

**APPROCHE ADAPTATIVE DE L'ESTIMATION DU MOUVEMENT  
DANS UNE SÉQUENCE D'IMAGES PAR FILTRAGE DE KALMAN  
TEMPOREL**

Florence GERMAIN \* # - Daniel BAUDOIS \*

# ITMI, Zirst BP 177, 38244 Meylan Cedex

\* CEPHAG, BP 46, 38 402 - Saint Martin d'hères Cedex

RÉSUMÉ

Du fait de la représentation explicite de l'incertitude, le filtre de Kalman est un outil puissant pour l'estimation du mouvement dans les images. Nous décrivons comment ce filtre, rendu adaptatif par la prise en compte des ruptures sur le signal d'innovation, permet un couplage étroit entre les sous-processus d'estimation et de segmentation du champ des vitesses. Le principe de détection des ruptures sur l'innovation exploite la relation entre l'évolution temporelle et la distribution spatiale du champ de vecteurs vitesse, au voisinage d'une frontière de mouvement. Ce même principe confère à notre estimateur de bonnes capacités de prise en compte des discontinuités du champ de vecteurs vitesse image.

1. INTRODUCTION

On note un intérêt croissant pour l'analyse du mouvement dans les images, émergeant simultanément de divers domaines d'application, depuis l'analyse de scènes dynamiques jusqu'au codage pour la transmission des séquences d'images [1][3][4]. Dans le cadre de l'analyse d'images naturelles, en particulier d'images non structurées, l'information dense de mouvement image est un puissant support d'aide à la segmentation et à l'interprétation qualitative de scènes. Cet article traite de l'estimation du champ de vitesses apparentes, déduit des variations de l'intensité lumineuse, dans le cas d'une séquence d'images.

Problématique

L'information de mouvement apportée par le gradient de l'intensité en un point donné - dite **information mesurable** - est entachée d'une incertitude anisotrope, variable selon les zones de l'image, et inhérente à la nature du problème. La connaissance a priori de certaines propriétés locales du champ de vecteur à estimer peut contribuer à réduire cette incertitude:

- (i) - continuité du vecteur vitesse, le long de la trajectoire d'un point lumineux image,
- (ii) - continuité spatiale du champ de vecteurs vitesse image.

Sous réserve de la vérification des hypothèses de continuité, l'incertitude peut être réduite en fonction directe de la diversité spatio-temporelle de la direction du gradient spatial de l'intensité. Lorsque la diversité est insuffisante, aucune

ABSTRACT

Because of the explicit representation of uncertainty, the Kalman filtering technique is well adapted to image motion estimation. We describe how an adaptive Kalman filter, using the detection of abrupt changes in the innovation signal, allows a strong interaction between motion estimation and motion segmentation subprocesses. The detection of change is based on the link between the temporal evolution and the spatial distribution of the estimated motion field, near a motion discontinuity. This adaptivity principle makes Kalman filtering perfectly suitable for handling motion frontiers.

considération de nature locale, quelle qu'elle soit, ne saurait permettre de lever l'incertitude résiduelle.

Dans la suite, on considère la propriété (i) vérifiée. La propriété (ii) est, elle, supposée vérifiée à l'exception d'un sous-ensemble de dimension un, inconnu a priori: les **frontières de mouvement**. L'étroite imbrication des sous-problèmes d'estimation du mouvement d'une part, de détection des frontières de mouvement d'autre part, constitue une des difficultés majeures de l'analyse du mouvement image.

Principe de notre approche

Le problème d'estimation du flot optique se décompose en deux phases:

- 1- mesure de l'information de mouvement portée par le gradient de l'intensité, avec évaluation de l'incertitude associée,
- 2- propagation de cette information en direction des zones à forte incertitude associée. Cette phase de propagation, fondée sur des critères locaux de continuité, est appelée **régularisation**. La régularisation est doublement contrainte par l'obligation de ne pas altérer l'information dans les zones à faible incertitude (C1), et de respecter des discontinuités de mouvement (C2). La difficulté majeure de l'analyse du mouvement réside dans le respect de la contrainte C2.

Notre approche consiste à intégrer les phases d'estimation du mouvement et de détection des discontinuités, au sein d'un processus incrémental d'estimation locale. L'utilisation du filtre de Kalman permet de prendre en compte explicitement l'incertitude liée à l'information mesurée, garantissant ainsi le respect de la contrainte C1. La régularisation est réalisée dans



le modèle dynamique du système à estimer. La détection des frontières de mouvement est menée, tout au long de l'estimation, au moyen d'une détection de ruptures sur le signal d'innovation. Après franchissement d'une frontière de mouvement, le filtre est réinitialisé en fonction de la frontière rencontrée, ce qui assure qu'aucune régularisation n'a lieu à travers une frontière de mouvement (respect de C2).

## 2. FILTRE DE KALMAN TEMPOREL POUR L'ESTIMATION INCRÉMENTALE DU MOUVEMENT

Nous considérons dans un premier temps une séquence ne contenant pas de frontière de mouvement. On notera  $I(p,t)$  la fonction qui à tout pixel  $p$  et à tout instant  $t$  associe l'intensité du point lumineux présent en  $p(x,y)$  à  $t$ , et  $\nabla I$  le gradient spatial de  $I$ . On notera  $w(p,t)$  le vecteur déplacement instantané associé au point lumineux coïncidant avec le pixel  $p$ , à l'instant  $t$ . Le filtre décrit réalise l'estimation temporelle de la vitesse  $w(p,t)$  en un pixel  $p$  fixé. Le champ de vecteurs est obtenu par activation d'un filtre en chaque pixel.

### Vitesse orthogonale

L'intensité lumineuse est supposée constante le long de la trajectoire d'un point image:

$$I(p+w(p,t), t+1) = I(p,t) \quad (1)$$

Par développement limité d'ordre 1 de la fonction d'intensité, on dérive de l'équation (1) l'Equation de Contrainte du Mouvement (ECM):

$$\nabla I \cdot w + \delta I / \delta t = 0 \quad (2)$$

qui donne la composante dite **composante orthogonale** du vecteur vitesse local en fonction des dérivées spatiales et temporelles de l'intensité.

Plus généralement, on montre que l'incertitude associée à l'information mesurable est minimale dans la direction du gradient spatial de l'intensité, et maximale dans la direction orthogonale. Cette propriété est connue sous le nom de **phénomène d'aperture** [4].

### Equations du filtre

L'équation d'évolution du filtre intègre un modèle d'état constant, à un bruit blanc gaussien centré près:

$$w(p,t+1) = w(p,t) + r(p,t) \quad r(p,t) \approx (0, R(p,t)) \quad (3)$$

Le choix d'une covariance  $R(p,t)$  non nulle permet au filtre de suivre l'évolution d'un vecteur vitesse  $w(p,t)$  faiblement variable dans le temps.

L'observation  $z(p,t)$  en un point  $p$  à l'instant  $t$  est la composante orthogonale du déplacement cherché, entachée d'un bruit blanc gaussien centré.

$$z(p,t) = \nabla I(p,t) \cdot w(p,t) + s(p,t) \quad s(p,t) \approx (0, S(p,t)) \quad (4)$$

On montre que l'incertitude sur l'information de mouvement mesurée est une fonction décroissante du gradient spatial local de l'intensité, dans la limite de validité de l'ECM. Une

expression analytique de l'incertitude en fonction de  $\|\nabla I\|$  est donnée par Simoncelli et Al. [2]. La covariance  $S(p,t)$  peut de ce fait être explicitée, en chaque point  $(p,t)$ , et injectée en continu dans le filtre.

La matrice de covariance  $P_{t|t}$  sur l'erreur d'estimation est diagonale dans le repère  $(\nabla I, \nabla I^\perp)$  [3]. La plus grande des valeurs propres est trouvée dans la direction  $\nabla I^\perp$ , et traduit l'incertitude liée au problème d'aperture. Le suivi temporel des valeurs propres de  $P_{t|t}$  est un traceur de la levée d'ambiguïté locale (voir fig 2).

## 3. PRISE EN COMPTE DES FRONTIERES DE MOUVEMENT

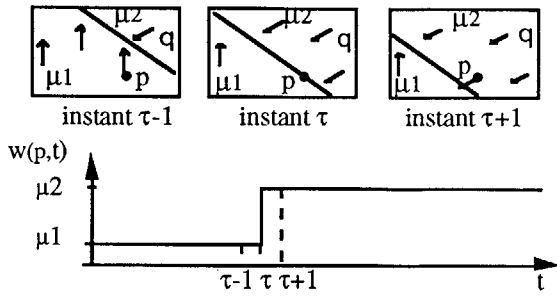
### Principe général

En un pixel donné  $p$ , le passage d'une frontière de mouvement se traduit par une inadéquation brutale et temporaire entre le modèle d'état du filtre et l'évolution du vecteur  $w(p,t)$ . Cette modification s'exprime localement par un saut de nature additive sur l'équation de modèle, et entraîne une modification brutale de la statistique du signal d'innovation du filtre (fig 4a). Une analyse en ligne de l'innovation peut permettre de détecter, puis de quantifier la perturbation portant sur le modèle d'état [6]. Dès que la rupture de modèle est connue, on peut réinitialiser le filtre en fonction de la perturbation enregistrée, et ainsi diminuer considérablement la durée du transitoire qu'elle engendre.

Sans information a priori sur l'amplitude de la rupture, la mise en oeuvre algorithmique d'une quantification de rupture de modèle pose un problème de retard à la détection, doublé d'un manque de robustesse. Dans le cas qui nous intéresse, la prise en compte d'une connaissance a priori de la vitesse "après rupture" permet de s'affranchir de ces problèmes. Plus précisément, considérons une rupture de modèle, se produisant en  $p$  à l'instant  $\tau$ . La distribution des vecteurs vitesse au voisinage du point  $p$ , à l'instant  $\tau-1$  fournit une connaissance a priori sur la vitesse  $w(p,\tau+1)$  après passage de la frontière (voir schéma ci-dessous). Dans la suite, nous supposons l'existence d'un pixel  $q$  appartenant à un voisinage de  $p$ , tel que  $w(q,\tau-1)$  soit un bon prédicteur de  $w(p,\tau+1)$  (hypothèse H1).

Dès lors, notre démarche pour l'identification d'une rupture en un point  $p$  est la suivante:

- Une recherche de changement brutal de statistiques est menée en continu sur le signal d'innovation
- Dès qu'une rupture est détectée (instant  $\tau$ ), on identifie les diverses composantes de la distribution des vitesses sur un voisinage  $V(p)$ , à l'instant  $\tau-1$  précédant l'instant courant. Notons  $N$  le nombre de composantes distinctes identifiées.
- Un banc de  $N$  filtres, déduits de chacune des  $N$  hypothèses de mouvement "après rupture", est activé à partir de  $\tau$ . Le filtre s'avérant le plus probable, compte-tenu des observations immédiatement postérieures à  $\tau$  sera retenu.



Passage d'une frontière de mouvement, en un pixel  $p$  :  
il existe  $q \in V(p)$  tel que  $w(q, \tau-1)$  soit proche de  $w(p, \tau+1)$

#### Principe détaillé

Dans ce qui suit, on considère un point  $p$  fixé. On notera  $V(p)$  le voisinage spatial de  $p$ , de taille  $T_v$  fixée a priori.

##### Etape 1: détection d'une frontière de mouvement

La présence d'une perturbation additive sur l'équation de modèle se traduit additivement sur l'innovation  $\gamma(p,t)$ . Le passage en  $p$  d'une frontière de mouvement peut alors être mis en évidence par un détecteur de saut de moyenne appliqué à  $\gamma(p,t)$  [6].

##### Etape 2: génération d'une famille d'hypothèses "après rupture"

Supposons qu'une rupture soit détectée en  $p$ , à l'instant  $\tau$ . On effectue une segmentation de  $V(p)$  à  $\tau-1$ , en  $N$  zones  $V_i$  (le nombre  $N$  de zones obtenues résulte de la segmentation) sur lesquelles le champ de vecteurs vitesse s'identifie à un vecteur constant, à un bruit gaussien centré près:

$$w(q) = w_i + b(q), b(q) \approx (0, \sigma^2), w_i = \text{Cst} \quad (5)$$

Une technique de segmentation au sens du mouvement, basée sur la seule connaissance de l'information mesurable, est décrite par J.Rivero et P.Bouthemy [5]. Cette étape délivre  $N$  hypothèses concurrentes  $\{w_i\}$  de mouvement après rupture.

##### Etape 3: activation d'un banc de $N$ filtres candidats

On définit  $N$  filtres  $\{F_i\}$ , déduits du filtre courant par réinitialisation de l'erreur d'estimation et de l'estimé courant:

$$P_{t|t} = P_{0|0}, w_{t|t} = w_i \quad (6)$$

Les  $N$  filtres sont activés en parallèle, à partir de l'instant  $\tau$ . Une probabilité conditionnelle  $P(F_i|z(p,t))$  est calculée à chaque instant, et pour chaque filtre, sur la base des observations  $z(p,\tau), z(p,\tau+1)\dots$ . L'application du théorème de Bayes donne:

$$P(F_i|z(p,t+1)) = \frac{P(z(p,t+1)|F_i) P(F_i|z(p,t))}{\sum_{j=1}^N P(z(p,t+1)|F_j) P(F_j|z(p,t))} \quad (7)$$

avec équiprobabilité initiale (à l'instant  $\tau$ ):

$$P(F_i|z(p,\tau)) = 1/N, \forall i.$$

Notons  $\gamma_i(p,t)$  l'innovation du filtre  $F_i$  et  $\sigma_i^2(p,t)$  la variance de l'erreur sur  $\gamma_i(p,t)$ . La probabilité  $P(z(p,t) | F_i)$  s'évalue à [7]:

$$P(z(p,t) | F_i) = \left| 2\pi\sigma_i^2(p,t) \right|^{-1/2} \exp\left\{ -\frac{\gamma_i^2(p,t)}{2\sigma_i^2(p,t)} \right\} \quad (8)$$

Un seul des  $N$  filtres en concurrence se révèle être un bon estimateur de l'évolution après rupture (voir hypothèse H1). La probabilité associée à ce dernier converge rapidement vers 1, tandis que les  $N-1$  autres convergent vers 0. Dès que la convergence est établie, le filtre correspondant est retenu comme estimateur "après rupture".

## 4. RESULTATS EXPERIMENTAUX

Cette section illustre notre approche, dans le cas d'une texture naturelle (figure1), animée d'un déplacement simulé.

### Estimation dans le cas d'une zone de mouvement homogène

Le comportement du filtre de Kalman est ici illustré dans le cas d'une translation uniforme, sans frontière de mouvement. La figure 1 donne le champ estimé aux instants  $t=1, 2$  et  $10$ .

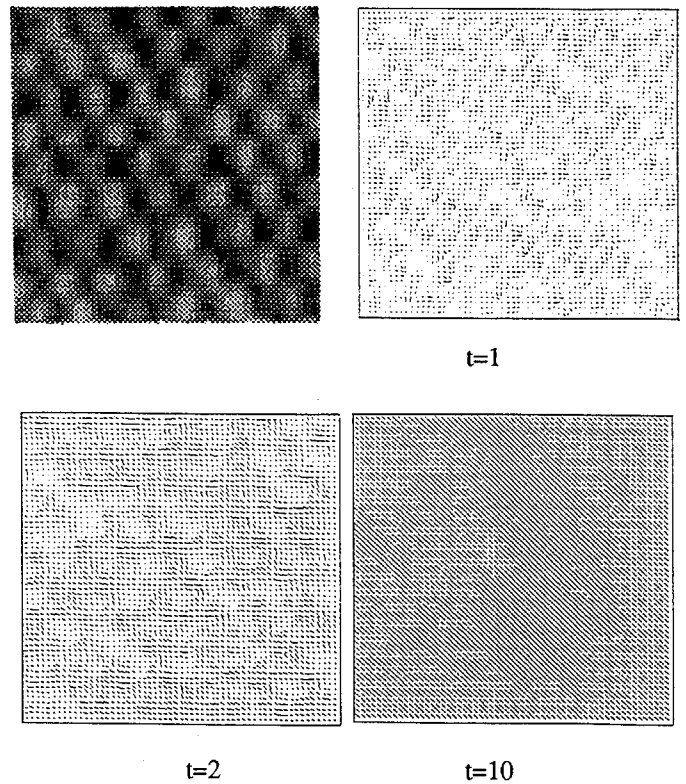


Fig.1 Image de base; champ estimé à  $t=1, 2, 10$

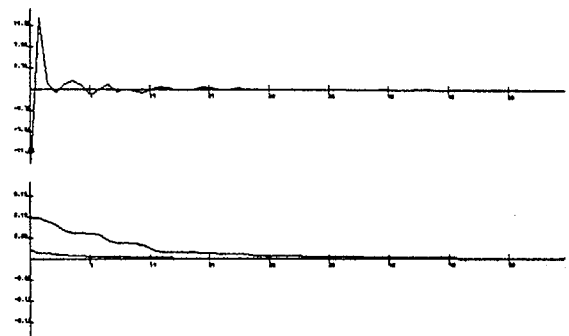


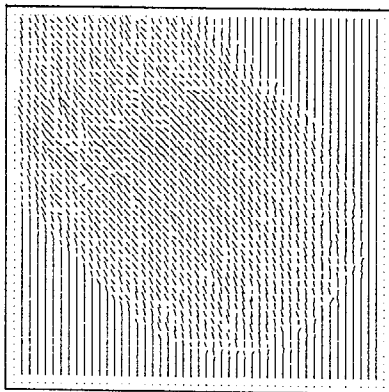
Fig. 2 Evolution temporelle de l'innovation (haut) et des valeurs propres de la matrice de covariance sur l'erreur d'estimation (bas), en un pixel  $p$  fixé



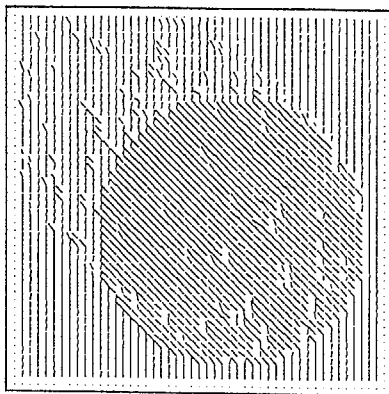
L'évolution temporelle de l'innovation ainsi que des valeurs propres de la matrice de covariance de l'erreur d'estimation, en un pixel  $p$  donné, sont présentées en figure 2. On notera la différence de vitesse de décroissance entre les deux valeurs propres, traduisant le problème d'Aperture .

#### Cas des frontières de mouvement

On considère maintenant le cas d'un disque texturé (texture de la figure 1), en translation oblique sur un fond de même texture. Le fond est lui-même animé d'un mouvement de translation verticale. D'un point de vue photométrique, disque et fond sont indiscernables. La figure 3 illustre le champ estimé à  $t=40$ , avec filtrage non adaptatif puis avec filtrage adaptatif. Dans le premier cas, on remarque une "trainée" constituée de pixels perturbés par le passage récent de la frontière de mouvement. Cet effet est annulé par la mise en oeuvre d'un filtre adaptatif.



a- filtre non adaptatif

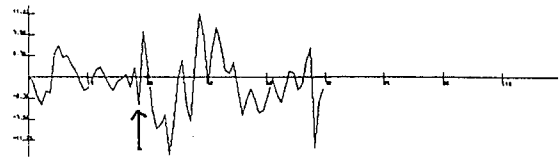


b- filtre adaptatif

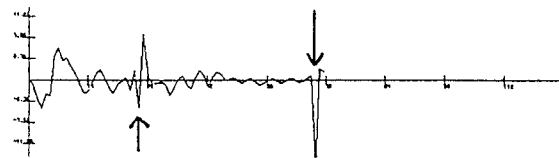
Fig. 3 Champ de vecteurs vitesse estimé ( $t=40$ )

L'évolution temporelle de l'innovation, en un point appartenant initialement au fond puis recouvert par le disque, et enfin à nouveau découvert, est donnée à la figure 4 dans le cas non adaptatif, puis adaptatif. Dans le premier cas, on constate une très longue perturbation de l'innovation, après passage de la première frontière de mouvement. Cette perturbation masque le passage de la seconde frontière. A l'inverse, les deux frontières apparaissent clairement sous forme de rupture sur l'innovation,

dans le cas adaptatif. Dans cet exemple, deux filtres candidats sont initialisés à chaque détection de rupture. Le temps moyen de convergence du "bon filtre" après rupture est de 3 itérations, pour une probabilité conditionnelle (7) supérieure à 0.99.



a- filtre non adaptatif



b- filtre adaptatif

Fig. 4 Evolution temporelle de l'innovation en un point du fond successivement recouvert par le disque, puis découvert

## 5. CONCLUSION

Le filtre de Kalman, grâce à la représentation explicite de l'incertitude, est un outil puissant pour l'estimation du mouvement dans les zones homogènes. Nous avons décrit comment ce filtre, rendu adaptatif par la prise en compte des ruptures sur l'innovation, permet un couplage étroit des sous-processus d'estimation et de segmentation du champ des vitesses. Le principe de détection des ruptures sur l'innovation utilise le lien existant entre l'évolution temporelle et la distribution spatiale du champ de vecteurs vitesse, au voisinage d'une frontière de mouvement. Ce même principe confère à notre estimateur une très bonne prise en compte des discontinuités du champ de vecteurs vitesse image.

## 6. REFERENCES

- [1] B.K.P Horn and B.G.Schuck, "Determining Optical Flow", *Artificial Intelligence*, 17, pp185-203, 1981
- [2] E.P.Simoncelli, E. H. Adelson and D. J. Heeger, "Probability Distributions of Optical Flow", *Computer Vision and Pattern Recognition*, pp310 - 315,1991.
- [3] J. Stuller and G. Krishnamurthy, "Kalman Filter Formulation of Low-Level Television Image Motion Estimation", *Computer Vision, Graphics and Image Processing*, Vol 21, n° 2, 1983.
- [4] A. Singh, "Optic Flow Computation: A Unified Perspective", IEEE Computer Society Press, 1991
- [5] J. S. Rivero and P. Bouthemy, "Region Segmentation According to Motion-Based Criteria", *PI* 340, 1987
- [6] M. Basseville, "Detecting Changes in Signals and Systems - A Survey", *Automatica*, Vol24, n°3, pp309-326, 1988
- [7] S. Jawhari, "Sur la Commande Stochastique Adaptative Multimodèles", Thèse de 3ième cycle, INPG, 1978