

Morphing 3D Automatique de Cartes de Profondeurs Maillées

Raphaèle Balter¹

Franck Galpin²

Luce Morin³

¹France Telecom RD, 4 rue du Clos Courtel, 35512 Cesson-Sévigné

²Tohoku University, Aramaki-aza Aoba01, Aoba-ku, Sendai 980-8579, Japan

¹ et ³ IRISA-INRIA, Campus de Beaulieu, avenue du Général Leclerc, 35042 Rennes, France

rbalter@rd.francetelecom.com, galpin@fractal.is.tohoku.ac.jp, lmorin@irisa.fr

Résumé

Cet article traite de la représentation de séquences vidéo de scènes statiques pour le codage très bas débit. A partir de vidéo on peut extraire automatiquement un flux de modèles 3D dont la reprojexion permet d'obtenir une séquence virtuelle identique à l'originale. Mais le fait d'utiliser plusieurs modèles introduit des artefacts visuels lors de la transition d'un modèle à l'autre. Le morphing 3D semble être la solution adaptée à ce problème. Or les méthodes de morphing 3D de maillages classiques nécessitent une intervention manuelle ce qui est en contradiction avec l'automatisme de la chaîne existante. On donne ici une solution à cette limite en présentant une méthode de morphing 3D de maillages entièrement automatique.

Mots clefs

Reconstruction 3D, codage très bas débit, morphing.

1 Introduction

Cet article se place dans le contexte du codage bas débit de séquences vidéo. Dans [Gal] Franck Galpin a proposé un schéma de codage basé sur un flux de modèles 3D extraits automatiquement à partir d'une vidéo acquise par une caméra monoculaire en mouvement. Les modèles 3D estimés sont en fait valides pour un groupe d'images (GOP). Ce schéma a de nombreux avantages pour la compression très bas débit (16 kb/s pour des images au format CIF à 25 Hertz) et l'utilisation de plusieurs modèles successifs pour la même scène permet une grande robustesse et une généralité pour les longues séquences (aucune hypothèse n'est faite sur le contenu de la scène ou sur la calibration de la caméra). Enfin le fait de passer par des modèles 3D permet des fonctionnalités de haut niveau par rapport aux modèles 2D comme les changements d'illumination, la stabilisation, l'insertion d'objets ou la navigation libre.

La navigation libre correspond à un changement de trajectoire de la caméra par rapport au chemin original. Dans ce contexte l'utilisation de plusieurs modèles 3D introduit des limites lorsque l'on s'éloigne de la trajectoire originale de la caméra. Pour garantir une transition douce entre deux modèles successifs il est donc nécessaire d'ajouter un post-

traitement. Il y a alors deux possibilités:

- le *fondus 3D* qui correspond à une transition progressive entre deux images successives
- le *morphing 3D* qui est basé sur la génération de modèles 3D intermédiaires entre les modèles donnés que l'on va utiliser pour générer la séquence décrivant la transformation.

Dans [Gal] un fondus 3D, qui consiste à calculer une moyenne pondérée de la projection des deux modèles considérés, a été proposé. Cette technique donne des résultats acceptables si la séquence reconstruite suit la trajectoire originale, mais cela introduit toujours des ruptures dans le cas de la navigation libre.

Le morphing 3D est plus adapté aux sauts de trajectoire mais les méthodes stables numériquement et performantes de la littérature reposent sur une intervention manuelle plus ou moins conséquente. Dans le cas plus particulier des maillages, le morphing permettant de passer d'un maillage source à un maillage cible comporte deux étapes principales:

- La mise en correspondance de chaque sommet du maillage source avec un sommet du maillage cible (et vice versa).
- L'interpolation qui définit la trajectoire entre chaque paire de points se correspondant.

Un bon état de l'art des techniques de morphing jusqu'en 1998 est donné dans [LV98]. Après cette date des approches intéressantes ont été proposées [ZSH00, LDSS99, MKFC01] ainsi qu'un autre état de l'art [Ale02], mais les méthodes proposées jusqu'à ce jour, et procurant des résultats de bonne qualité et numériquement stables reposent toutes sur une intervention manuelle plus ou moins lourde pour la première étape.

Cela nous pose problème puisque l'algorithme proposé dans [Gal] est totalement automatique et que l'on voudrait donc conserver cet avantage. Il nous faut donc mettre au point une méthode de morphing qui permettra d'obtenir une qualité visuelle satisfaisante mais qui, contrairement aux méthodes existantes, sera complètement automatique.

2 Le morphing prédictif

Dans notre cas les sommets de deux maillages définissant deux modèles successifs ne correspondaient pas aux mêmes points physiques. Pour établir la liste des correspondants nous utilisons une interpolation et un remaillage de la carte de profondeur obtenue par compensation du champ de mouvement. Les occultations sont gérées par un Z-buffer.

La deuxième étape est une interpolation linéaire entre les paires de points afin d'obtenir les modèles intermédiaires.

2.1 Description de la méthode

Dans notre cas les sommets de deux maillages définissant deux modèles successifs ne correspondent pas aux mêmes points physiques. En effet les modèles étant approximés à partir de maillages réguliers des images clés, les sommets de modèles différents ne correspondent pas au même point 3D.

Il faut donc établir une correspondance entre les points de M_{n+1} et ceux de M_n . On introduit pour cela des modèles obtenus à partir de M_n et M_{n+1} , que nous noterons respectivement M_n^i et M_{n+1}^i , le i signifiant que c'est un modèle qui va être exploité pour l'interpolation. Ainsi pour réaliser le morphing 3D de deux modèles successifs M_n et M_{n+1} on a choisi de calculer pour chaque position de la caméra intermédiaire un modèle courant M_c qui est une moyenne pondérée entre M_n^i et M_{n+1}^i .

Afin de permettre de travailler sur le même point physique le modèle M_n^i est un remaillage de M_n et M_{n+1}^i doit être un remaillage de M_{n+1} ayant les propriétés suivantes:

- les sommets de M_{n+1}^i sont les correspondants des sommets de M_n^i ,
- la topologie de M_{n+1}^i est identique à celle de M_n^i .

Ensuite une interpolation linéaire est appliquée entre les paires de points afin d'obtenir les modèles intermédiaires.

Les points 3D du modèle courant M_c pourront donc être obtenus à partir de ceux de M_n^i et M_{n+1}^i par la formule :

$$M_c = \alpha * M_n^i + (1 - \alpha) * M_{n+1}^i$$

avec

$$\alpha = \frac{\|t_{t_{n+1}} - t_{t_c}\|}{\|t_{t_{n+1}} - t_{t_n}\|}$$

où t_{t_c} représente la translation associée à la position de la caméra à l'instant courant, t_{t_n} celle correspondant à l'instant t_n (correspondant à K_n) et $t_{t_{n+1}}$ à l'instant t_{n+1} (correspondant à K_{n+1}). Par sa définition en fonction des translations t , α vaut 1 en début de GOP et permet donc de privilégier M_n , et décroît jusqu'à valoir 0 en fin de GOP afin d'augmenter progressivement l'influence de M_{n+1}^i .

Les images de la séquence reconstruite s'obtiennent donc par la projection du modèle courant M_c sur le point de vue associé à C_c et texturé avec T_c :

$$I_c = Pr(M_c, T_c, C_c)$$

L'image de texture T_c correspond à une interpolation linéaire entre T_n et T_{n+1} :

$$T_c = \alpha T_{t_n} + (1 - \alpha) T_{t_{n+1}}$$

Il reste maintenant à établir la méthode d'obtention de M_{n+1}^i .

2.2 Calcul de M_{n+1}^i

L'objectif est donc d'obtenir un remaillage de M_{n+1} correspondant aux mêmes points physiques que M_n et ayant même topologie que M_n .

La profondeur représente ici la distance entre un point et le centre de la caméra.

La méthode adoptée pour obtenir M_{n+1}^i est de construire la carte de profondeur de M_{n+1} vu de C_n et de remailler avec les mêmes sommets que M_n .

Pour cela on projette les sommets du maillage 3D M_{n+1} sur K_n . Les sommets obtenus sont placés de manière irrégulière et pour chacun d'eux on calcule la distance entre le point 3D et la caméra C_n . A partir de ces valeurs de profondeurs obtenues et de la position des sommets projetés on réalise une interpolation (triangle par triangle) afin d'obtenir une carte de profondeur dense pour M_{n+1}^i :

$$Z_{n+1}^n = ZMap(M_{n+1}, C_n)$$

Ensuite on remaille régulièrement l'image Z_{n+1}^n obtenue afin d'obtenir les profondeurs de M_{n+1} pour les sommets de M_n .

M_n^i est donc le sous-ensemble des points 3D de M_n qui ont un correspondant dans M_{n+1}^i (c'est à dire l'ensemble des points 3D correspondant aux sommets de M_n pour lesquels un correspondant existe dans Z_{n+1}^n).

On peut résumer la méthode que l'on va utiliser à l'aide des deux figures suivantes 1, 2 et 3.

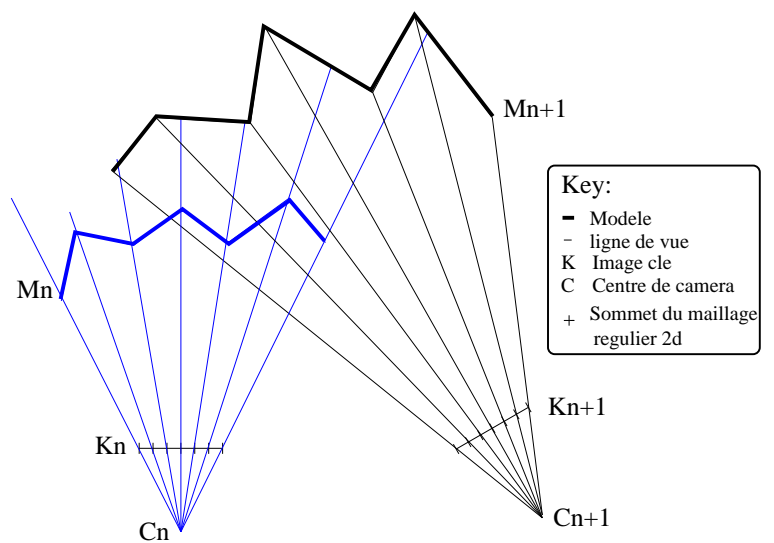


Figure 1 – Principe du morphing : les modèles d'entrée

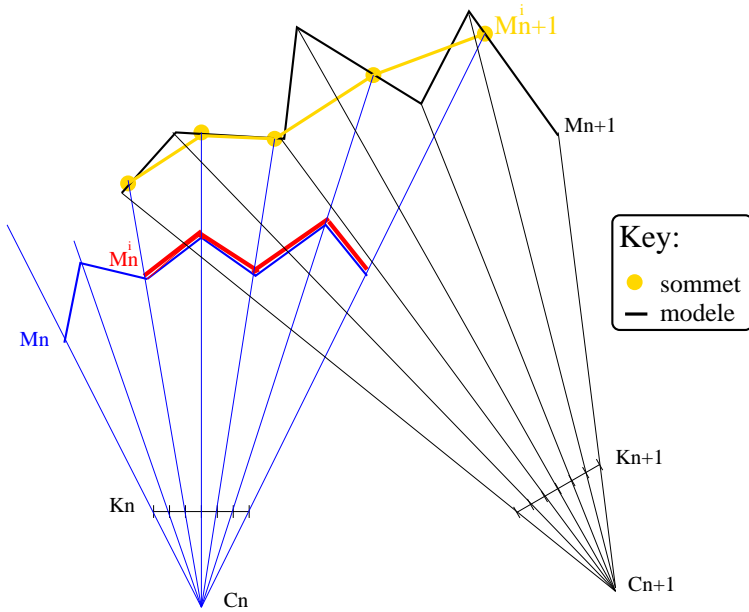


Figure 2 – Principe du morphing prédictif: M_n^i et M_{n+1}^i

L'interpolation linéaire entre M_n^i et M_{n+1}^i réalise un parcours des lignes de vue issues de C_n entre les sommets de ces deux modèles.

Il y a un remaillage à chaque fin de GOP qui correspond au passage du modèle M_{n+1}^i à M_{n+1} . Mais cette transition est douce car les deux modèles sont identiques du point de vue de C_{n+1} :

$$Pr(M_{n+1}^i, T_{n+1}, C_{n+1}) = Pr(M_{n+1}, T_{n+1}, C_{n+1})$$

Afin de palier au problème des points de M_{n+1} qui se projetteraient au même endroit dans K_n (car ils seraient situés sur la même ligne de vue issue de C_n), on ajoute un algorithme de Z-buffer au moment du remplissage de la carte de profondeur qui permet de ne garder que le point ayant la plus petite profondeur (on choisit de prendre la plus petite valeur car elle correspond au point le plus visible de C_n). Celui-ci compare pour chaque point sa profondeur à celle contenue dans la carte de profondeur Z_{n+1}^n à la position correspondant aux coordonnées 2D obtenues par projection. Z_{n+1}^n a été initialisée à l'infini et le Z-buffer permet de ne remplacer la profondeur dans la carte pour une position 2D donnée que si le point est plus près du centre de la caméra (c'est à dire celui qu'il a une profondeur plus petite) lorsqu'il y a un conflit. Le principe est décrit par la figure 4.

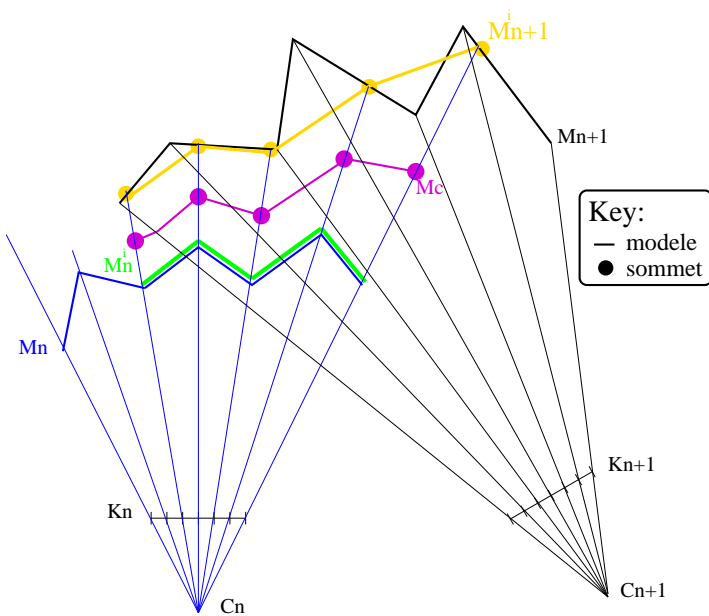


Figure 3 – Principe du morphing prédictif: le modèle couvrant M_c pour $\alpha = 0.5$

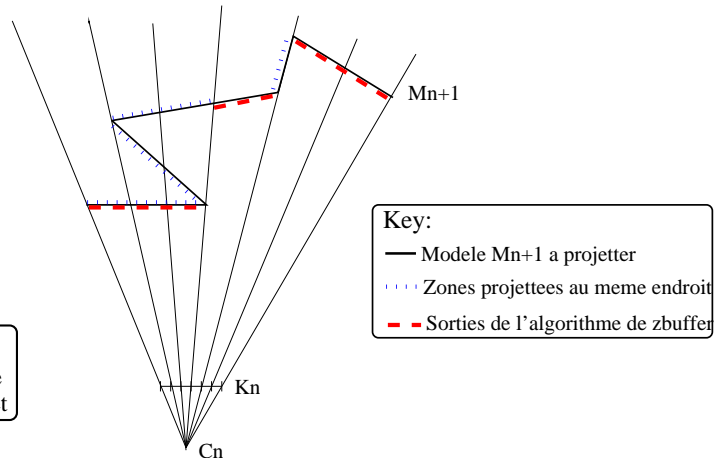


Figure 4 – Principe du Z-buffer

On a choisit de réaliser le morphing au moment du décodage plutôt que lors du codage car bien qu'il faille alors recalculer la carte de profondeur correspondant au niveau n, toutes les autres données (les modèles, les positions caméra, les matrices de projection et les images de texture) sont alors connues. L'avantage est donc que l'on a pas d'informations supplémentaires à transmettre ce qui nous permet de rester à très bas débit.

3 Résultats

C'est au niveau du rendu des images dans le cas de la navigation libre que l'on constate particulièrement l'apport de

notre morphing 3D comme le montrent les images de la figure 5:

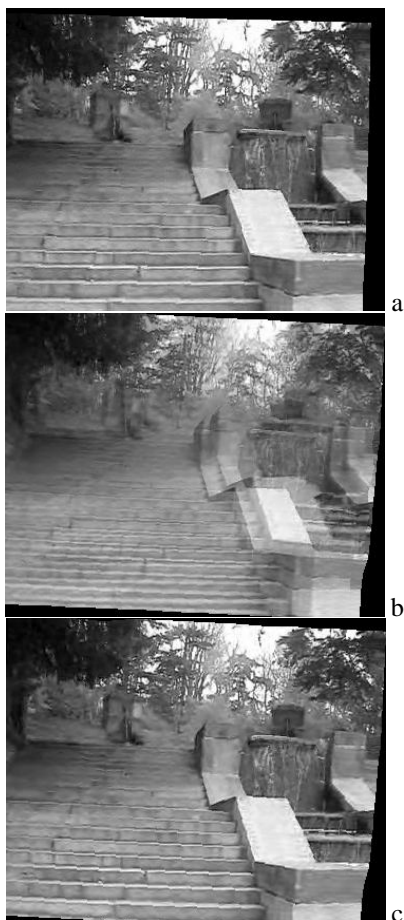


Figure 5 – Image 154 reconstruite correspondant à la position de caméra originale (a) et image 154 extraites d’une séquence de navigation libre avec le fondu 3D (b) et avec le morphing 3D (c).

Le fondu 3D introduit des artefacts visuels produits par la double projection. En effet comme M_{n+1} est valide entre K_{n+1} et K_{n+2} si on le projette sur des images situées entre K_n et K_{n+1} cela introduit un décalage. C’est ce que l’on peut observer sur l’image b de la figure 5 qui correspond à la moyenne de deux images issues de modèles différents et pas valides pour les mêmes parties de la séquence.

L’introduction du morphing 3D permet de supprimer ces artefacts visuels puisqu’on ne projette plus qu’un seul modèle comme on le voit bien sur l’image c de la figure 5.

Nos résultats sont comparables aux méthodes classiques de morphing 3D pour ce qui est du rendu visuel mais notre méthode a en plus le gros avantage d’être entièrement automatique.

4 Conclusion

Nous avons présenté une méthode efficace de morphing 3D de maillages dans le contexte de la représentation de

séquences vidéo et du codage très bas débit. Notre approche permet d’obtenir de manière entièrement automatique une séquence virtuelle dans laquelle on peut naviguer interactivement de manière libre.

L’ajout d’un tel morphing donne donc des résultats satisfaisants quant au rendu image dans le cadre de la navigation libre en particulier. Dans ce contexte il pourrait être intéressant d’essayer de gérer les zones qui se découvrent ou sont recouvertes d’un GOP à l’autre. En effet les points localisés dans ces zones ne sont pas du tout pris en compte au cours du parcours du GOP car ils sont éliminés par la présence d’un Z-buffer.

De plus une autre perspective est d’étendre cette méthode en y intégrant un calcul automatique de matrice de projection et de positions caméra associées pour pouvoir l’appliquer à n’importe quels modèles 3D maillés sans nécessiter de calculs préliminaires ou sans passer par la séquence vidéo de départ. De plus l’idée de remplacer la simple interpolation linéaire par un remaillage en 3D permettrait sans doute d’améliorer encore la qualité visuelle de la séquence reconstruite.

Références

- [Ale02] Marc Alexa. Recent advances in mesh morphing. In *Computer graphics forum*, pages 173–196, 2002.
- [Gal] Franck Galpin. *Représentation 3D de séquences vidéo: Schéma d’extraction automatique d’un flux de modèles 3D, applications à la compression et à la réalité virtuelle*. PhD thesis, Université de Rennes 1, janvier.
- [LDSS99] Aaron W.F Lee, David Dobkin, Wim Sweldens, and Peter Schröder. Multiresolution mesh morphing. In *Siggraph’99 Proceedings of the 26th Annual Conference on Computer Graphics, Los Angeles, USA*, pages 343–350, August 1999.
- [LV98] Francis Lazarus and Anne Verroust. Three-dimensional metamorphosis: a survey. *The Visual Computer*, pages 373–389, December 1998.
- [MKFC01] Takashi Michikawa, Takashi Kanai, Masahiro Fujita, and Hiroaki Chiyokura. Multiresolution interpolation meshes. In *Proc. 9th Pacific Graphics International Conference*, pages 60–69, October 2001.
- [ZSH00] Malte Zöckler, Detlev Stalling, and Hans-Christian Hege. Fast and intuitive generation of geometric shape transitions. *The Visual Computer*, pages 241–253, 2000.