

Influence de la Réduction des Couleurs sur la Détection des Changements de Plan dans les Films d'Animation

Bogdan IONESCU⁽¹⁾, Patrick LAMBERT⁽²⁾, Didier COQUIN⁽²⁾, Vasile BUZULOIU⁽¹⁾

⁽²⁾ LISTIC - Université de Savoie - BP 806 -74016 Annecy Cedex

⁽¹⁾ LAPI - Université POLITEHNICA de Bucarest

Bogdan-Emanuel.Ionescu@univ-savoie.fr, Patrick.Lambert@univ-savoie.fr

Résumé - Dans ce papier nous proposons une technique de détection des changements de plan appliquée aux films d'animation. Cette technique est basée sur la mesure de distance entre histogrammes couleur d'images voisines. L'approche envisagée demande la définition d'une palette couleur réduite sur laquelle sont calculés les histogrammes. La détection des changements est alors effectuée par seuillage, le seuil étant déterminé automatiquement. Nous présentons des résultats obtenus avec quelques films sur lesquels une segmentation temporelle a été effectuée manuellement pour disposer d'une vérité terrain. Ces tests permettent d'évaluer l'influence du choix de la palette couleur choisie.

Abstract - In this paper an improved cut detection algorithm is proposed. It is adapted to the segmentation of animation movies. As the color is an important feature in the animation movies (each movie has its own particular color distribution), metrics on color histograms are used in order to measure the visual discontinuity produced by video cuts. For the frame classification an automatic threshold estimation is proposed. Building the frame histograms requires reducing the number of colors. Based on experimental results a study of the influence of different color reduction algorithms on the cut detection error is presented.

1. Introduction

Le travail exposé ici a été développé en lien avec le CICA (Centre International du Cinéma d'Animation [1]). Le CICA organise chaque année, depuis 1960, un Festival International du Cinéma d'Animation à Annecy. Un très grand nombre de films ont ainsi pu être archivés et la difficulté est de pouvoir exploiter cette base. L'utilisation d'un système d'indexation est de toute évidence nécessaire. Les travaux présentés s'inscrivent dans un projet plus global dont l'objectif est de construire les méthodes et les outils permettant la caractérisation des films, la recherche ou la navigation dans la base.

Les films d'animation présentent un certain nombre de spécificités qui obligent à envisager des approches en partie différentes de celles utilisées dans l'analyse de vidéos plus classiques (informations, sports, etc. [2]) : les événements ne suivent pas forcément une logique particulière (des objets peuvent apparaître et disparaître sans respecter de règle physique), les déplacements peuvent être discontinus et les mouvements de caméra très complexes, des effets spéciaux peuvent apparaître à tout moment, les couleurs ne sont pas nécessairement des couleurs naturelles, etc. .

Dans l'analyse de telles séquences, le premier travail consiste à effectuer une segmentation temporelle en détectant les changements de plans. Les méthodes de détection des changements (ou «cuts») peuvent être classées en trois catégories selon les attributs qu'elles utilisent : la couleur, les contours

ou le mouvement. Une comparaison de ces trois types d'approche est proposée dans [3] et [4]. Les méthodes basées sur l'utilisation des histogrammes couleur sont largement utilisées : elles sont en effet peu sensibles aux transformations géométriques et fournissent en général de meilleurs résultats que les autres approches [5]. Différentes techniques à base d'histogrammes couleur ont été proposées dans la littérature : mesure de l'intersection d'histogrammes dans l'espace Cb-Cr et r-b [6], différence d'histogrammes dans l'espace Yuv [7], comparaison d'histogrammes calculés sur des sous-fenêtres [8], différence d'histogrammes utilisant une distance de Hausdorff multi-niveau [9].

Le travail présenté ici utilise également les histogrammes couleur pour détecter les discontinuités visuelles provoquées par les «cuts». La démarche envisagée consiste à réduire le nombre de couleurs, puis à calculer 4 histogrammes correspondant aux 4 quadrants de chaque image. Les différentes distances entre histogrammes d'images successives sont automatiquement seuillées, et permettent alors de décider de la présence d'un «cut». Le but de ce travail est d'analyser l'incidence de l'étape de réduction des couleurs sur les résultats de la segmentation temporelle.

Le papier est organisé de la manière suivante. Dans une première partie nous présentons les différentes réductions couleur utilisées. La deuxième partie détaille la technique de détection des changements de plan. Enfin, dans une dernière partie, nous donnons des résultats sur des séquences réelles.

2. La réduction couleur

Classiquement représentée sur 3x8 bits, la couleur peut donc prendre plus de 16 millions de valeurs. Dans la plupart des applications, la réduction des couleurs est une étape préalable indispensable. Les techniques de réduction couleur sont très nombreuses [10], [11], [12]. Elles sont basées sur le fait que l'œil humain ne perçoit pas les petites variations de couleur. On peut ainsi modifier la couleur de certains pixels sans modification majeure de la perception visuelle. Le choix dépend souvent de l'application et résulte d'un compromis entre la qualité de la préservation des couleurs et le temps de calcul. En général, la réduction couleur se déroule en 2 étapes : la construction d'une palette, puis l'attribution d'une des couleurs de la palette pour chaque couleur de l'image.

Ici, notre soucis n'est pas de préserver la qualité visuelle de l'image, mais de conserver les différences couleur pouvant exister entre images. Parmi les différentes approches possibles, on distingue deux types de stratégies :

- D'une part les stratégies pour lesquelles la réduction des couleurs est déterminée en fonction des couleurs de chaque image.

- D'autre part les stratégies basées sur l'utilisation d'une palette de couleurs déterminée a priori, indépendamment des couleurs des images à traiter. Ces techniques sont plus simples à mettre en œuvre et permettent une comparaison plus rapide des histogrammes calculés sur ce type de palette. Nous nous sommes limités à l'utilisation de ces techniques en envisageant quatre approches :

- une quantification uniforme dans l'espace RVB. En gardant 5 couleurs sur chaque composante, nous disposons de 125 couleurs. C'est la stratégie la plus simple, mais les différences perceptuelles ne sont pas toujours bien restituées avec cette palette.

- une quantification de l'espace Teinte, Luminance, Saturation (TLS). L'espace TLS est basée sur la perception humaine de la couleur. Les axes des Teintes et des Luminances sont divisés en 7 intervalles et celui de la Saturation en 4 intervalles, l'œil humain étant plus sensible à la Teinte et à la Luminance. 9 niveaux de gris, incluant le noir et le blanc, sont ajoutés, ce qui aboutit à 205 couleurs. Nous nous sommes limités à un nombre faible de couleurs pour garder une cohérence avec les autres palettes, même s'il est préconisé [2] d'utiliser une palette de plus de 561 couleurs pour préserver la Teinte.

- une palette construite dans l'espace Lab proposée dans [10]. Cette palette comporte 215 couleurs. Chaque couleur de l'image est remplacée par la couleur de la palette la plus proche mesurée selon la distance euclidienne calculée dans l'espace Lab. L'espace Lab présente une meilleure uniformité de la perception visuelle que les autres espaces : ainsi, les distances euclidiennes y sont en concordance avec les écarts perceptuels.

- la diffusion d'erreur dans l'espace XYZ. Après réduction

des couleurs par la palette Lab présentée ci-dessus, les couleurs sont modifiées en tenant compte de leur voisinage spatial avec l'algorithme de Floyd and Stenberg dans l'espace XYZ [13]. La diffusion permet d'atteindre une meilleure qualité visuelle au prix d'une plus grande complexité.

D'une manière générale, le choix de la palette dépend d'un compromis entre le nombre de couleurs conservées, la qualité visuelle que l'on désire obtenir et le coût en termes de temps de calcul. Les résultats obtenus avec les quatre réductions envisagées sont présentés en figure 1.

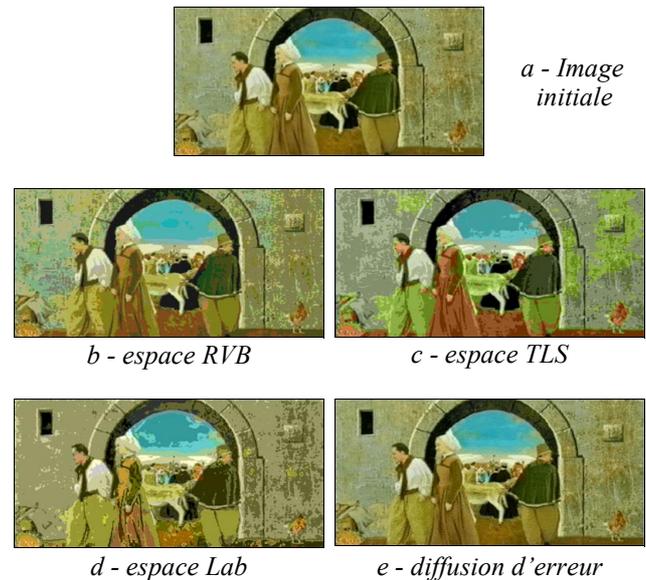


FIG 1 : Réduction couleur

Du point de vue du temps de calcul, les deux dernières méthodes sont beaucoup plus coûteuses puisqu'elles demandent, pour chaque couleur de l'image, la recherche de la couleur de la palette la plus proche.

L'ensemble de ces résultats est résumé dans le tableau ci-dessous où la qualité a été évaluée de manière subjective par une note allant de 1 (mauvais) à 4 (bon). Les temps de calcul sont mesurés sur un processeur PentiumM 1.6 GHz pour une image 178x81.

TAB 1 : Performances de la réduction couleur

Réduction	RVB	TLS	Lab	Diff. err.
Qualité	2	1	3	4
Tps calcul	0.007 s	0.017 s	4.1 s	4.2 s

3. La segmentation temporelle

La technique utilisée ici repose sur la comparaison d'histogrammes couleur. Elle présente l'avantage d'être rapide et de ne pas nécessiter l'estimation du mouvement contenu dans l'image. Elle est adaptée à la détection de changements rapi-

des. Les changements avec effets spéciaux (fade, dissolves, etc.) demandent un traitement particulier [14], [15]. Pour accélérer le traitement, les images sont sous-échantillonnées spatialement (pour se ramener à des images de taille approximative 160x90) et temporellement (prélèvement régulier d'une image sur n , n paramètre de la méthode). Ce sous-échantillonnage temporel ne peut être trop important car on trouve dans les films d'animation des plans de courte durée (moins de 5 images).

La détection de «cut» à l'instant courant k se fait en considérant les instants précédent et suivant, l'observation sur un horizon de 3 images limitant le risque de fausses détection. Nous noterons $I(k-n)$, $I(k)$ et $I(k+n)$ les images utilisées à l'instant k . Pour limiter l'influence du mouvement qui pourrait engendrer de fausses détections, les images sont découpées en quatre quadrants ($i = 1 \dots 4$). Un histogramme couleur (sur la palette réduite) est calculé sur chacun des quadrants, et l'on mesure les différences entre les histogrammes des quadrants respectifs des paires d'images consécutives. Le schéma de la figure 2 résume cette stratégie.

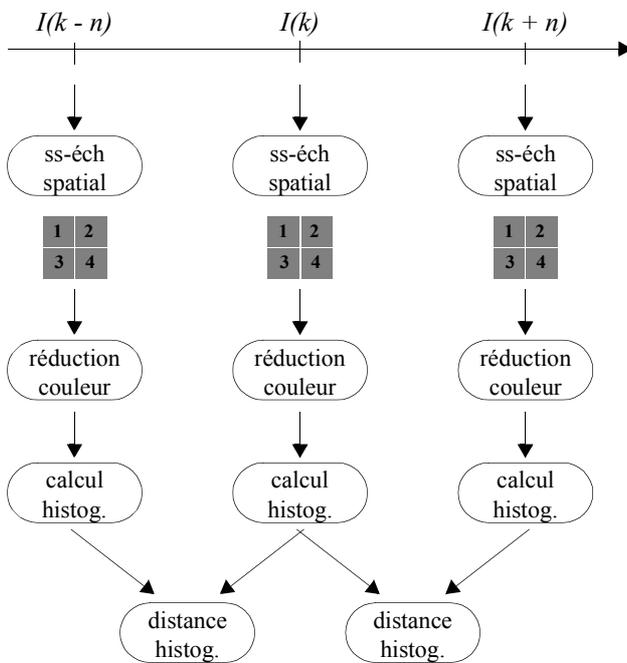


FIG 2 : Principe de la détection des «cuts»

Pour l'étude de l'instant k , on dispose donc des 8 distances suivantes :

$$D_i(k) = \text{dist}[\text{hist}_i(k-n), \text{hist}_i(k)]$$

$$D_i(k+n) = \text{dist}[\text{hist}_i(k), \text{hist}_i(k+n)]$$

avec $i \in [1, 4]$

La distance utilisée est la distance Euclidienne «bin» à «bin». Un changement de plan est détecté si la condition :

$$D_i(k) > S \quad \text{ET} \quad D_i(k+n) < S$$

est vérifiée sur au moins 3 des 4 quadrants, c'est-à-dire au

moins pour 3 des 4 valeurs de i .

Une procédure d'estimation automatique du seuil S a été validée. Elle s'appuie sur une analyse de la moyenne D_{moy} des distance D_i , $i = 1 \dots 4$, et fonctionne en deux temps. Une première étape élimine toutes les distances inférieures à la moyenne des distances. Le seuil est alors déterminé comme la moyenne des distances supérieures à cette première moyenne. Un exemple est présenté dans la figure 3.

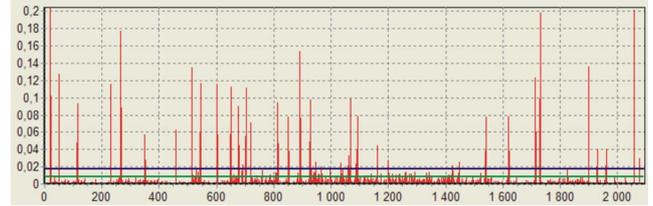


FIG 3 : Estimation automatique du seuil de détection : L'axe des abscisses correspond au temps, l'axe des ordonnées aux distances D_{moy} . Le seuil S est représenté par la ligne horizontale bleue. La ligne verte indique le premier seuil obtenu en effectuant la moyenne de toutes les distances. Les pics isolés correspondent à des «cuts»

La méthode ne présente donc pas de paramètre de réglage, à l'exception des valeurs de sous-échantillonnage spatial et temporel. Ces deux paramètres n'ont qu'une incidence sur la rapidité des calculs mais n'influent pas beaucoup sur la qualité de la détection si leurs valeurs restent dans des plages raisonnables.

4. Résultats expérimentaux comparatifs

La technique de détection des changements de plan a été testée sur deux films d'animation de longue durée (film1 : 84 mn 46s et 1597 changements; film2 : 73mn18s et 1569 changements) en utilisant les 4 techniques de réduction couleur et un sous-échantillonnage temporel n égal à 2. Les résultats obtenus sont présentés en figure 4. Après une étape de segmentation temporelle manuelle, nous avons pu calculer les taux de non détections et de fausses alarmes présentés dans la figure 4.

On constate que la réduction couleur utilisée joue un rôle important sur le taux de non-détections : variation allant de 13.30 % à 7.96 %. C'est la palette TLS qui fournit les meilleurs résultats, ce qui peut s'expliquer par le fait que cette palette a tendance à préserver, voire accentuer, les différences couleur. La conséquence est que c'est cette même palette qui donne le taux de fausses détections le plus élevé. Cependant le taux de fausses détections semble moins sensible à la méthode de réduction couleur employée. Les taux restent compris entre 5.34 % et 6.85 %, la diffusion d'erreur donnant le moins de fausses détections. Il faut noter que la plupart des fausses alarmes sont dues à des mouvements de caméra, et que les non détections proviennent de changements de plan sans modification importante des couleurs.

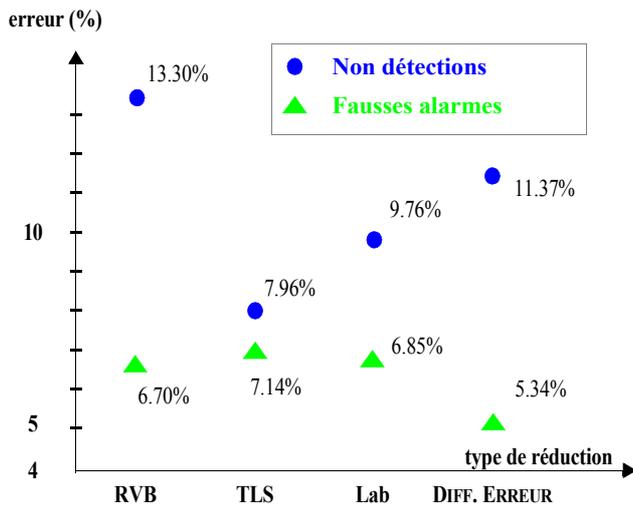


FIG 4 : Performances de la méthode de détection des changements de plan en fonction de la palette couleur utilisée

5. Conclusion

Dans ce papier nous avons proposé une méthode automatique de détection des changements de plans («cut») dans les films d'animation. La discontinuité visuelle provoquée par un «cut» est évaluée à travers la mesure de la distance Euclidienne entre les histogrammes couleur d'images successives. Cette distance est comparée à un seuil calculé automatiquement.

Pour limiter le temps de calcul nécessaire à l'analyse d'une séquence, un certain nombre de réductions ont été effectuées : sous-échantillonnage spatial, réduction de la taille des images et réduction des couleurs.

Une analyse de l'influence du choix de la technique de réduction couleur sur la qualité de la détection des «cuts» a été réalisée. Les erreurs de fausses alarmes restent faibles, entre 5 et 7%, et ne semblent pas dépendre beaucoup de la palette utilisée. En revanche, les non-détections ont des valeurs plus importantes et le choix de la palette couleur semble être un facteur important. On constate également que la plus faible erreur de détection est obtenue avec la palette TLS dont la qualité visuelle reste moyenne du fait de sa tendance à accentuer les différences visuelles entre les couleurs. La palette obtenue par diffusion d'erreur sur l'espace XYZ, même si elle assure une bonne qualité visuelle, aboutit à un taux de non-détection relativement important.

Les perspectives de ce travail ont pour objectif la prise en compte des informations de mouvement de manière à compenser les variations dues aux déplacements caméra / objets.

Remerciements : les auteurs tiennent à remercier le CICA (Centre International du Cinéma d'Animation) pour la mise à disposition des films d'animation.

Références

- [1] Centre International du Cinéma d'Animation. <http://www.annecy.org>.
- [2] A.D. Bimbo. *Visual information retrieval*. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, pp. 81-113, 1999.
- [3] R.Lienhart. *Reliable Transition Detection in Videos: A Survey and Practitioner's Guide*. International. Journal of Image and Graphics, 1(3), pp. 469-486, 2001.
- [4] R. Lienhart. *Comparison of Automatic Shot Boundary Detection Algorithms*. SPIE Storage and Retrieval for Still Image and Video Databases VII, 3656, pp. 290-301, January 1999.
- [5] U. Gargi, R. Kasturi, S.H. Strayer. *Performance Characterization and Comparison of Video-Shot-Change Detection Methods*. IEEE Transaction on Circuits and Systems for Video Technology, 10(1), February 2000.
- [6] M.S. Drew, Z.N. Li, X. Zhong. *Video dissolve and wipe detection via spatio-temporal images of chromatic histogram differences*. Proc. of IEEE International Conference on Image Processing, 3, pp. 929-932, 2000.
- [7] S. Nayaga, S. Seki, R. Oka. *Temporal video segmentation using unsupervised clustering and semantic object tracking*. SPIE Journal of Electronic Imaging, 7(3), pp. 592-604, July 1998.
- [8] A. Nagasaka, Y. Tanaka. *Automatic video indexing and full-video search for object appearances*. Proc. of IFIP TC2/G2.6 Second Working Conference on Visual Database Systems, pp. 113-127, 1991.
- [9] B. Shen. *HDH based compressed video cut detection*. Proc. of Visual 97 San Diego CA, pp. 149-156, December 1997.
- [10] *Worldnet User's Reference Desk*. <http://www.wurd.com>, 2004.
- [11] P. Bonton, C. Fernandez-Maloigne, A. Trémeau, *IMAGE NUMERIQUE COULEUR De l'acquisition au traitement*, DUNOD, ISBN 2 10 006843 1, 2004.
- [12] K. Kanjanawanishkul, B. Uyyanonvara. *Novel fast color reduction algorithm for time-constrained applications*. Journal of Visual Communication and Image Representation, In Press, Corrected Proof, Available online 12 October 2004,.
- [13] B. L. Evans, V. Monga, N. Damera-Venkata, *Variations on error diffusion: retrospectives and future trends*, Proc. SPIE Vol. 5008, p. 371-389, Color Imaging VIII: Processing, Hardcopy, and Applications; Reiner Eschbach, Gabriel G. Marcu; Eds, jan 2003.
- [14] M.S. Drew, Z.N. Li, X. Zhong. *Video dissolve and wipe detection via spatio-temporal images of chromatic histogram differences*. Proc. of IEEE International Conference on Image Processing, 3, pp. 929-932, 2000.
- [15] B. Ionescu, D. Coquin, P. Lambert, V. Buzuloiu, "Analyse et Caractérisation de séquences de films d'animation", 9^{ème} journées ORASIS 2005.