

Relaxation Markovienne et seuillage par hystérésis pour une détection de mouvements temps réel dans des séquences couleur

Ghilès Mostafaoui¹

¹LISIF case courrier 252

4 place Jussieu 75252 Cedex 05

ghiles.mostafaoui@lis.jussieu.fr

Thomas Kunlin² Lionel lacassagne²

²IEF- AXIS, Université Paris-Sud, Centre d'Orsay,

91405 Orsay cedex

lionel.lacassagne@ief.u-psud.fr

Résumé

Cet article décrit une approche originale en couleur pour l'extraction d'objets en mouvement par champs de Markov. La robustesse de la relaxation markovienne est augmentée par un post-traitement de type seuillage par hystérésis tirant partie de l'information couleur. La cadence de traitement ainsi obtenue (80 images monochromes 512x512 par seconde et 25 images couleur 256x256) permet l'utilisation de notre algorithme dans diverses applications telles que la télé-surveillance ou la compression d'images.

1. Introduction

La détection de mouvements est un domaine de recherche de plus en plus répondu en traitement d'images. Les applications de ce dernier sont nombreuses et variées allant de la réalité virtuelle à l'interaction homme-machine en passant par la vidéo surveillance et la compression d'images.

Il existe dans la littérature plusieurs techniques de détection de mouvement, nous pouvons citer: la différence d'images, le flot optique introduit par [4], la mise en correspondance de points ou de blocs, la détection de contours en mouvement, les filtres de Gabor, les lignes de niveaux, la transformée de Hough...etc.

En parallèle de ces méthodes se sont développés des algorithmes à base de champs de Markov (*Markov Random Field*) dans de nombreux domaines du traitement d'image (détection de contours, segmentation de régions, restaurations d'images bruitées [4]).

Ces méthodes ont été appliquées à la détection de mouvement afin d'améliorer la qualité des résultats (augmentation du ratio signal utile sur bruit)

Nous présentons ici une extension aux séquences couleur de la détection de mouvements par une méthode dite « relaxation markovienne » basée sur des champs de Markov. Différents « post-traitements », tirant profit de l'information couleur, sont appliqués à la fin de chaque relaxation, améliorant ainsi la qualité de la détection.

La première partie de cet article décrira les méthodes markoviennes monochromes classiques de détection de mouvement. La seconde, présentera une version couleur de ces algorithmes. Nous montrerons ensuite des résultats expérimentaux avant de conclure.

2. Markov monochrome

La relaxation markovienne présentée ici a été introduite par [1], le modèle d'énergie utilisé dérive de celui présenté dans [7].

Soient à l'instant t , $I_{(t)}$, l'image en niveau de gris et $O^{(t)}$, l'observation c'est à dire la différence absolue de deux images consécutives (ou de l'image courante et une image de référence). Ces observations, une fois binarisées, servent d'estimation au champ d'étiquettes notées $\hat{E}^{(t)}$. L'ICM (Iterated Conditional Modes), algorithme de relaxation déterministe est utilisé pour déterminer le champ d'étiquettes relaxées $E^{(t-1)}$ (résultat final de la détection) à partir de $\hat{E}^{(t)}$, $\hat{E}^{(t)}$ et $\hat{E}^{(t)}$ (figure 1).

La relaxation étant déterministe, l'algorithme converge en quelques itérations, typiquement quatre, vers le premier minimum local (voir [6] et [2] pour plus de détails sur la relaxation markovienne).

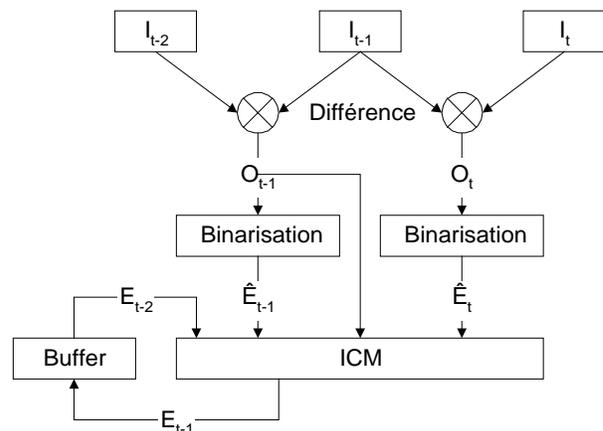


Figure 1. Relaxation Markovienne par ICM

3. Markov couleur

Nous avons introduit précédemment les bases de la détection de mouvement par relaxation markovienne pour les séquences d'images monochromes. Si l'efficacité de cette technique n'est plus à prouver, pour certaines séquences très bruitées (mauvaises conditions d'acquisition, variations brusques de luminosité, etc.), la détection peut être imprécise conduisant à une mauvaise segmentation des objets en mouvement.

Afin de pouvoir combler ces lacunes, nous nous sommes intéressés à l'apport de la couleur pour la relaxation markovienne ainsi qu'aux post-traitements tirant profit de cette nouvelle source d'information.

3.1 Pré-traitement et ICM couleur

Le principe est identique à celui du Markov monochrome. On notera $I_{(t)}$ l'image à l'instant t , $I_{(t)}$ est composée de trois plans :

$$I_{(t)} = \begin{pmatrix} I_{(t)}^R \\ I_{(t)}^G \\ I_{(t)}^B \end{pmatrix}$$

L'image d'observation $O_{(t)}$ est calculée entre 2 images consécutives, ou entre l'image courante et une image de référence.

$$O_{(t)} = |I_{(t)} - I_{(t-1)}|, \quad O_{(t)} = |I_{(t)} - I_{(ref)}|$$

L'image d'étiquettes estimées $\hat{E}_{(t)}$ est obtenue par seuillage binaire de chacun des trois plans R, G et B avec trois seuils θ_R , θ_G et θ_B .

L'image obtenue est visuellement interprétable : c'est une image où la couleur de chaque pixel (de l'image $\hat{E}_{(t)}$) indique les plans initialisés à un. Les trois couleurs primaires (rouge, vert et bleu) indiquent que le pixel comporte un seul plan à 1. Les couleurs secondaires jaune ($R+G$) magenta ($R+B$), cyan ($G+B$) indiquent que deux plans sont à un, et la couleur blanche que les trois plans sont à un.

Afin d'obtenir l'image relaxée (elle aussi visuellement interprétable), on applique (sur l'image $\hat{E}_{(t)}$) l'ICM plan par plan indépendamment des deux autres. On gardera ici la notation $\hat{E}_{(t)}$ pour l'image qui résulte de la relaxation couleur afin de la différencier de l'image finale de la détection (après les post-traitements).

3.2 Post-traitement

Le but de ce post-traitement est double : filtrer le bruit ayant résisté à la relaxation couleur et récupérer des pixels en mouvement, qui n'auraient pas été classés comme tels par une relaxation monochrome. L'image finalement obtenue est binaire, le pixel étant classé comme étant en mouvement ou non.

Nous avons développé trois types de post-traitement :

- la *diffusion* de niveau k ,
- le *ceiling* de niveau k ,
- le *seuillage par hystérésis*

Rappelons que le résultat de la relaxation markovienne couleur (avant les post-traitements) est une image ($\hat{E}_{(t)}$) de type (R,G,B) binaire sur chaque plan.

3.2.1 La diffusion et le ceiling de niveau k

Soient (R,G,B) les composantes couleur d'un pixel p de l'image $\hat{E}_{(t)}$ (R, G et B ne peuvent prendre ici que les valeurs zéro ou un).

La *diffusion* de niveau k ($k=1..3$) est appliquée à la fin de la relaxation (ICM), son principe est le suivant : Les composantes (R,G,B) du pixel p passent à un (pixel en mouvement) si $R+G+B \geq k$ et sont mis à zéro sinon.

Le *ceiling* de niveau k ($k=2..3$) est appliquée à l'intérieur de la boucle de l'ICM, son principe est le suivant : Les composantes (R,G,B) du pixel p restent invariants si $R+G+B \geq k$ et sont mis à zéro sinon.

Les avantages et les inconvénients observés d'après expériences sont les suivants :

-La *diffusion de niveau 1* (un pixel est dit en mouvement s'il est détecté en mouvement sur l'un des plans R,G,B) réalise une détection correcte : les régions détectées en mouvement sont bien remplies, en revanche, le bruit reste présent. Ce post-traitement est donc à réserver aux séquences faiblement bruitées.

-La *diffusion de niveau 2* (un pixel est dit en mouvement s'il est détecté en mouvement sur au moins deux des plans R,G,B) diminue significativement le bruit par rapport à la *diffusion 1*, néanmoins la perte d'information est visible : les régions détectées en mouvement sont moins homogènes.

-Le *ceiling 2* est plus robuste que la *diffusion 2*, en effet, le fait de ne pas changer les composantes du pixel évite de propager du bruit supplémentaire aux autres composantes. La robustesse au bruit est aussi accrue par la mise à zéro des pixels de bruit durant les 4 itérations de l'ICM.

La *diffusion de niveau 3* (un pixel est dit en mouvement s'il est détecté en mouvement sur les trois plans R,G,B) et le *ceiling 3* apportent une très bonne robustesse au bruit. Une perte importante de pixels en mouvement est cependant visible, spécialement pour le *ceiling 3*.

Ces post-traitements améliorent nettement la détection de mouvement dans les séquences bruitées. L'information couleur est donc effectivement bien prise en compte et permet d'aboutir à une meilleure détection.

Néanmoins, le dilemme "diminuer le bruit ou récupérer de l'information" se pose toujours. Le choix est souvent

difficile sachant que le résultat de la détection est généralement envoyé à une couche "haut niveau" d'analyse de mouvement. Cette dernière peut en effet être handicapée, aussi bien par le bruit que par le manque d'informations utiles. Une solution éventuelle à ce problème, est d'intervenir non pas en aval de la relaxation, mais en amont, en ayant des seuils de binarisation qui sont en fonction du passé du pixel et d'un système de pistage [5].

3.2.2 Seuillage par hystérésis

Une autre solution permettant de diminuer le bruit sans perdre de l'information, serait une combinaison entre la *diffusion 3*, pour les zones fortement bruitées et la *diffusion 1*, pour les zones faiblement bruitées. Cette stratégie s'apparente donc à l'application d'un hystérésis aux étiquettes relaxées.

Le principe du seuillage par hystérésis est le suivant : étant donnés deux seuils haut et bas s_H, s_B

- si la valeur du pixel est supérieure à s_H , le pixel est conservé,
- si la valeur du pixel est inférieure à s_B , le pixel est rejeté,
- si la valeur du pixel est comprise entre s_B et s_H , le pixel est conservé s'il existe un chemin (un ensemble de pixels connexes) menant à un pixel de valeur supérieure à s_H .

Afin d'appliquer le seuillage par hystérésis à notre problématique, on utilisera une image monochrome noté $H_{(t)}$ créée à partir de l'image $\hat{E}_{(t)}$.

Soient (R,G,B) les composantes couleur d'un pixel p de l'image $\hat{E}_{(t)}$ (binaire sur chaque plan). Le niveau de gris du pixel correspondant à p dans l'image $H_{(t)}$ est alors égale à $R+G+B$. Les niveaux de gris dans l'image $H_{(t)}$ varient alors entre zéro et trois.

On applique alors un seuillage par hystérésis sur $H_{(t)}$. Les trois couples de valeurs (s_B, s_H) possibles sont $(1,2)$, $(1,3)$ et $(2,3)$. Le résultat de ce traitement nous donnera alors les images binaires finales des étiquettes relaxées.

Ce traitement, plus local (critère de connexité), apporte à l'algorithme une meilleure robustesse aux bruits en plus d'un caractère adaptatif aux différentes séquences.

4. Résultats expérimentaux

Notre algorithme a été testé sur un grand nombre de séquences avec diverses conditions d'acquisition. Nous présentons ici les résultats obtenus sur deux d'entre elles. Dans la première, sont représentées deux personnes en mouvement. Cette séquence présente un bruit non négligeable dû aux variations de luminosité.

La figure 2 montre le résultats de la relaxation markovienne (avant post-traitements) sur chacun des trois plans (R,G,B) avec deux différents seuillages des images

de différence. Le premier assez bas avec $\theta_R=\theta_G=\theta_B=15$ et le second plus haut avec $\theta_R=\theta_G=\theta_B=30$.

Nous pouvons constater que l'utilisation d'un seuillage bas a pour effet de générer un bruit assez

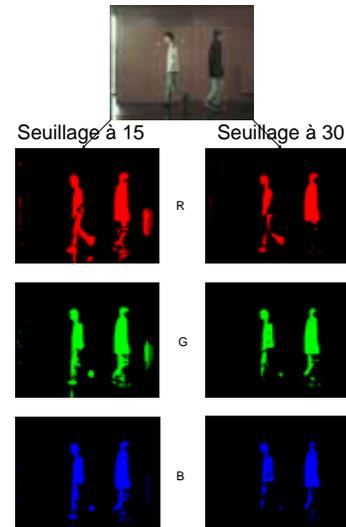


Figure 2. Résultat de la relaxation markovienne sur les trois plans avec deux différents seuils

important (à droite de l'image) notamment sur la composante R . Un seuillage haut permet de supprimer ce bruit mais implique une perte de l'information utile (les personnes en mouvement sont mal segmentées).

Une première solution est ici d'appliquer un seuillage différent sur chacun des plan (R,G,B) . En effet, avec $\theta_R=30$ et $\theta_G=\theta_B=15$ nous pouvons diminuer le bruit de la composante R et garder de l'information utile sur les deux autres composantes. L'image $\hat{E}_{(t)}$ (binaire sur chacun de ses plans couleur) qui en résulte est illustrée figure 3.



Figure 3. Image des étiquettes relaxées avant post-traitements

On remarquera qu'il subsiste encore du bruit sur les composantes G et B , ceci est dû au seuil bas appliqué sur ces composantes.

Le dilemme "diminuer le bruit ou récupérer de l'information" est ici bien illustré. En effet, l'application de post-traitements de type *diffusion* ou *ceiling* de niveau 1 ou 2 ne permettra pas de supprimer le bruit car il est présent sur deux plans couleur. Une *diffusion* ou un

ceiling de niveau 3 ne permettra de récupérer que les pixels « blancs » dans l'image figure 3. Ce qui implique une grande perte de l'information utile.

Cette problématique justifie l'utilisation d'un post-traitement de type *hystérésis*.

Prenons $s_B=1$ et $s_H=3$. Les pixels détecté en mouvement sur ses trois composantes sont alors conservés (pixels blancs dans la figure 3). Les pixels ayant au moins été détectés en mouvement sur l'un des plans (R,G,B) ne sont conservé que si ces derniers sont connexes à des pixels détectés en mouvement sur les trois plans.

Le résultat de ce post-traitement, illustré figure 4, montre bien que le bruit est fortement diminué tout en gardant l'information utile.

Il est néanmoins important de noter que pour certaines séquences très bruitées, les étiquettes sont mal initialisées. On préférera alors une *diffusion* ou un *ceiling* de niveau 3 à un seuillage par hystérésis. En effet, ce dernier aura pour effet (si il y a connexité des pixels de bruit avec ceux des objets en mouvement) de propager le bruit.



Figure 4. Résultat de la relaxation markovienne avec un seuillage par hystérésis

Afin de pouvoir comparer les résultats des différents post-traitements, une mesure de l'erreur a été effectuée grâce à une vérité de terrain grossière.

Dans un premier temps, nous avons « détourné » à la main sur une vingtaine d'image de chaque séquence les contours des objets en mouvement. Ces contours (fermés) nous permettrons d'avoir les coordonnées des pixels appartenant aux objets en mouvement (pixels des régions entourées par les contours) et celles des pixels faisant partie du fond immobile.

Nous effectuons alors deux mesures. La première représente le pourcentage de pixels appartenant aux objets en mouvement ayant été mis à zéro (notée E_f) et la seconde le pourcentage de pixels du fond immobile ayant été mis à un (notée E_b).

Le tableau 1 illustre les résultats obtenus pour les différents post-traitements appliqués sur la séquence cité ci-dessus (notée seq1) et sur une séquence (figure 4), provenant de la conférence Pets2002. Cette dernière présente un certain nombre de difficultés tels que les variations de luminosité, l'occultation des personnes en mouvement par des objets immobiles (lettres sur la vitre), la présence de reflets sur les vitres...etc



Figure 5. Image de la séquence pets2002

	Diffusion1	Ceiling2	Ceiling3	Hystérésis
Seq1	$E_f=0.5\%$ $E_b=3\%$	$E_f=0.8\%$ $E_b=0.9\%$	$E_f=1\%$ $E_b=0.5\%$	$E_f=0.5\%$ $E_b=0.7\%$
Pets2002	$E_f=1\%$ $E_b=5\%$	$E_f=1.5\%$ $E_b=1.5\%$	$E_f=4\%$ $E_b=0.6\%$	$E_f=1\%$ $E_b=0.8\%$

Tableau 1

Conclusion

Dans cet article nous avons présenté un nouvel algorithme couleur de détection de mouvement par champs de Markov et hystérésis. L'utilisation de trois plans couleur, de leur combinaison et d'un seuillage par hystérésis rend cette technique de détection robuste, y compris dans des conditions d'éclairage difficiles.

Comme perspectives nous envisageons une gestion automatique des seuils de binarisation (par un processus EM par exemple) afin d'améliorer l'initialisation des champs d'étiquettes. Une combinaison de la version couleur avec une version monochrome hiérarchique permettra d'augmenter la robustesse de l'algorithme.

References

- [1] A. Caplier. "Modèles markoviens de détection de mouvements dans les séquences d'images : approche spatio-temporelle et mise en oeuvre temps réel". Thèse INPG 1995.
- [2] A.Caplier, F.Luthon, C.Dumontier, "Real-Time implementations of an MRF-Based Motion Detection Algorithm", RealTimeImg(4), No. 1, February 1998, pp. 41-54.
- [3] S Geman, D. Geman "Stochastic relaxation, Gibbs distributions, and the bayesian Restoration of images". IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol 6, No 6, November 1984, pp 721-741
- [4] B. K. P Horn, B. G. Schunk, "Determining optical Flow", Artificial Intelligence, No 17, pp 185-203, 1981
- [5] Y. Z. Hsu, H. H. Nagel, G. Rekers. "New likelihood test methods for change detection in image sequences" CVGIP 26, 1984.
- [6] L. Lacassagne, M. Milgram, P. Garda, "Motion detection, labeling, data association and tracking in real time on RISC computer", International Conference on Image Analysis and Processing, 1999.
- [7] P. Lalande. "Détection du mouvement dans les séquences d'images selon une approche markovienne, application à la robotique sous marine", Thèse Université de Renne I, 1990.