

Détection temps-réel de changements de plan dans une séquence d'images couleur

Sébastien Lefèvre
lefevre@univ-tours.fr

Laboratoire d'Informatique, Université de Tours / E3i,
64, Avenue Jean Portalis - 37200 Tours - France

Atos Services
19, rue de la Vallée Maillard - BP 1311 - 41013 Blois Cedex - France

Mots-clés

Indexation vidéo, Segmentation temporelle, Temps-réel

1 Introduction

L'émergence de la vidéo comme média d'information s'accompagne d'un besoin croissant d'outils pour segmenter ces séquences d'images dans le domaine temporel. Cette segmentation temporelle a pour but de détecter les changements de plan, généralement caractérisés par des cuts ou par d'autres effets plus complexes. Elle est par exemple nécessaire dans des applications telles que la gestion de bases de données multimédia, la création automatique de résumés de films ou de séquences télévisées, la compression de séquences vidéo, ou encore le suivi d'objet en temps-réel dans des scènes dynamiques.

Nous sommes particulièrement intéressés par ce dernier type d'application, c'est-à-dire le suivi d'objet en temps-réel. Une détection préalable des changements de plan est nécessaire pour assurer une bonne qualité de suivi et une exécution en temps-réel.

Après avoir rappelé les principales solutions existantes, nous allons présenter ici notre méthode de détection de changement de plan en détaillant les principales étapes de l'algorithme.

2 Travaux antérieurs

La plupart des méthodes de détection de changements de plan peuvent se regrouper en plusieurs catégories : différence pixel à pixel, comparaison d'histogrammes, estimation de mouvement. Ces catégories vont être présentées brièvement (pour une revue plus complète, le lecteur pourra se reporter à [1],[2],[3]).

Les méthodes de différence pixel à pixel comparent les pixels de deux images successives. Si le nombre de pixels différents est supérieur à un seuil, alors un changement de plan est détecté. Ces méthodes sont assez sensibles au bruit et aux mouvements importants dans la scène.

Le deuxième type de méthodes utilise une comparaison d'histogrammes : si les histogrammes de deux images successives sont suffisamment différents, alors on affirme qu'un changement de plan est présent dans la séquence. L'inconvénient de ces méthodes est lié à la non-prise en compte de l'information spatiale contenue dans les images.

Enfin, les méthodes basées sur l'estimation de mouvement estiment le flux optique dans deux images successives. Une comparaison du nombre de mouvements différents entre les deux images à un seuil permet de déterminer la présence d'un changement de plan. Ces méthodes sont cependant coûteuses en temps de calcul.

3 Description de la méthode

La méthode présentée ici doit pouvoir s'exécuter en temps-réel. Pour cela une diminution de la résolution spatiale des images est effectuée, ce qui permet aussi de limiter la sensibilité au bruit. Cette réduction de la résolution spatiale est effectuée de deux façons, selon que la séquence source soit compressée ou non. En effet, dans le cas d'une séquence d'images compressée (comme MPEG ou MJPEG), l'utilisation des coefficients DC permet de reconstruire une image basse résolution ([4]). Dans le cas d'une séquence d'images non compressée, on divise l'image en blocs et on calcule la moyenne des intensités des pixels pour chaque bloc.

Une fois cette diminution de la résolution spatiale effectuée, il est nécessaire de diminuer le nombre de composants utilisés pour représenter chaque pixel. On modifie l'espace de représentation couleur utilisé (depuis RGB vers HSV), ce qui nous permet d'éliminer les effets de l'illumination en ne conservant que les composantes H et S (respectivement teinte et saturation).

Les changements de plan sont détectés en comparant les images basse résolution obtenus précédemment. La probabilité de la présence d'un changement de plan est proportionnelle à la différence entre deux images successives. Cette différence est calculée en utilisant la mesure de distance suivante :

$$d(t) = \sum_i |P_i(t) - P_i(t-1)|^n$$

avec $n = 0,1,2$ et la convention $0^0 = 0$ La contrainte d'exécution en temps-réel

implique la minimisation des données à traiter. Ici une différence entre deux images se résume à une simple valeur, ce qui permet des calculs rapides.

Afin de détecter les cuts et les effets plus complexes, on étudie l'évolution de la valeur d calculée précédemment ([5],[6]). Les cuts sont détectés en calculant la différence $D(t)$ entre $d(t)$ et $d(t-1)$ et en comparant cette valeur avec un seuil $T1$. Si $D(t)$ est supérieur à $T1$, alors un cut a été détecté. En effet, si $D(t)$ est important, alors la différence entre deux images successives n'est pas constante, contrairement à ce que l'on peut observer dans le cas de mouvements de caméra ou d'effets de transitions progressifs. Si un cut n'a pas été détecté, il est encore possible d'être en présence d'un effet progressif. Afin de détecter ce type d'effet, notre méthode s'inspire de [5]. On utilise un second seuil $T2$ et on le compare à la valeur $d(t)$ représentant la différence entre les images t et $t-1$. Si $d(t) > T2$, alors il est possible qu'un changement de plan soit présent. On calcule alors la valeur cumulée de $D(t)$ jusqu'à obtenir une valeur $d(t)$ inférieur à $T2$. On compare finalement la valeur cumulée obtenue pour D avec le seuil $T1$ fixé décrit précédemment. Un changement de plan est détecté si la valeur est supérieur au seuil.

4 Conclusion

La méthode proposée ici est capable de détecter, en temps réel, des changements de plan dans des séquences d'images couleur. Elle est relativement robuste au bruit, aux mouvements, ainsi qu'aux changements d'illumination. Des études sont menées actuellement afin de tester d'autres mesures de distance, d'autres espaces de représentation couleur, et une possibilité d'implémentation sur une station multiprocesseurs. Enfin, l'utilisation directe des blocs de différence est envisagée pour effectuer un suivi d'objets basse résolution.

Références

- [1] F. Idris and S. Panchanathan, "Review of Image and Video Indexing Techniques", Journal of Visual Communication and Image Representation, Vol. 8, No. 2, pp. 146-166, June 1997.
- [2] G. Ahanger and T.D.C. Little, "A Survey of Technologies for Parsing and Indexing Digital Video", Journal of Visual Communication and Image Representation, Vol. 7, No. 1, pp. 28-43, March 1996.
- [3] R. Brunelli, O. Mich, and C.M. Modena, "A Survey on the Automatic Indexing of Video Data", Journal of Visual Communication and Image Representation, Vol. 10, No. 2, pp. 78-112, June 1999.
- [4] J. Song and B.L. Yeo, "Spatially Reduced Image Extraction from MPEG-2 Video : Fast Algorithms and Applications", in Proceedings Storage and Retrieval for Image and Video Database VI, pp. 93-107, January 1998.
- [5] HJ Zhang, A. Kankanhalli, and S.W. Smolliar, "Automatic Partitioning of Full-Motion Video", Multimedia Systems, Vol. 1, No. 1, pp. 10-28, 1993.

- [6] CH Demarty and S. Beucher, "Morphological tools for indexing video documents", in Proceedings IEEE International Conference on Multimedia Computing and Systems - ICMCS'99, pp. 991-992, June 1999.