

# Approximation d'images par minimisation d'une distance géométrique appliquée à un modèle d'éléments finis 3D

Patrick Lechat, Henri Sanson et Lilian Labelle

France Télécom - CNET / DIH, Laboratoire du CCETT  
4, rue du Clos Courtel - 35512 Cesson Sévigné Cedex - FRANCE  
{plechat|sanson}@ccett.fr

## RÉSUMÉ

Ce papier présente un modèle géométrique pour le codage d'images, permettant un échantillonnage adaptatif, et se présentant sous la forme d'une structure de maillage 3D de type éléments finis. Nous avons défini une distance géométrique caractérisant l'erreur entre l'image originale et son modèle et développé un algorithme de gradient efficace permettant une déformation adéquate du modèle. Le rendu 2D résultant constitue une représentation performante en terme de qualité subjective, offrant une bonne restitution des caractéristiques visuellement importantes.

## ABSTRACT

We propose in this paper a method for image coding based on an adaptive non-uniform sampling scheme. The aim is to optimize a 3D mesh-based model such as those involved in finite elements method. To this end, a geometrical distance has been defined, measuring the error between the model and the original image, and a gradient algorithm has been derived to obtain an accurate model deformation. The corresponding 2D rendering leads to an efficient image representation : salient features such as edges are preserved, offering a good subjective quality.

## 1 Introduction

Les développements audio-visuels récents mettent en évidence la nécessité de transmettre et stocker toujours davantage d'information. De nombreux efforts ont donc été entrepris en terme de codage et des méthodes éprouvées, maintenant normalisées, ont vu le jour. Cependant, ces méthodes traditionnelles, dites de première génération, s'appuient sur des transformations en cosinus discrètes (DCT) ou sous-bandes, efficaces pour un signal stationnaire mais peu adaptées à de fortes variations locales et générant des effets de bloc. De plus, elles offrent moins de souplesse pour l'implantation de nouvelles fonctionnalités, par exemple celles définies par MPEG4.

Contrairement à ces méthodes traditionnelles qui échantillonnent régulièrement le signal, nous proposons ici une approche de codage d'images s'appuyant sur un échantillonnage adaptatif de celui-ci. Ce signal est ensuite reconstruit par utilisation de fonctions polynômiales par morceaux, de type éléments finis. Le modèle résultant apparaît sous la forme d'un maillage tridimensionnel de type triangulaire composé d'un ensemble de nœuds appelés points de contrôle, connectés via un ensemble d'arcs. Ce modèle, de par sa structure, permet la mise en œuvre de nombreuses fonctionnalités définies par le projet de standardisation ISO-MPEG4 :

- La compression : le nombre relativement peu important de points de contrôle constituant le modèle laisse présager des taux de compression intéressants.
- Le codage basé région : toute région de contour polygonal peut être extraite, manipulée, déformée et codée de façon indépendante.
- Le changement d'échelle spatiale : par l'utilisation d'un maillage hiérarchique pyramidal.

- L'interpolation temporelle : l'estimation de mouvement est appliquée directement au maillage.
- Le codage hybride : la structure de notre modèle permet une intégration rapide dans une scène synthétique.

Cet article présente les fondements théoriques de notre méthode d'approximation : comment, à partir de la définition d'une distance géométrique mesurant l'erreur entre le modèle et l'image originale, nous avons déduit une formulation de la variation appliquée à chacun des points de contrôle permettant de la minimiser.

Nous nous attacherons ensuite à montrer les moyens algorithmiques ayant permis l'implémentation de la méthode, et particulièrement les mécanismes de rendu permettant l'obtention de l'image résultante à partir de notre modèle tridimensionnel.

Enfin, nous présenterons les résultats : ceux obtenus après application de la méthode sur une grille régulière afin d'évaluer les performances en terme de rapidité de convergence, puis ceux obtenus après une étape de division-fusion.

## 2 Fondements théoriques

Les images sont naturellement représentées sous forme explicite  $z = g(x, y)$  : si  $\hat{g}(x, y)$  représente la fonction approchée, nous pouvons alors minimiser la norme  $L^2$  :  $\|g - \hat{g}\|_2$ , pour réaliser notre approximation. Cependant nous voyons que ce critère n'autorise qu'une variation verticale et ne permet pas une évolution dans le plan  $(x, y)$  des points de contrôle du maillage.

Il est donc préférable de représenter la surface de l'image sous forme paramétrée :  $m(u, v) = (x(u, v), y(u, v), z(u, v))$ . Cependant, la minimisation de  $\|m - s\|_2$ ,  $s(u, v)$  étant la

surface du modèle, est très dépendante de la paramétrisation choisie et n'autorise pas une évolution optimale des points du maillage [4].

La solution retenue a donc consisté à définir une distance géométrique  $E(m, s)$ . Cette distance correspond au volume  $V(m, s)$  séparant la surface originale  $s$  de son modèle  $m$  (figure 1)[3] [2].

$$E(m, s) = \int_{\mathbb{R}^3} I_{V(m,s)}(q) d^3q \quad (1)$$

Où  $I_V(q)$  est la fonction indicatrice et  $q$  un point de  $\mathbb{R}^3$ .

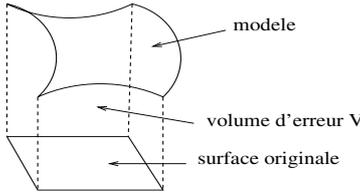


FIG. 1 — Représentation du volume d'erreur V

En considérant  $X_i$  comme les points de contrôle du maillage et  $B_i^k(u, v)$  une base polynômiale par morceaux de degré  $k$ , on en déduit la variation de  $E$  par rapport à ces points :

$$\frac{\partial E}{\partial X_i} = - \iint_{u,v} \eta(u, v) n(u, v) \cdot B_i^k(u, v) d\sigma(u, v) \quad (2)$$

Où la fonction  $\eta(u, v) = 1$  si la normale unitaire  $n(u, v)$  est orientée vers l'intérieur du volume, et  $-1$  dans le cas contraire. Cette expression est calculée par morceaux sur les surfaces élémentaires (éléments finis).

L'algorithme de pseudo-gradient déduit, appliqué aux points  $X_i$  à l'itération  $k$ , est le suivant :

$$X_i^{k+1} = X_i^k - \gamma_i \frac{\partial E^k}{\partial X_i} / \left\| \frac{\partial E^k}{\partial X_i} \right\|_{\infty} \quad (3)$$

où les  $\gamma_i$  constituent des pas indépendants. Ceux-ci sont divisés par 2 dans le cas où le vecteur de déplacement  $\frac{\partial E^k}{\partial X_i}$  (dim. 3) du point  $X_i^k$  vérifie  $\frac{\partial E^{k+1}}{\partial X_i} \cdot \frac{\partial E^k}{\partial X_i} < 0$ , permettant d'accélérer la convergence et de limiter le comportement oscillatoire de l'algorithme.

### 3 Implémentation algorithmique

La principale difficulté de la méthode consiste à évaluer la quantité  $\frac{\partial E}{\partial X_i}$  (2). La méthode proposée a été implémentée dans le cas d'un modèle d'éléments finis de continuité  $C^0$  où la base  $B_i^k(u, v)$  est affine ( $k = 1$ ). La normale  $n(u, v)$  est donc constante sur chacune des surfaces élémentaires constituées par les éléments finis du maillage.

#### 3.1 Optimisation du modèle

Pour chacune des surfaces triangulaires élémentaires, nous calculons la normale  $n$  permettant l'identification d'un repère local  $(X_1, u, v, w)$  tel que  $w = n$ ,  $u = \frac{X_1 X_2}{\|X_1 X_2\|}$  et  $v = w \wedge u$ , où  $(X_1 X_2 X_3)$  représente l'élément fini local (voir figure 2).

Nous en déduisons la transformation suivante :

$$X = T + R\chi \quad (4)$$

Où :  $X(x, y, z)^t$  dans  $(O, i, j, k)$   
 $\chi(x', y', z')^t$  dans  $(X_1, u, v, w)$

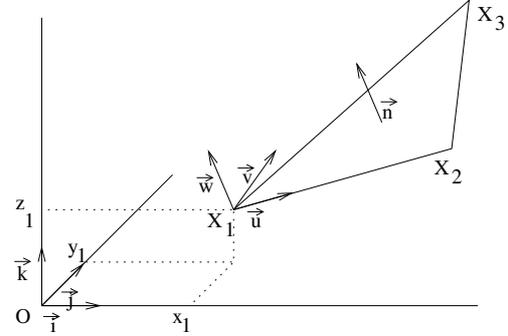


FIG. 2 — Discrétisation de  $\frac{\partial E}{\partial X_i}$

Un rendu 2D est réalisé dans le plan  $(X_1, u, v)$ , rendu qui fournit une discrétisation adéquate de l'intégrale intervenant dans le calcul de  $\frac{\partial E}{\partial X_i}$  (choix des points de discrétisation et calcul des  $\sigma(u, v)$  correspondants). La base  $B_i^k$  utilisée étant affine, elle correspond à l'utilisation de coordonnées barycentriques calculées sur  $(X_1 X_2 X_3)$ . La transformation (4) permet de connaître le point correspondant  $X(x, y, z)^t$  de chacun des points  $\chi(x', y', z' = 0)^t$  du rendu 2D, de façon à identifier  $\eta(u, v)$ .

Une fois l'expression  $\frac{\partial E}{\partial X_i}$  calculée, le déplacement correspondant (3) est appliqué au point de contrôle  $X_i$  avec, si nécessaire, réduction du pas  $\gamma_i$  pour conserver la topologie du maillage (figure 3).

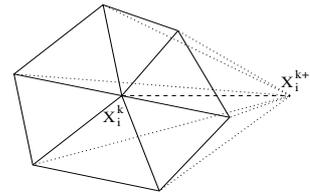


FIG. 3 — Cas de non-conservation de la topologie : réduction du pas  $\gamma_i$  associé à  $X_i$  nécessaire

#### 3.2 Rendu 2D du modèle

Deux types d'approches ont été testées pour réaliser le rendu 2D de notre modèle tridimensionnel, l'une de continuité  $C^0$  directement calculée à partir de notre modèle, l'autre de continuité  $C^1$  permettant de limiter la perception visuelle des arêtes du maillage.

**Approches  $C^0$ .** Une première technique rapide utilisant un rendu de type *Gouraud shading* a été implémentée, utilisant l'algorithme de *Bresenham* pour le tracé des lignes et des dégradés de couleur. Cette approche est cependant inadaptée lorsque les facettes planes du modèle s'approchent de la verticale.

La seconde méthode est plus coûteuse en calculs mais exacte : elle consiste à estimer  $z$  de  $X(x, y, z)^t$  connaissant  $x$  et  $y$ . La topologie du maillage initiale étant à tout moment respectée, à chaque point discret  $X(x, y, 0)^t$  de l'image correspond un unique élément fini où la transformation (4) est connue : sachant que  $z' = 0$ , on en déduit  $z$  (luminance exacte).

De façon équivalente, on peut utiliser la fonction d'interpolation  $\pi$  appliquée sur la surface  $z = v(x, y)$  :

$$\pi(v) = \sum_{i=1}^3 \lambda_i v(a_i)$$

Où  $(a_1 a_2 a_3)$  représente la projection de l'élément  $(X_1 X_2 X_3)$  sur le plan  $(x, y)$  de l'image et  $(\lambda_1, \lambda_2, \lambda_3)$  les coordonnées barycentriques associées à  $(a_1 a_2 a_3)$ .

**Approche  $C^1$ .** La continuité  $C^1$  est réalisée en utilisant l'élément fini de *Hermitte*. Celui-ci consiste à interpoler la surface de chaque élément triangulaire à l'aide d'une base polynômiale du troisième degré. La continuité est assurée via le calcul des deux tangentes directionnelles  $Dv(a_i)(a_{i+1} - a_i)$  et  $Dv(a_i)(a_{i-1} - a_i)$  en chacun des trois points  $a_i$  constituant l'élément (figure 4). Ces tangentes sont prises orthogonales au vecteur normal moyen de tous les éléments triangulaires adjacents au point  $a_i$  considéré.

La fonction d'interpolation est la suivante :

$$\pi(v) = \sum_{i=1}^3 v(a_i)q_i^0 + Dv(a_i)(a_{i-1} - a_i)q_{i,i-1}^1 + Dv(a_i)(a_{i+1} - a_i)q_{i,i+1}^1$$

Où :

$$q_i^0 = -2\lambda_i^3 + 3\lambda_i^2 + 2\lambda_1\lambda_2\lambda_3$$

$$q_{i,i-1}^1 = \frac{1}{2}\lambda_i\lambda_{i-1}(\lambda_i - \lambda_{i-1} + 1)$$

$$q_{i,i+1}^1 = \frac{1}{2}\lambda_i\lambda_{i+1}(\lambda_i - \lambda_{i+1} + 1)$$

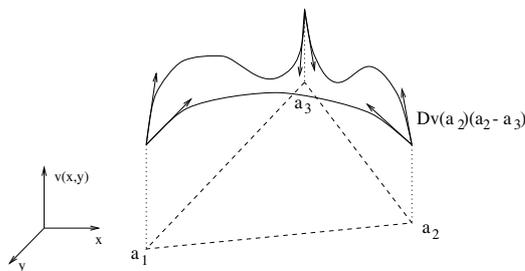


FIG. 4 — Interpolation  $C^1$  de *Hermitte*

## 4 Résultats

Notre méthode est applicable sur un maillage de topologie quelconque : il est donc approprié de faire précéder notre optimisation par une étape de division-fusion [1]. Celle-ci permet de concentrer les points de contrôle dans les zones critiques (contours, arêtes) et de les disséminer dans les zones uniformes. Notre optimisation appliquée sur un tel maillage génère alors un placement optimal des points permettant la reconstruction la plus fidèle possible.

Pour mesurer l'efficacité de notre méthode d'optimisation, nous l'avons appliquée sur une grille régulière 40x40 (figures 6). Dix itérations successives ont été réalisées : le calcul du PSNR (figure 5) donne une mesure objective du gain en qualité

et permet de vérifier la convergence rapide de l'algorithme, bien qu'il ne corresponde pas à la distance géométrique minimisée. Cette convergence peut être accélérée ou ralentie par le pas  $\lambda_i$  initial choisi.

Cependant, l'apport principal de la méthode constitue le gain en qualité subjective : le maillage déformé est cohérent avec le contenu sémantique de l'image, les arêtes se placent le long des contours, de sorte à offrir un meilleur aspect visuel (figures 7). Les figures 8 montrent une optimisation après une étape de division-fusion.

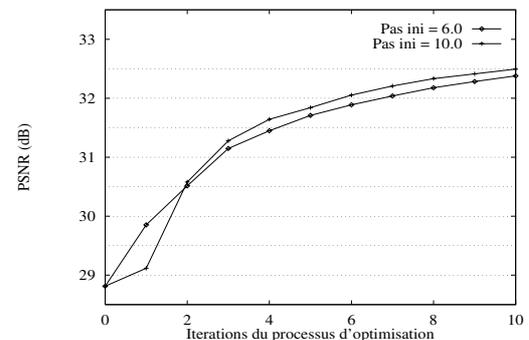


FIG. 5 — Evolution du PSNR suivant le nombre d'itérations (image fig. 6, pas  $\lambda_i$  initialisé à 6 et 10 pixels).

## 5 Conclusion

Nous avons présenté dans cet article une méthode de déformation de maillage triangulaire quelconque permettant une représentation efficace d'images. Celle-ci s'inscrit en complémentarité des approches division-fusion et offre de réelles perspectives dans un schéma de codage de séquences vidéo orienté région : la structure de maillage se prête très facilement au découpage en régions polygonales et le déplacement calculé en chacun des noeuds du modèle permet une estimation de mouvement inter-images.

La méthode peut également être intrinsèquement améliorée par une optimisation du maillage basée sur le rendu  $C^1$  pour une approximation optimale des régions uniformes, et un interpolateur local hautes fréquences pour coder les textures et zones à forte densité de détails.

## Références

- [1] K. W. Chum, B. Jeon, and J. M. Jo. Irregular triangular Mesh Representation Based on Adaptive Control Point Removal. *SPIE*, vol. 2727, 844-853, 1996.
- [2] L. Labelle. *Représentation Adaptative d'Images Appliquée à un Schéma de Codage Orienté Régions de Séquences TV*. PhD thesis, Université de Rennes I, 1996.
- [3] L. Labelle and H. Sanson. Computation of Image Representation Based on Active Triangular Meshes Through Geometrical Surface Evolution. In *Proc. of VCIP'97*, San Jose, CA, USA, Feb 10 - 14 1997.
- [4] O. Lee and Y. Wang. Non Uniform Image Sampling and Interpolation Over Deformed Meshes and its hierarchical extension **2501** part I conf. 3, pp. 389-400. *Proc. conf. VCIP, Taiwan*, 1995.



FIG. 6 — Grille 40x40 originale (a) et rendu associé (b)



FIG. 7 — Grille optimisée (10 itérations) (a), rendus associés  $C^0$  (b) et  $C^1$  (c)

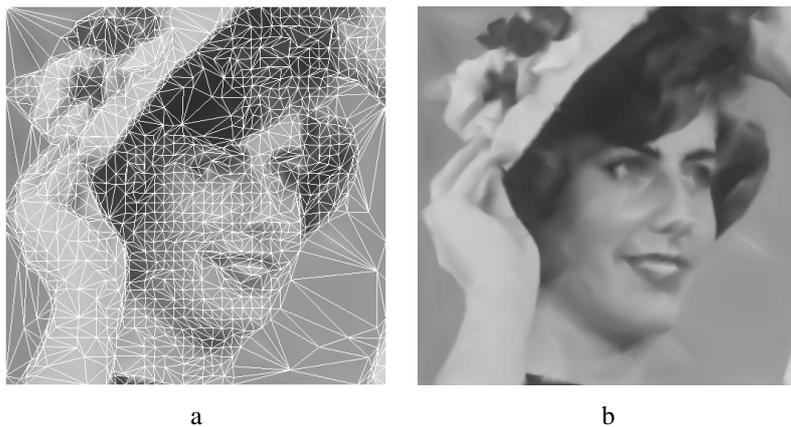


FIG. 8 — Grille non régulière (a) et rendu associé (b)