

# Extraction d'objets vidéo : Une approche combinant les contours actifs et le flot optique

Youssef Zinbi\*, Youssef Chahir\* et Abder Elmoatz \*\*

\* GREYC - URA CNRS 6072  
Campus II - BP 5186  
Université de Caen  
14032 Caen Cedex

\*\* LUSAC , Site Universitaire, rue Luis Aragon  
50130 Cherbourg  
{ yzinbi,chahir }@info.unicaen.fr  
abder@greyc.ismra.fr

**Résumé.** Dans cet article, nous présentons une méthode mixte de segmentation d'objets visuels dans une séquence d'images d'une vidéo combinant à la fois une segmentation basée régions et l'estimation de mouvement par flot optique. L'approche développée est basé sur une minimisation d'une fonctionnelle d'énergie ( $E$ ) qui fait intervenir les probabilités d'appartenance (densité) avec une gaussienne, en tenant compte des informations perceptuelles de couleur et de texture des régions d'intérêt. Pour améliorer la méthode de détection et de suivi, nous avons étendu la formulation énergétique de notre modèle de contour actif en incluant une force supplémentaire issue du calcul du flot optique. Nous montrons l'intérêt de cette approche mixte en terme de temps de calcul et d'extraction d'objets vidéo complexes, et nous présentons les résultats obtenus sur des séquences de corpus vidéo couleur.

## 1 Introduction

La recherche d'objets vidéo est une tâche difficile compte tenu de la richesse des informations multiples dans l'image. Pour trouver de manière automatique ces objets vidéo, il est important de tenir compte de trois étapes principales qui sont la segmentation, l'identification et le suivi d'objets en mouvement par flot optique.

Le but de la segmentation active est de détecter et d'extraire des informations pertinentes dans une image. Différents modèles de contours actifs ont été proposés dans la littérature, mais on peut distinguer deux principales approches: Des approches basées contours et d'autres basées régions. L'implémentation de n'importe quel modèle de contour actif exige la minimisation d'une fonctionnelle d'énergie. Cette énergie a deux composantes: énergie externe, qui est caractérisée par la régularité de la courbe et l'énergie interne qui a pour fonction d'attirer la courbe vers les gradients les plus forts (les forts contraste de l'image).

Les contours actifs classiques ont été proposés pour la première fois par Kass et al (Kass et al., 1987) pour la segmentation d'images médicales. L'idée de base consiste à faire évoluer la courbe vers la frontière de l'objet à détecter. Ce modèle a été confronté à plusieurs

contraintes, liées à l'initialisation, au paramétrage et à l'impossibilité de changement de topologie du snake. Une autre méthode a été introduite par Osher & Sethian (Sethian, 1999) connue par la méthode des ensembles de niveaux. Son principe consiste à faire évoluer une courbe initiale jusqu'à ce qu'elle détecte la forme de l'objet à extraire. Ensuite, les contours actifs géodésiques (Caselles et al., 1997) ont été présentés comme une alternative géométrique aux snakes, présentant l'avantage d'être indépendant du paramétrage. Les contours actifs basés régions adoptés par Barlaud et Jehan-Besson (Jehan-Besson et al., 2002). Ils utilisent des descripteurs statistiques des régions, de manière générale, on peut dire que cette méthode est efficace, quand l'ensemble des objets à segmenter est homogène.

Le flot optique se calcule entre deux images: c'est le champs de vecteurs mouvement, rapportés aux pixels, pour passer d'une image à l'autre. Pour faire ce calcul plusieurs approches ont été proposés (Horn et al., 1981). Il existe aussi une autre classe de méthodes pour estimer le mouvement, telles que celles utilisées en compression: le bloc-matching (Koga et al., 1981).

Dans la section suivante nous présentons notre approche de segmentation basée région qui est une implémentation rapide d'un modèle de contour actif qui permet de tenir compte d'informations de couleur et de texture. Ensuite, nous abordons le problème général du flot optique et présentons notre implémentation de l'estimation du mouvement par une approche basée sur la méthode d'Horn & Schunck. Dans la section 4, nous proposons une méthode de mise en évidence (à partir de la détection et de suivi d'objets visuels) en utilisant une approche mixte qui combine à la fois les contours actifs et le flot optique. La section 5 sera consacrée aux résultats obtenus à partir des différentes approches présentées, pour la segmentation des objets en mouvement sur une série d'images vidéo. Enfin, nous terminons notre article par une conclusion en indiquant quelques améliorations possibles.

## 1.1 Segmentation par contour actif

Nous avons amélioré la méthode de Chan & Vese (Zinbi et al., 2004) en utilisant une fonction générale de la gaussienne qui permet de mieux tenir compte des caractéristiques divers de texture et de couleur dans l'image. Pour les images en couleur, on a choisi de travailler dans un espace de couleur perceptuel (tels que Lab), ainsi  $\vec{I}_j$  est le vecteur couleur du pixel  $j$ .  $\vec{c}_i$  le vecteur moyenne des composantes couleur de la région  $\Omega$  et  $\Sigma$  la matrice de covariance. Ces deux derniers caractérisent le comportement des deux régions de couleur et la probabilité d'appartenance d'un pixel à une région donnée est présentée dans la fonction ci-dessous:

$$P_{\Omega_i}(j / (\mu_i, \sigma_i)) = \frac{1}{\sqrt{(2\pi)^3 \left| \sum_j \right|_i}} e^{-\frac{(\vec{I}(i,j) - \vec{c}_i)^T \sum_i^{-1} (\vec{I}(i,j) - \vec{c}_i)}{2\sigma_i^2}} \quad (1)$$

Alors la fonction d'énergie est définie comme suite:

$$E_{CA}(\Omega_{in}, \Omega_{out}, P_{\Omega_{in}}, P_{\Omega_{out}}) = \mu \cdot (\text{long}(C))^p + \int_{\text{int}\Omega} \sum_1^N \log(P_{\Omega_{in}}(j / \mu_{in}, \sigma_{in})) dx - \int_{\text{ext}\Omega} \sum_1^N \log(P_{\Omega_{out}}(j / \mu_{out}, \sigma_{out})) dx \quad (2)$$

C est la courbe (ou un ensemble de courbes) qui doit évoluer dans le temps en fonction des régions in/out. Un pixel change d'état en fonction de sa position (intérieur/extérieur) et de l'énergie calculée  $E_{CA}$  (positive ou négative).

## 2 Estimation de mouvement par flot optique: Modèle de Horn & Schunck

Les méthodes de détermination du flot optique font partie des principales contributions qui ont été présentées pour extraire une information dense et précise du mouvement, sans nécessairement se fier à une connaissance a priori. Horn & Schunck (Horn et al., 1981) ont partis de l'ECMA (contrainte d'intensité) et ont introduit une contrainte supplémentaire dite contrainte de lissage quadratique. Ainsi, les combinaisons de ces deux contraintes s'expriment par la fonction, donnée ci-dessous :

$$E_{FO} = \iint [(uIx + vIy + It)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2)] dx dy \quad (3)$$

avec  $\alpha = \frac{I_x \cdot u_{avg} + I_y \cdot v_{avg} + I_t}{I_x^2 + I_y^2}$

Dans notre implémentation, au lieu de choisir le nombre d'itérations, nous avons décidé d'arrêter le programme en fonction du taux de convergence par rapport à un seuil. Ce taux se calcule à partir de l'erreur moyenne entre deux itérations:

$$e^{k+1} = \sum_{image} \frac{(u^{k+1} - u^k)^2 + (v^{k+1} - v^k)^2}{np} \quad (4)$$

Avec np le nombre de pixels de l'image. Le taux de convergence est donc:  $\rho = \frac{e^{k+1}}{e^k}$ .

## 3 Détection et suivi d'objets visuels par une approche mixte «CAFO»

L'idée principale consiste à utiliser une segmentation active des régions d'intérêt avec un critère qui est fonction du mouvement. L'algorithme de base est divisé en deux parties: l'estimation du mouvement et la segmentation par contour actif. Pour simplifier la tâche et gagner en rapidité et en efficacité, nous allons minimiser une fonctionnelle d'énergie unique pour la segmentation et l'estimation:

$$F = \alpha \cdot E_{CA} (\Omega_{in}, \Omega_{out}, P_{\Omega_{in}}, P_{\Omega_{out}}) + (1 - \alpha) \cdot E_{FO} \quad (5)$$

Il s'agit d'une résolution simultanée du problème d'estimation de mouvement et de segmentation active d'une image vidéo, par minimisation de l'énergie F. La première étape consiste en l'initialisation du contour sur l'image courante. Elle peut être effectuée de manière automatique en utilisant le résultat d'une étape de séparation du fond et des objets. Une fois le contour initialisé, un processus de déformation intervient jusqu'à convergence en

utilisant les forces décrites précédemment. Il va nous permettre de déterminer la position d'un objet(t) à l'instant t en se basant sur sa position précédente objet(t-1).

L'intérêt de notre méthode consiste à extraire des objets visuels en mouvement de mêmes contraste que le fond de l'image. L'approche CAFO combine les avantages des deux méthodes «Contour Actif» (CA) et «Flot optique» (FO). En effet, la segmentation active (CA) donne des résultats satisfaisants sur des images complexes (objet + décors). Alors que l'estimation du mouvement par flot optique donne d'assez bons résultats quand les régions d'intérêt sont texturées. L'approche CAFO améliore le processus de segmentation et résout les problèmes d'occlusion et d'ouverture connus quand on estime le mouvement des pixels entre deux images.

D'un autre côté, cette approche est particulièrement adaptée pour le suivi d'objets. La méthode peut être composé de deux étapes qui sont l'initialisation et la déformation successivement sur chaque image de la séquence vidéo. Tout d'abord le contour est initialisé en utilisant le résultat obtenu à l'image précédente. Il est ensuite déformé en utilisant à la fois des énergies issue du modèle de contours actifs et de la force issue du calcul du flot optique. Le contour final de l'image précédente sera utilisé comme contour initial sur l'image courante. Une fois le contour initialisé, un processus de déformation intervient jusqu'à convergence en utilisant les forces décrites précédemment. Cette méthode va nous permettre de déterminer la position d'un objet(t) à l'instant t en se basant sur sa position précédente objet(t-1).

## 4 Expérimentation:

Nous avons testé nos algorithmes sur un corpus connu d'images vidéo ( cf. figure 1.a). Dans la figure 1, nous montrons les résultats de nos méthodes de segmentation par contour actif, d'estimation du mouvement par flot optique et enfin notre approche mixte appelée «CAFO». L'intérêt de ces images vidéo réside dans le fait qu'elles contiennent des objets visuels de premier plan avec des fonds complexes mais stables. Dans cette série d'images, nous avons segmenté les régions en minimisant la fonctionnelle via la propagation du contour actif en tenant compte des descripteurs dépendant des régions. Les résultats de segmentation par contour actif, toujours, à partir d'une courbe initialisée automatiquement (rectangle). On remarque que la méthode détecte tous les objets contrastés ou texturés. La figure 1.b montre le résultat de segmentation sur des images clés de vidéos tests. Un des principaux avantages de cette description implicite du contour est qu'elle gère intrinsèquement les changements topologiques en cours de convergence. Par exemple, si un processus de segmentation est initialisé par des germes multiples, les collisions et fusions des composantes connexes sont gérées sans modification de l'algorithme (Zinbi et al., 2004).

Cependant, les objets d'intérêt ne sont pas toujours très visibles et très distinguables du fond. Toutes les méthodes de segmentation trouvent leurs limites quand il y a une occlusion ou quand l'objet possède les mêmes caractéristiques statistiques (moyenne, ...etc.) que son voisinage ou le fond. D'où la phase d'estimation de mouvement par flot optique qui permet de séparer les objets qui bougent d'un fond statique. Pour une bonne visibilité, nous avons opté pour l'affichage d'un masque binaire paramétrable en fonction de la l'amplitude des vecteurs de déplacements (la longueur estimée des déplacements) qui montre les pixels qui ont bougé. La figure 1.c montre le résultat d'estimation de mouvement sur les séquences d'Akiyo, d'Erik et du Hall (entre deux images consécutives). On remarque d'ores et déjà, que seules les pixels qui ont changé de couleur, sont gardés.

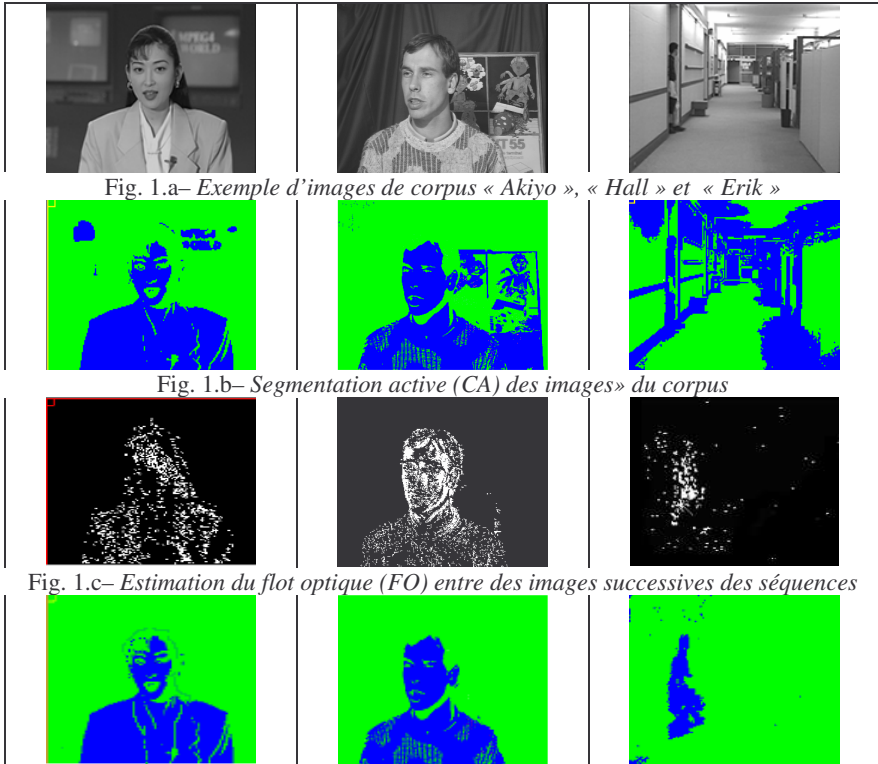


FIG. 1 – CAFO : Segmentation des objets visuels par combinaison d'une approche basée Contour Actif et de méthode de Flot Optique

En général, les méthodes différentielles échouent avec ces images. L'estimation de mouvement (ou l'appariement de blocs) entre deux images successives, donne des informations supplémentaires mais insuffisantes pour l'extraction du contenu visuel des séquences. La figure 1.d montre le résultat de l'approche mixte qui intègre à la fonctionnelle d'énergie des régions les forces issues du flot optique. Cela a permis de mettre en évidence plus nettement des régions les forces issues du flot optique. Cela a permis de mettre en évidence plus nettement des régions les forces issues du flot optique. Cela a permis de mettre en évidence plus nettement des régions les forces issues du flot optique. Cela a permis de mettre en évidence plus nettement des régions les forces issues du flot optique. L'avantage de combiner les informations du mouvement avec la segmentation est multiple. D'abord, cela peut contribuer à l'indépendance de l'approche de segmentation par contour actif, de la phase d'initialisation du contour initial (par exemple, en utilisant le FO comme initialisation de l'image suivante). Ensuite, elle permet de compenser les informations manquantes dans une image, ou non détectables (détails) à cause la moyenne. Enfin, elle permet de privilégier l'une des deux approches (grâce aux poids) en fonction des applications. Toutes les expériences ont été faites en utilisant l'approche de segmentation active, et le calcul du flot optique développés en langage C++ dans l'environnement Pandore.

## 5 Conclusion et perspectives

Dans cet article, nous avons présenté une méthode de segmentation active qui est basée sur la formulation énergétique de notre modèle (Zinbi et al., 2004), en incluant une force supplémentaire issue du calcul du flot optique. Nous avons montré l'intérêt de cette approche dans l'extraction d'objets vidéo complexes, et nous avons présenté les tests et résultats obtenus sur des vidéos. Dans le cadre de nos travaux futurs, nous prolongerons notre modèle pour prendre en considération un autre descripteur la forme (Gastaud et al., 2004).

## Références

- Caselles, V., V. Caselles, R. Kimmel, G. Sapiro (1997) .Geodesic active contours. *International Journal of Computer Vision*, vol. 22(1), pp. 61-79.
- Koga, T., K. Linuma, A. Hirano, Y. Iijima, T. Ishiguro, (1981) Motion-compensated interframe coding for video conferencing, *Proc. NTC*, New Orleans, LA, pp. G5.3.1--5.
- Horn, B.K.P., B.G. Schunck, (1981). Determining Optical Flow. *Artificial Intelligence*, vol. 2. pp. 185-203
- Gastaud, M., M. Barlaud and G. Aubert, (2004). Combining shape prior and statistical features for active contour segmentation, *IEEE TCSVT special session on Audio and Video Analysis for Interactive Multimedia Services*. 14(5): 726-734
- Jehan-Besson S., M. Barlaud, G. Aubert, (2002). A 3-step algorithm using region-based active contours for video objects detection. *JASP*, 2002(6): 572-581.
- Kass, M., M. Kass, A. Witkin, D. Terzopoulos, (1987). Snakes: Active Contour Models” *First International Conference On Computer Vision*: 259-268.
- Sethian, J.A. (1999). Level Set Methods and Fast Marching Methods. Evolving Interfaces in Computational Geometry, Fluid Mechanics, *Computer Vision, and Materials Science* .*Cambridge Monograph on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, UK.
- Zinbi, Y., Y. Chahir and A. Elmoataz (2004). Visual Object Detection Using General Gaussian in Active Region Model, *SEE/IEEE Int. Conf. on Complex Systems, Intelligence and Modern Technology Applications (CSIMTA'2004)*, Cherbourg.

## Summary

In this paper, we present a new approach to segment visual objects in a video combining active contours with an implementation of optical flow. The developed approach is based on a minimization of a functional of energy (E) which uses perceptual information in ROI in an image, in conjunction with a mixture of Gaussian to model pixels of the background image and those of the semantic objects. To improve the detection stage, we extended the energy formulation of our model of active contour by including an additional force resulting from the calculation of the optical flow. We demonstrate the robustness of the proposed algorithm on real sequences.