

Détection stochastique du mouvement dans les séquences vidéos

Application à la compression

Michel Collobert, Olivier Bernier, Daniel Collobert

France Telecom R&D

FTR&D Pièce 112LF rue P. Marzin 22307 Lannion Cedex
michel.collobert@francetelecom.com

Résumé

Nous proposons dans cet article une nouvelle méthode de détection de mouvement par différence d'image qui, en utilisant des propriétés statistiques locales, permet de s'affranchir de la notion de seuil de différence.

Une notion de « zone » de différence est inhérente à cette méthode. Une première évaluation dans le cadre de la visiophonie montre que ces zones ne représentent que de 5 à 10 % de l'image initiale.

Nous proposons cette méthode comme brique pouvant servir à la compression de flux vidéo dans le cas particulier de prise de vue avec caméra fixe.

Mots Clef

Détection de mouvement, statistique, vidéo, compression.

1 Travaux antérieurs

De nombreux algorithmes de détection du mouvement existent, par exemple les méthodes basées sur le flux optique, sur les appariements de blocs, de contours etc.. [1][2][3].

Globalement ce sont des algorithmes coûteux en temps de calcul.

La méthode la plus simple et la plus rapide est de calculer l'image de différence de luminance en chaque pixel entre deux images consécutives de la séquence vidéo. A cause du bruit induit par l'électronique et les capteurs de la caméra, un seuil S doit être appliqué. L'image de différence est ensuite binarisée suivant des relations de la forme :

pour un pixel $I(i, j)$

$I(i, j) = 0$ si valeur absolue (différence) $\leq S$

$I(i, j) = 1$ si valeur absolue (différence) $> S$

Typiquement, pour les caméras mono CCD, ce seuil est compris dans une fourchette de 3 à 5% de la luminance maximale possible.

Pour une image vidéo comportant un personnage en mouvement tel que représenté figure 1 (image obtenue avec une caméra Canon VC-C1)



fig 1 : image vidéo de référence

on obtient l'image de différence seuillée de la figure 2.



fig 2 : image de différence seuillée

2 Méthode proposée

Nous proposons une méthode permettant de se passer de ce seuil en prenant en compte le fait que, spatialement, la moyenne du bruit sur la différence (non absolue) est nulle si il n'y a pas de mouvement.

Il existe donc, pour un pixel de l'image de différence correspondant à une zone en mouvement, une dissymétrie entre la probabilité d'être positif et celle d'être négatif.

C'est ce que nous allons mettre en évidence par l'algorithme suivant qui comporte trois phases principales :

- Quantification ternaire de la différence d'image.
- Filtrage pour supprimer les effets de corrélation spatiale du bruit.
- « Zonage » permettant de mettre en évidence toutes les zones de l'image dont le produit scalaire entre le gradient spatial local et le mouvement des objets est non nul.

2.1 - Quantification ternaire

Comme, pour un pixel de l'image différence et malgré le bruit, il existe une dissymétrie entre la probabilité d'être positif et la probabilité d'être négatif sur toutes les zones où localement la projection du vecteur mouvement sur le vecteur gradient de luminosité est non nul, on utilise une quantification conservant cette information dans l'image de différence :

pour un pixel $I(i, j)$

- $I(i, j) = 0$ si la différence est nulle
- $I(i, j) = 1$ si différence est positive
- $I(i, j) = -1$ si différence est négative

Un exemple de résultat est donné figure 3

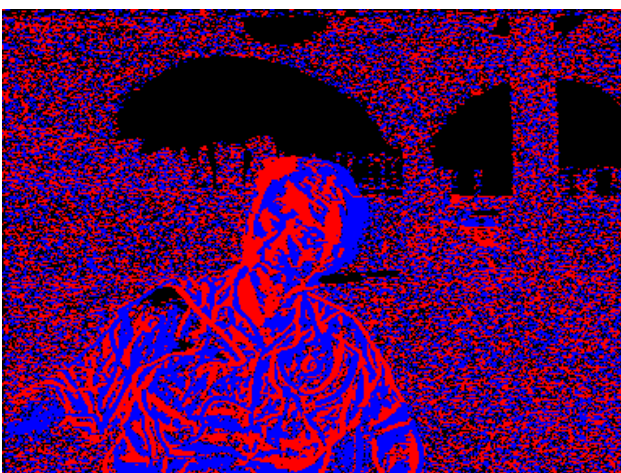


fig3 : en rouge différence positive, en bleu différence négative

Pour isoler le mouvement qui nous intéresse, il faut donc supprimer de cette image de différence tous les pixels

correspondant au bruit. C'est ce que réalisent les opérations présentées dans les deux paragraphes suivants :

2.2 Filtrage pour suppression des effets de corrélation du bruit

Un phénomène de « percolation », dû à des phénomènes physique de corrélation spatiale, introduit des zones « filamenteuses » ainsi que le montre l'image suivante montrant le détail d'une zone de l'image de différence due au bruit (figure 4).



fig 4 : détail d'une zone correspondant au bruit de notre caméra

Pour éviter ce phénomène, un filtrage très simple (en une passe) est appliqué à cette image de différence ternaire pour « casser » ces filaments.

Pour chaque pixel (valeur $\in \{-1, +1\}$)
si ce pixel se trouve entre deux pixels,
horizontalement ou verticalement,
de valeurs différentes de la sienne,
alors ce pixel est mis à 0

fin si
fin pour

On réduit ainsi le bruit précédent en bruits isolés (fig 5).

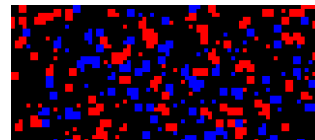


fig 5 : détail d'une zone de bruit après filtrage

2.3- Segmentation des zones correspondant à du mouvement

Sur la figure précédente (fig 5), on voit que les zones correspondant au bruit de la caméra sont 'atomisées'.

L'étape suivante du traitement consiste donc, à partir de l'image de différence ternaire ayant subi le filtrage précédent, à rajouter un autre filtrage supprimant ces 'atomes'.

Nous avons effectué ce nouveau filtrage en utilisant une technique de croissance de région basique utilisant une connexité d'ordre 4 sur les pixels de l'image de différence qui sont à +1 ou à -1 et pas à 0. Ces petites zones sont supprimées en fixant un seuil cette fois-ci sur la taille de la zone d'intensité. A l'usage, la valeur de ce seuil n'est pas très critique.

On obtient ainsi la segmentation des objets en mouvement représenté figure 6.

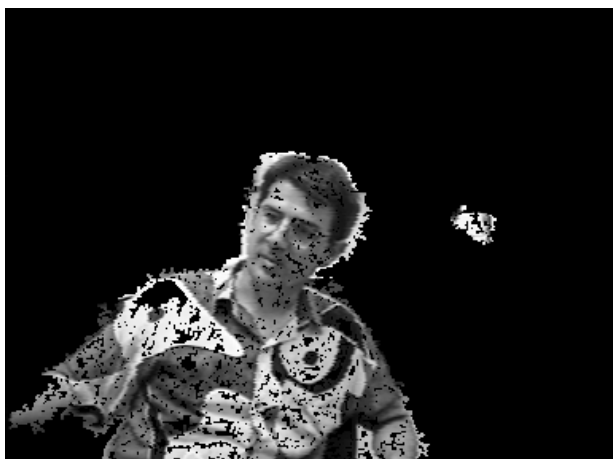


fig 6 : en résultat de l'ensemble des traitements, 2 objets sont segmentés. La personne grâce à son mouvement, l'écran du moniteur grâce au balayage vidéo non synchronisé à la caméra de prise de vue.

2.3. Conclusion sur la méthode

La comparaison des figures 2 et 6 montre clairement que notre méthode apporte une segmentation beaucoup plus pertinente que la méthode par différence seuillée classique. En effet, au lieu d'obtenir le mouvement surtout sur les endroits proches des contours, une grande partie des zones composant un objet en mouvement est facilement mise en évidence.

Le temps de calcul est intéressant (25 img/sec, processeur Alpha21164, 600Mhz) et la technique très sensible : la détection de variations globalement en dessous du bruit de la caméra peut être effectuée, comme la montre cette visualisation de la fumée d'une cigarette qui est à peine observable par un opérateur humain regardant l'image sur un moniteur vidéo (fig 7 et 8)

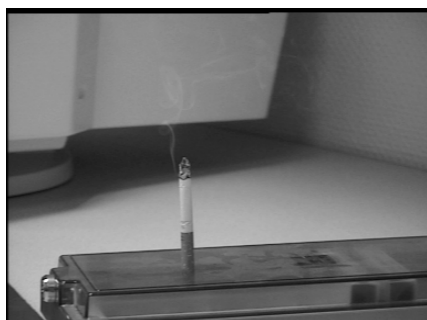


fig 7 et 8 détection de fumées.

3. Application à la compression

Par cette technique, même les petits mouvements des lèvres, des sourcils, ou de la respiration sont repérés et isolés comme le montre l'image figure 9. Les zones concernées peuvent facilement être caractérisées au cours du processus de croissance de région par des hexagones exinscrits.

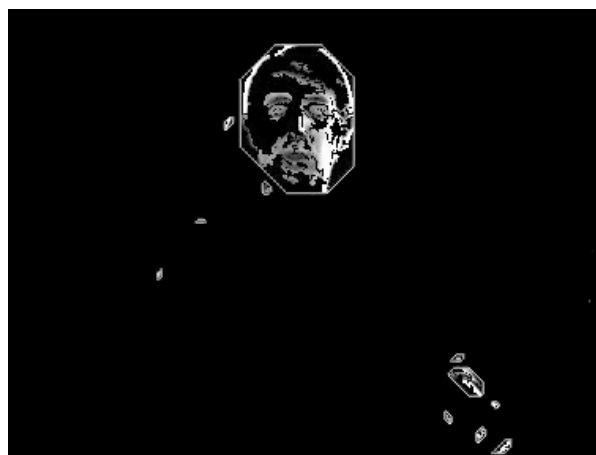


fig 9 : détections de petits mouvements

Dans le cadre d'une visio-phonie ou visio-conférence typique (caméra fixe) on peut constater que les personnes bougent peu en moyenne. Une première estimation sur des séquences de ce type, montre que seuls 5 à 10 % des pixels sont affectés en moyenne par les mouvement d'un locuteur.

L'élaboration d'un algorithme de compression basé sur la seule transmission des zones ainsi déterminées peut donc nous laisser espérer un taux de compression sans, ou quasi-sans perte (par compression supplémentaire des zones elles-mêmes), de l'ordre de 10 ou 20.

Cette transmission quasi sans perte des petits mouvements (mouvements des sourcils et des lèvres, moues, respirations...) est très importante pour la qualité perceptive de la communication.

Il peut donc être intéressant d'explorer plus en avant cette voie, notamment par des expérimentations sur des séquences vidéos types, dans le cadre par exemple

de visiophonie en intra-net (débit variable, PC de bureau avec caméra fixe).

Bibliographie

[1] L. Torres and M. Kunt. *Vidéo Coding, The Second Generation Approach*, pp. 219-263 (Kluwer Academic Press 1996).

[2] B.K.P. Horn, B.G. Schunck. « Determinig optical flow » : a retrospective. *Artificial Intelligence* 59 (1993) pp. 81-87.

[3]O. Bernier « Analyse du flot optique dans les séquences d'images par réseau d'automates » *Quinzième Colloque sur le Traitement du Signal et des Images* Juan-les-pins France 1995.