Rappelons que le maillage de référence est construit à partir des informations liées au gradient de l'image ∇I_0 . Trois types de découpages sont analysés et comparés dans cette section.

Commençons par présenter le maillage hiérarchique adapté. L'image de gradient ∇I_0 est tout d'abord subdivisée en plusieurs mailles de taille identique. Dans cet exemple, nous avons fixé la taille initiale de chaque maille à 49×49 . Certaines de ces mailles initiales, en accord avec l'image de gradient ∇I_0 , peuvent être ensuite subdivisées en mailles rectangulaires. Dans cet exemple, la taille de la plus petite maille est égale à 4×4 . La Figure 5 présente le découpage superposé sur le gradient de l'image I_0 . Il serait judicieux que les tailles initiale et finale des mailles puissent être ajustées par le codeur en fonction par exemple du débit imposé et de la qualité désirée de la séquence vidéo décodée. Cet aspect

d'optimisation n'est pas considéré dans cet article et fera l'objet d'investigations futures.

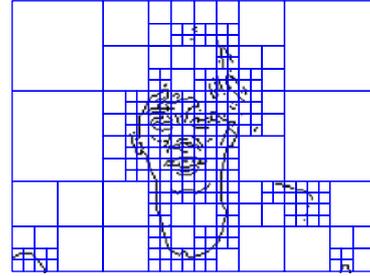


Figure 5 – Maillage rectangulaire de la grille de référence

La Figure 6 présente la déformation de la grille de référence G_0 . Dans cet exemple, la grille G_0 est composée de 779 noeuds. Pour la phase d'estimation de mouvement décrite au paragraphe 3, parmi les 779 noeuds un premier tri des noeuds montre que seuls 524 noeuds sont autorisés à être éventuellement déplacé si les conditions (i) et (iii) sont vérifiées. Pour les composantes de chrominance (Cb_6, Cr_6), il suffit de redimensionner la grille déformée G_6 par un facteur 4 puisque le format utilisé correspond à 4:2:0.

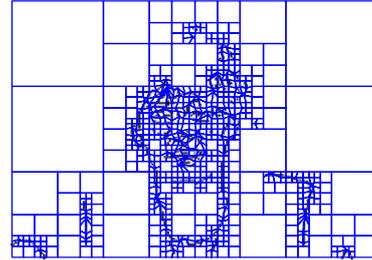


Figure 6 – Maillage hiérarchique G_6 associé à l'image courante I_6

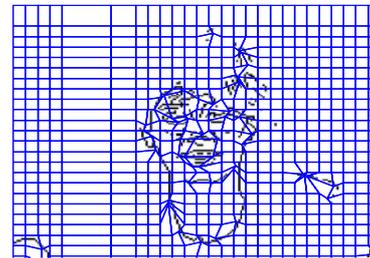


Figure 7 – Maillage semi régulier G_6 associé à l'image courante I_6

Afin de comparer les performances des trois types de maillage, nous avons ajusté le nombre de noeuds de chaque grille de référence de façon à travailler dans des conditions identiques. Pour un maillage rectangulaire uniforme, la taille de chaque maille est fixée à 7×7 (voir

Figure 8). Dans ce cas la grille régulière de référence G_0 est composée de 775 nœuds. La Figure 7 présente la déformation de la grille G_6 dans le cas d'un maillage semi régulier adapté composé de 784 nœuds.

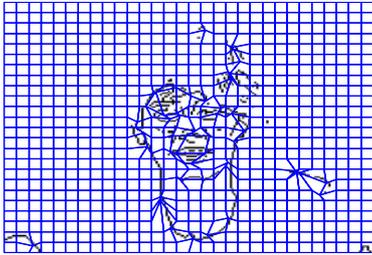


Figure 8 – Maillage uniforme G_6 associé à l'image courante I_6

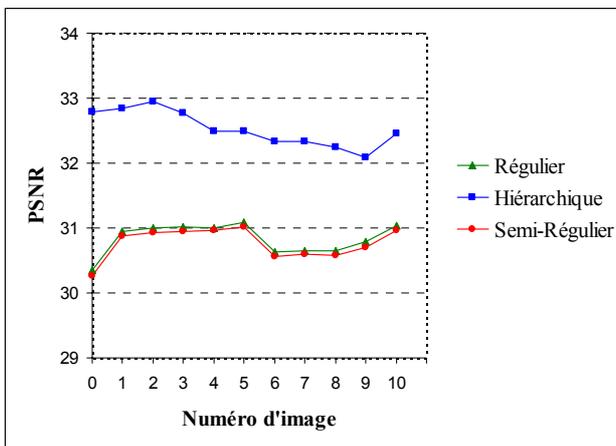


Figure 9 – PSNR en fonction du numéro de l'image prédite dans la séquence vidéo ($I_0 - I_{10}$)

La Figure 9 compare les erreurs quadratiques moyennes de prédiction (PSNR) des images d'intensité engendrées par les 3 types de déformation de maillage : uniforme, semi régulier et hiérarchique. Ces courbes montrent que nous obtenons de meilleurs résultats pour le maillage hiérarchique adapté. De plus, le temps nécessaire pour la phase de recherche des positions optimales est moins important comparé aux deux autres types maillages. Dans cet exemple particulier, parmi les 775 nœuds de la grille régulière, seuls 667 nœuds sont susceptibles d'être déplacé (contre 524 nœuds pour le maillage hiérarchique).

5 Conclusion

Dans cet article, nous nous sommes intéressés à l'estimation de mouvement à partir de maillage déformable adapté au gradient de l'image dans le but d'améliorer l'approche classique basée sur l'appariement de blocs. Notre méthode d'estimation nécessite, dans une première phase, la construction de la grille de référence associée à une image sélectionnée (de type Intra) dans la

séquence vidéo. Le déplacement des nœuds de cette grille permet de modéliser le mouvement des objets dans l'image courante. La recherche de la position optimale du nœud à déplacer est dictée par la position qui engendre la plus petite erreur quadratique moyenne de prédiction dans les mailles. Les résultats de simulations ont montré que le maillage hiérarchique est le mieux adapté à l'estimation de mouvement. De plus, ce type de maillage nécessite beaucoup moins de calcul lors de la phase de recherche de la position optimale des nœuds.

Cette étude préliminaire sera complétée dans des travaux de recherche futurs. Nous envisageons de traiter un ensemble de points importants. Nous analyserons l'influence du degré de la fonction spline sur l'erreur de prédiction puisque cette erreur sera codée ultérieurement pour réaliser la compensation de mouvement. Nous étudierons également les problèmes de codage induits par le maillage. L'ajustement des paramètres, tels que le seuil de détection de contours, la taille des mailles initiale et finale, en accordance avec le débit offert sera considéré. Le but final de ce travail serait de proposer une chaîne complète de codage où l'estimation et la compensation de mouvement seront basées sur le principe de maillage déformable.

Références

- [1] ISO/IEC 13818. *Information Technology – Generic Coding of Moving Pictures and Associated Audio Information*. 2000.
- [2] ITU-T, Recommendation H.263 (version 2). *Video Coding for low Bitrate Communications*. February 1998.
- [3] ISO/IEC 14496-2. *Coding of Audio-Visual Objects – Part 2: Visual*. 2001.
- [4] F. Pereira and T. Ebrahimi. *The MPEG-4 Book*. IMSC Press, 2002.
- [5] ISO/IEC 14496-10 and ITU-T Rec. H.264. *Advanced Video Coding*. 2003.
- [6] Iain E. G. Richardson. *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia.*, Ed. John Wiley & Sons, 2003.
- [7] Haibo Li, A. Lundmark, R. Forchheimer. Image sequence coding at very low bit rates: a review. *IEEE Transactions on image processing* 3(5) : 589–609, September 1994.
- [8] J. R. Jain, A. K. Jain. Displacement measure and its application in interframe image coding. *IEEE Transactions on communication*, 29(12) : 1799–1806, December 1981.
- [9] M. H. Chan, Y. B. Yu, A.G. Constantinides. Variable size block matching motion compensation with applications to video coding. *Communications, Speech and Vision, IEE Proceedings*, 137(4) : 205–212, August 1990.
- [10] H. Brusewitz. Motion compensation with triangles. *Proc. 3rd Inter. Conf on 64 kbit Coding of moving*

- Video*, Rotterdam, Netherlands, September. 1990.
- [11] Y. Nakaya, H. Harashima. An iterative motion estimation method using triangular patches for motion compensation. *Proc. SPIE Visual Comm. and Image* 546–557, Boston, MA, November. 1991.
- [12] G. J. Sullivan, R. L. Baker. Motion compensation for video compression using grid interpolation. *ICASSP'91*, pages 2713–2716, Toronto, Canada, May 1991.
- [13] Y. Nakaya, H. Harashima. Motion compensation based on spatial transformation. *IEEE Transactions on circuits and systems for video technology*, 4(3) : 339–356, June 1994.
- [14] Aria Nosratinia, M. T. Orchard. Optimal warping prediction for video coding. *ICASSP'96*, pages 1986–1989, Atlanta, May 1996.
- [15] Aria Nosratinia. New Kernels for Fast Mesh-Based Motion Estimation. *IEEE Transactions on circuits and systems for video technology*, 11(1) : 40–51, January 2001.
- [16] C. L. Huang, C. Y. Hsu. A new motion compensation method for image sequence coding using hierarchical grid interpolation. *IEEE Transactions on circuits and systems for video technology*, 4(1) : 13–52, February 1994.
- [17] Y. Wang, O. Lee. Active mesh-A feature seeking and tracking image sequence representation scheme. *IEEE Transactions on image processing*, 3(5) : 610–624, September 1994.
- [18] Gary J. Sullivan, Richard L. Baker. Efficient Quadtree Coding of Images and Video. *IEEE Transactions on image processing*, 3(3) : 327–331, May 1994.
- [19] Haibo Li, Astrid Lundmark, Robert Forchheimer. Image Sequence Coding at Very Low Bitrates: A Review. *IEEE Transactions on image processing*, 3(5) : 589–609, September 1994.
- [20] C. Toklu, A. T. Erdem, M. I. Sezan, A. M. Tekalp. Tracking motion and intensity variations using hierarchical 2D mesh modelling for synthetic object transfiguration. *Graph. Models Image Process*, 58(6) : 553–573, November 1996.
- [21] Y. Altunbask and A. M. Tekalp. Closed-form connectivity-Preserving solutions for motion compensation using 2-D meshes. *IEEE Transactions on image processing*, 6(9) : 1255–1269, September 1997.
- [22] P. V. Beek, A. Tekalp, N. Zhang, I. Celasum, M. Xia. Hierarchical 2D-mesh representation tracking and compression for object video. *IEEE Transactions on image processing*, 9(2) : 353–369, Mars 1999.
- [23] Ghassan Al-Regib, Yucel Altunbasak, Russell M. Mersereau. Hierarchical Motion Estimation With Content-Based Meshes. *IEEE Transactions on circuits ad systems for video technology*, 13(10) : 1000–1005, October 2003.
- [24] C. De Boor. *A practical guide to splines*, revised edition, New York, Springer-verlag, 2001.
- [25] N. Chihab, A. Zergainoh, P. Duhamel, J-P. Astruc. The influence of the non-uniform spline basis on the approximation signal. *EUSIPCO'04*, Vienna, Austria, September 2004.
- [26] E. B. Becker, G. F. Gatey, J. T. Oden. *Finite Elements, An introduction*. Englewood Cliffs, NJ Prentice-Hall, 1982.
- [27] K. J. Bathe. *Finite Element procedures in engineering analysis*. Englewood Cliffs, NJ Prentice-Hall, 1982.
- [28] Y. Wang and O. Lee. Use of two-dimensional deformable mesh structures for video coding, part I-The synthesis problem: Mesh-based function approximation and mapping. *IEEE Transactions on circuits and systems for video technology*, 6(6) : 636–646, December 1996.
- [29] Y. Wang, O. Lee, A. Vetro. Use of two-dimensional deformable mesh structures for video coding, part II-The analysis problem and a region based coder employing an active mesh representation. *IEEE Transactions on circuits and systems for Video Technology*, 6(6) : 647–659, December 1996.