



Introduction au filtrage particulaire

Exemples d'application

Christian Musso (DTIM/EVS)
christian.musso@onera.fr



retour sur innovation

Plan

- Introduction
- Filtre bayésien et particulaire
- Algorithmes
- Regularisation
- Point dur
- Borne de Cramer-Rao a posteriori
- Applications

Principe

Filtrage particulaire : calcul numérique approché du filtre Bayésien, estimation de l'état caché dans un système dynamique bruité, modélisé par une chaîne de Markov à l'aide d'un système de particules en interaction.

Les particules explorent l'espace d'état de manière indépendante selon le système dynamique (noyau markovien) les particules interagissent sous l'effet d'un mécanisme de sélection de pondération, de correction en fonction de l'adéquation de chaque particule avec la mesure (vraisemblance)

Historique des méthodes de Monte Carlo en filtrage

- filtre pondéré séquentiel (Handschin et Mayne, 1969)
- filtre particulaire (Del Moral, Rigal et Salut, 1992)
- bootstrap filter, SIR (Gordon, Salmond et Smith, 1993)
- Monte Carlo filter (Kitagawa, 1996)
- condensation (Blake, Isard, 1996)

Améliorations des algorithmes du FP

- Linéarisation locale
 - Doucet (96) pour l'EKF, Doucet (00) pour l'UKF
- Auxilliary particle filter
 - Pitt & Shephard (99)
- Kernel smoothing (regularisation)
 - Hurzeler & Kunsch (98), Musso & Le Gland & Oudjane (00), Liu & West (00), Pham & Dahia (03)
- MCMC
 - Gordon & Whitby (99), Gilks & Berzuini (00), Andrieu (00)
-

Pourquoi le FP

Avantages du FP

- facilité de mise en œuvre
- prise en compte de modèles complexes (hybrides avec contraintes)
- méthode non paramétrique, estimation de la densité conditionnelle complète. Cadre non-linéaire, non gaussien

Applications

- localisation, navigation, poursuite
- vidéo
- robotique mobile
- intégration longue, track before detect (radar, GPS)

But du FP

- Equation de dynamique

$$x_k = F(x_{k-1}) + v_k$$

- Equation de mesure

$$y_k = H(x_k) + w_k$$

Hypothèse markovienne

Vraisemblance

$$y_k / x_k \sim q(y_k - H(x_k))$$

Estimer la loi conditionnelle $p_{k/k}$, loi de : $x_k / y^k = (y_1, \dots, y_k)$

Filtre bayésien optimal

- prédiction

$$p(x_k / y^{k-1}) = \int p(x_k / x_{k-1}) p(x_{k-1} / y^{k-1}) dx_{k-1}$$

- correction

$$p(x_k / y^k) = \frac{p(y_k / x_k) p(x_k / y^{k-1})}{p(y_k / y^{k-1})}$$

où :
$$p(y_k / y^{k-1}) = \int p(y_k / x_k) p(x_k / y^{k-1}) dx_k$$

Estimation par échantillonnage

Prédiction suivant la dynamique $(x_{k/k-1}^{(1)}, \dots, x_{k/k-1}^{(N)})$

Correction suivant la mesure (poids) $w_j = p(y_k / x_{k/k-1}^{(j)}) / \sum_{j=1}^N p(y_k / x_{k/k-1}^{(j)})$

$$\hat{p}_{k/k}(x / y^k) = \sum_{j=1}^N w_j \delta(x = x_{k/k-1}^{(j)})$$

Estimation par échantillonnage

Prédiction

$$p_{k/k-1}(x) = \int_{R^d} p_{w_{k-1}}(x - F_k(u)) p_{k-1/k-1}(u) du$$

Correction

$$p_{k/k}(x) = \frac{\int_{R^d} p_{v_k}(y_k - H_k(x)) p_{k/k-1}(x) p_{v_k}(y_k - H_k(u)) p_{k/k-1}(u) du}{\int_{R^d} p_{v_k}(y_k - H_k(u)) p_{k/k-1}(u) du}$$

$(w_{k/k-1}^i, x_{k-1/k-1}^i)$ donné

$$x_{k/k-1}^i = F_k(x_{k-1/k-1}^i) + w_{k-1}^i \longrightarrow w_{k/k}^i \propto w_{k/k-1}^i p_{v_k}(y_k - H_k(x_{k/k-1}^i)) \quad (i = 1, \dots, N)$$

$$\hat{p}_{k/k}^N(x) = \sum_{i=1}^N w_{k/k}^i \delta(x = x_{k/k-1}^i)$$

Estimation :

$$\hat{x}_{k/k} = \sum_{i=1}^N w_{k/k} x_{k/k-1}^{(i)}$$

Problème de dégénérescence

Les poids se multiplient et au bout d'un certain temps une seule trajectoire prend tout le poids

Ré-échantillonnage : lorsque les poids sont trop concentrés on redistribue les particules en dupliquant celles de poids grand et en éliminant celles de faible poids

Illustration du FP

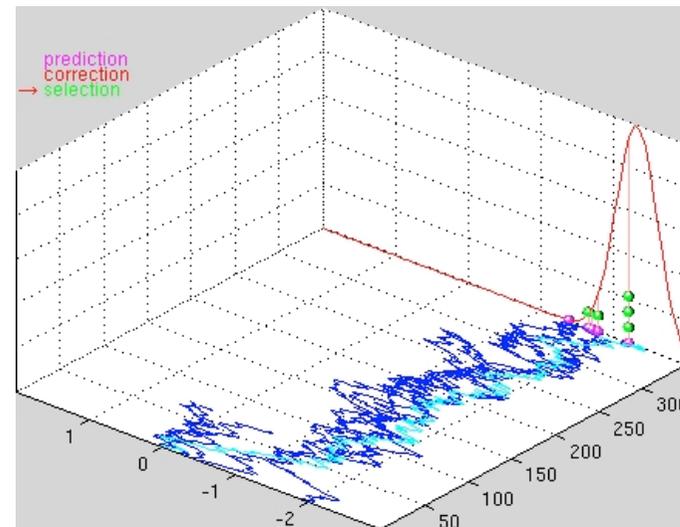
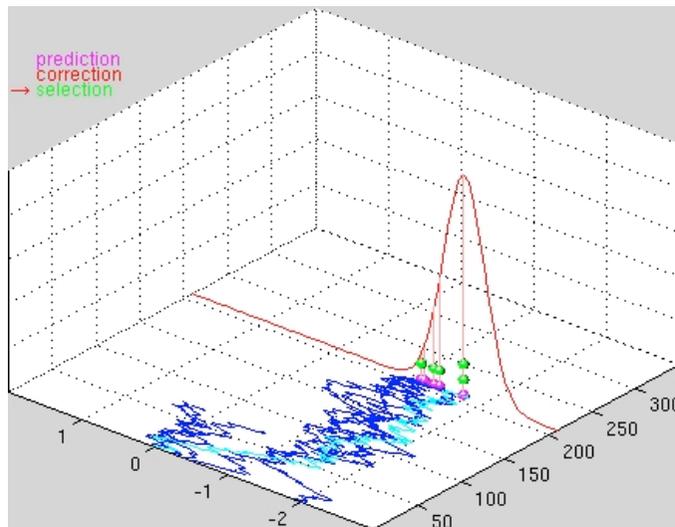
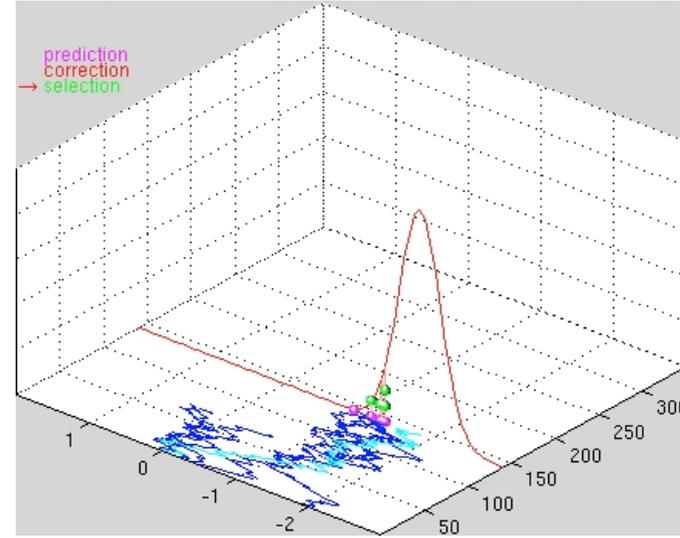
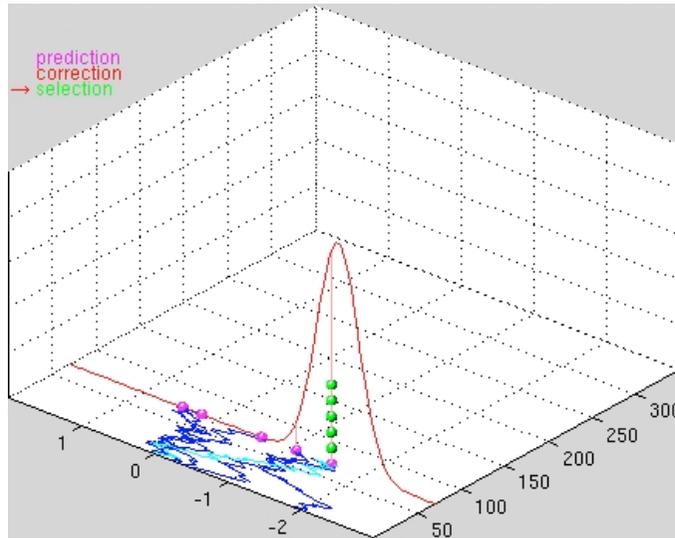
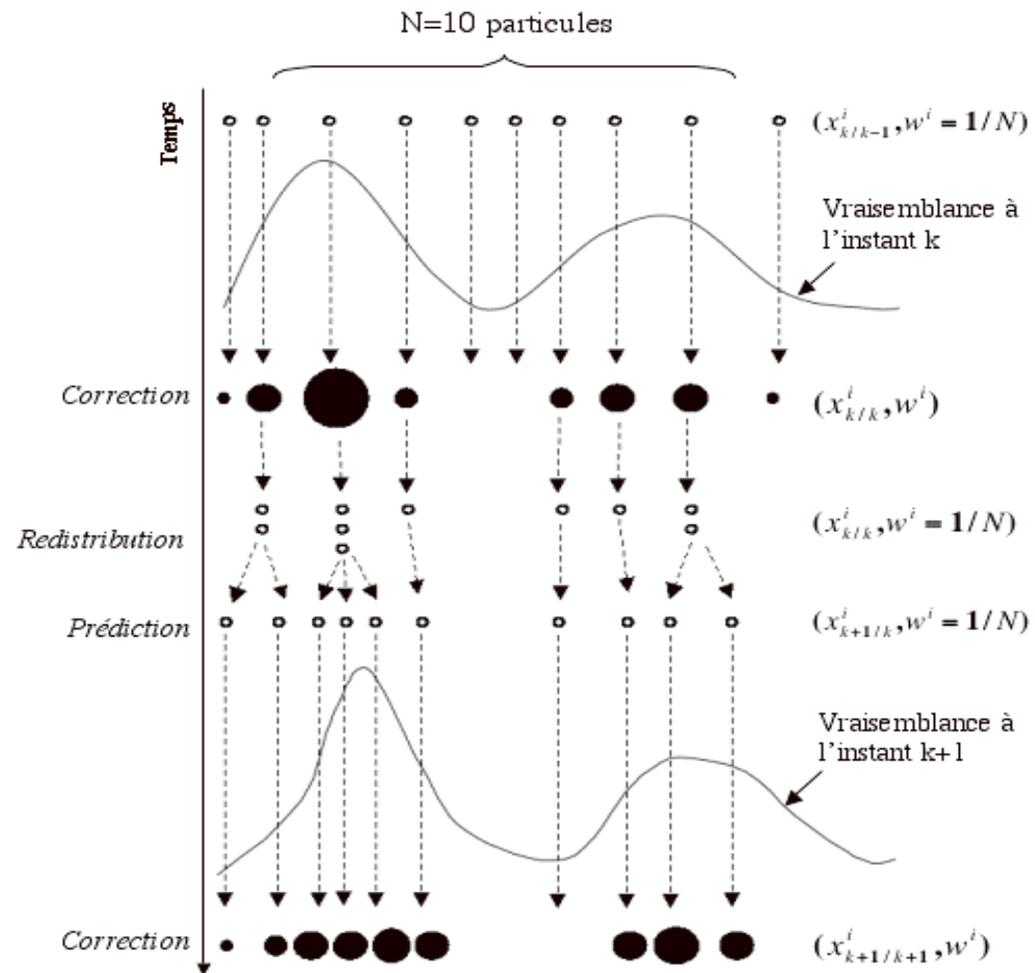


Illustration du FP



Appauvrissement du système de particules

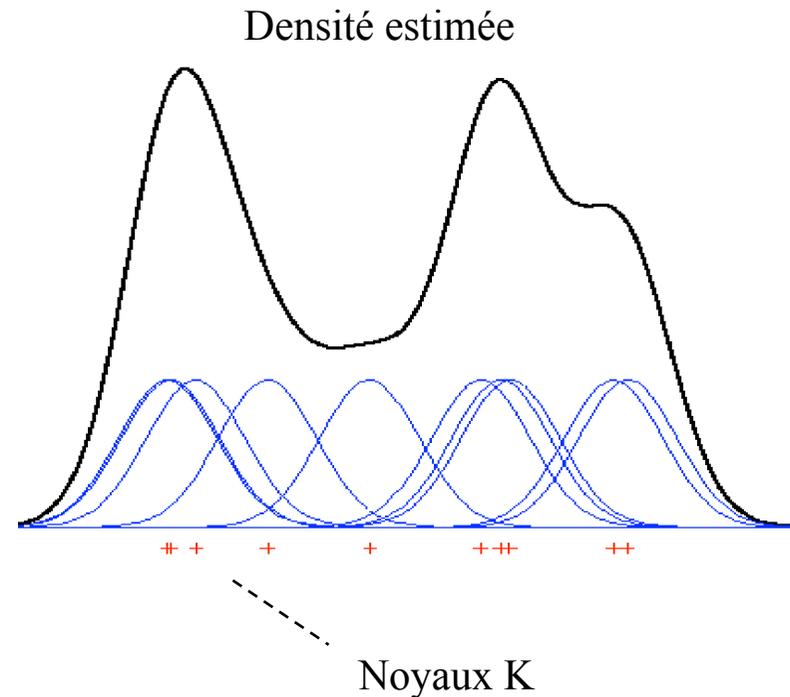
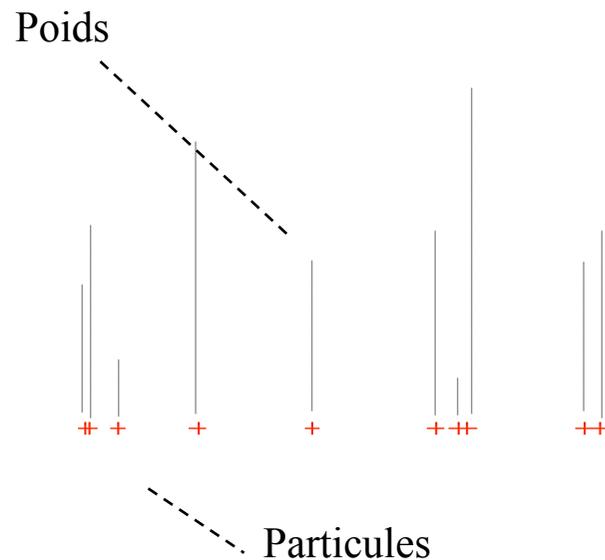
Lors des ré-échantillonnages successifs, la duplication entraîne un appauvrissement du système de particules, en particulier si le bruit de dynamique est faible

Une solution : le lissage de la densité empirique (régularisation)

Le filtre régularisé : principe

$$\hat{p}_{k/k}(x) = \sum_{j=1}^N w_j \delta(x = x^{(j)})$$

$$\hat{p}_{k/k}(x) = \frac{1}{h^d} \sum_{j=1}^N w_j K(A^{-1}h^{-1}(x - x^{(j)}))$$



Le filtre régularisé : fonctionnement

Sample from $\hat{p}_{k/k}(x) = \frac{1}{h^d} \sum_{j=1}^N w_j K(A^{-1}h^{-1}(x - x^{(j)}))$

$$\left\{ \begin{array}{ll} P(J = j) = w_j & \text{(Multinomial)} \\ Z \sim K(x)dx & \text{(Epanechnikov)} \\ x = x_{k/k-1}^{(J)} + hA_k Z & \text{(Regularization)} \end{array} \right.$$

Le filtre régularisé : algorithme

$$\left\{ \begin{array}{l} K(x) = \frac{1}{2} c_d^{-1} (d+2) (1 - \|x\|^2) \delta(\|x\| \leq 1) \\ h = \mu [(8(d+4)(2\sqrt{\pi})^d / c_d) / N]^{1/(d+4)} \quad (0.2 \leq \mu \leq 0.6) \\ A_k A_k^T = \Omega_e \quad \text{où} \quad \Omega_e = \frac{1}{N} \sum_{j=1}^N (x_{k/k-1}^{(j)} - \bar{x}_{k/k-1}^{(j)}) (x_{k/k-1}^{(j)} - \bar{x}_{k/k-1}^{(j)})^T \end{array} \right.$$

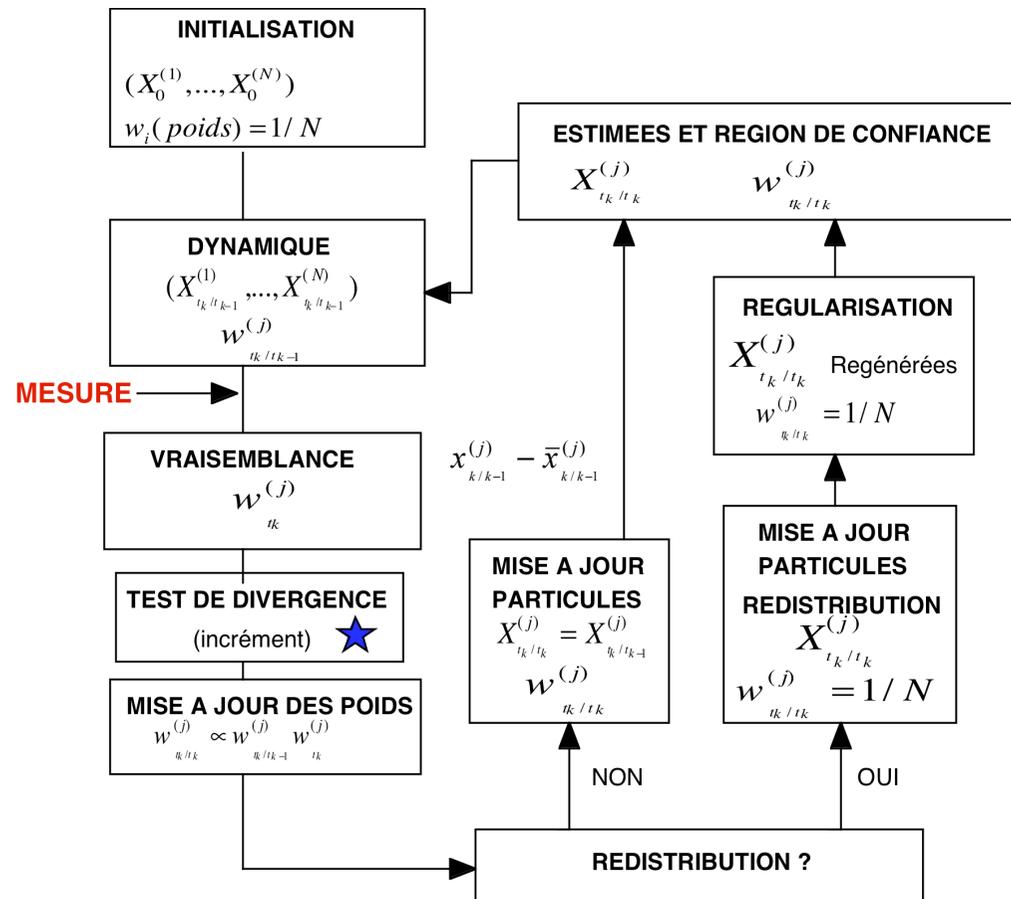
$$\hat{p}_{k/k}(x) = \frac{1}{h^d} \sum_{j=1}^N w_j K(A^{-1} h^{-1} (x - x^{(j)}))$$

$$C_{ent} = \log(N) + \sum_{i=1}^N w_i \log(w_i) > \text{threshold}$$

Entropy

Resampling

Le filtre régularisé : algorithme

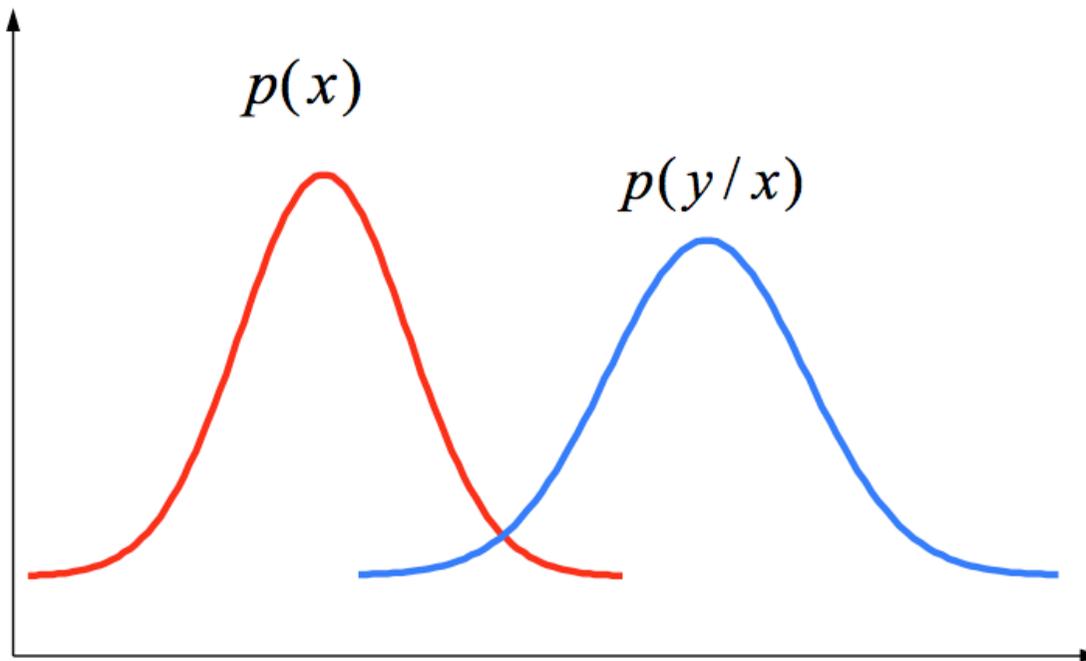


★ Si le nombre d'incrément de divergence > seuil alors retour en INITIALISATION

FP : le point dur

Incohérence entre la densité a priori et la vraisemblance

Beaucoup de particules ont un poids nul



FP : le point dur

Solutions :

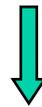
- choix d'une bonne fonction d'importance. La FI n'est plus la dynamique mais une fonction tenant compte de la mesure courante
- correction progressive : si la vraisemblance est pointue, on corrige progressivement à partir d'une suite d'écart-types de plus en plus petits

FP : résultats théoriques sur l'erreur

Erreur sur le filtre approché au sens faible

$$\sup_{\|\phi\|=1} E[|\langle p_{k/k} - \hat{p}_{k/k}^N, \phi \rangle| / y^k] \leq 2 \frac{\bar{\gamma}_k}{\sqrt{N}}$$

où $\bar{\gamma}_k = \prod_{i=1}^k \gamma_i$; $\gamma_i = \frac{\sup_x p(y_i/x)}{\int p(y_i/x) p_{i/i-1}(x) dx}$



l'intégrale quantifie l'incohérence

En particulier, l'erreur augmente avec la dimension de l'état (thèse en cours à l'Onera de Paul Bui Quang)

Borne de Cramer-Rao a posteriori

Intérêt :

Evaluation de la performance du filtre, prédiction de performance (hors ligne), zone de confiance

$$J_{ij} = E_{X,Y} \left[- \frac{\partial^2 \log p_{X,Y}(X,Y)}{\partial X_i \partial X_j} \right] \quad \text{Information matrix (Fisher)}$$

Matrice de covariance

$$E\{(g(Y) - X)(g(Y) - X)^T\} \geq J^{-1} = \text{PCRB}$$

Estimateur de X

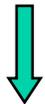
Borne de Cramer-Rao a posteriori

Forme récursive, Tichavsky (98)

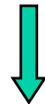
Cas d'une dynamique linéaire

$$\begin{cases} X_{k+1} = F_k X_k + W_k \\ y_k = H_k(X_k) + V_k \end{cases}$$

$$J_{k+1} = E_{X_{k+1}} \{ [\nabla_{X_{k+1}} H_k^T(X_{k+1})] R_k^{-1} [\nabla_{X_{k+1}} H_k^T(X_{k+1})]^T \} + (F_k J_k^{-1} F_k^T + S_k)^{-1}$$



MC évaluation



Gain (variation de h)



Perte (dynamique)

- Generalisation de l'équation de Riccati
- formulation informationnelle du Kalman si h linéaire

Recalage par corrélation de terrain : altimétrie

- **Dynamique**

$$X_t = F(X_{t-1}) + v_t$$

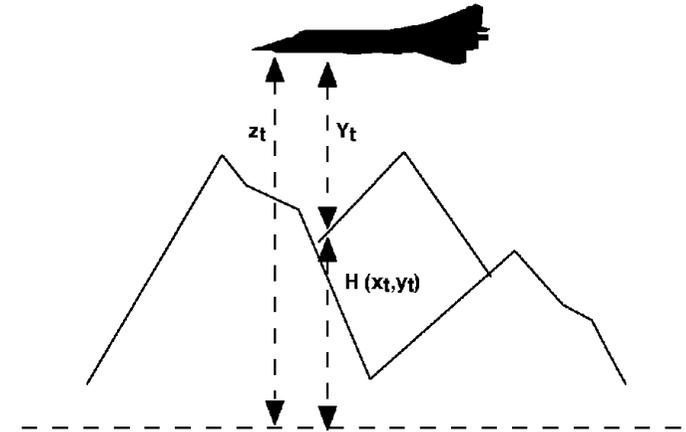
Echantillon prédit (particules bruitées) :

$$(X_{t/t-1}^{(1)}, \dots, X_{t/t-1}^{(N)}) = (F(X_{t-1/t-1}^{(1)}), \dots, F(X_{t-1/t-1}^{(N)}))$$

- **Mesure** (ex : altimétrie, gravimétrie)

$$y_t = z_t - MNT(x_t, y_t) + \eta_t = H(X_t) + \eta_t$$

Poids des particules $w_i \propto q_{\eta_t}(z_t - H(X_{t/t-1}^{(i)}))$



Estimée de l'état (positions, vitesses) :

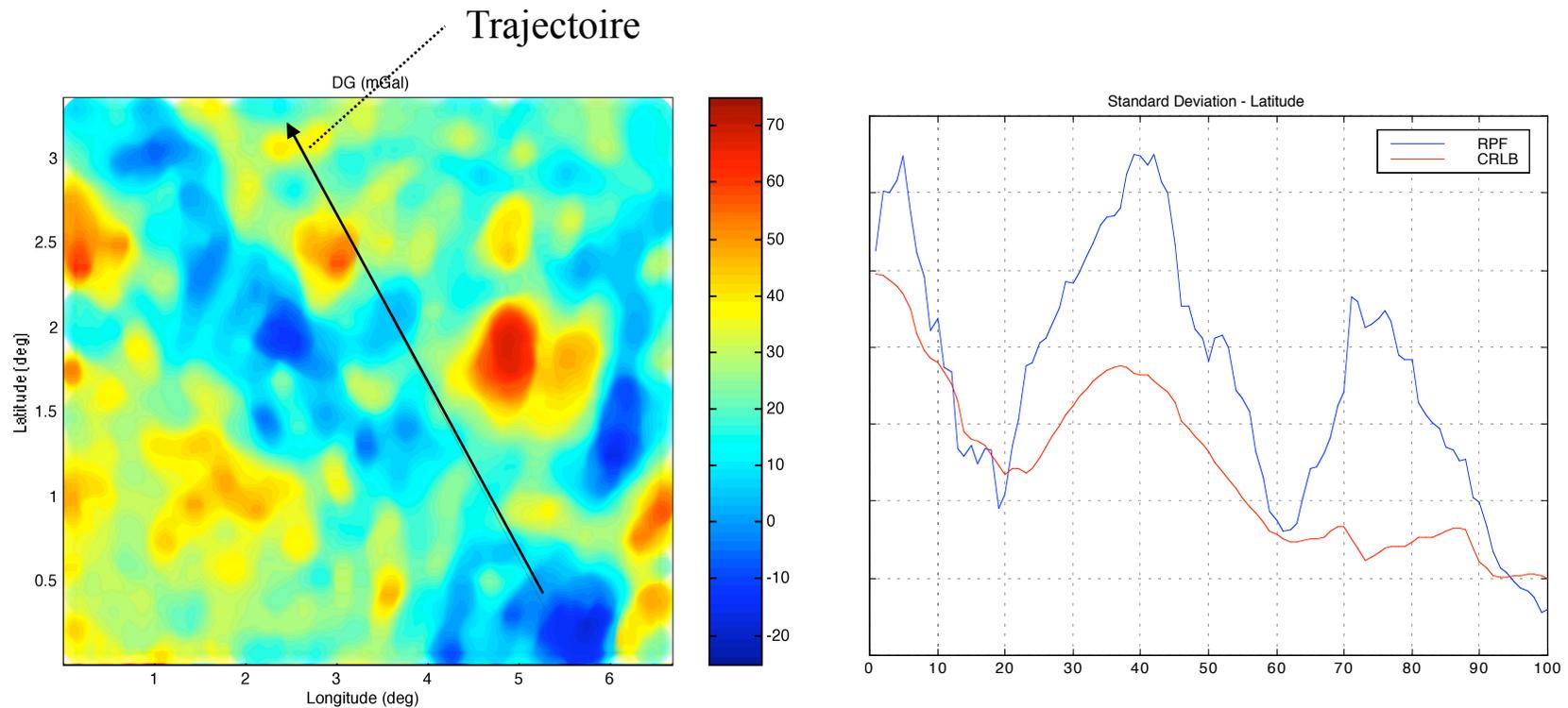
$$\hat{X}_{t/t} = \sum_{i=1}^N w_i X_{t/t-1}^{(i)}$$

Recalage par corrélation de terrain : gravimétrie

Erreur du filtre : comparée à l'erreur minimale

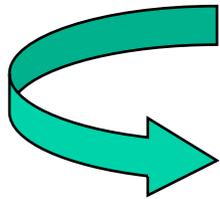
Vecteur d'état : dérive inertielle (dim 7)

Mesure : anomalie gravimétrique



FP pour le GPS : perturbations

- Biais atmosphériques
- Biais et dérives d'horloge
- Multi-trajets
- Brouillage



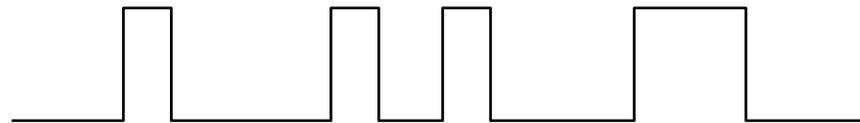
Filtre robuste : intégration longue par filtrage particulière

Méthode classique : recherche exhaustive dans le plan (retard/Doppler) pour la détection

Insuffisant

FP pour le GPS ; modélisation

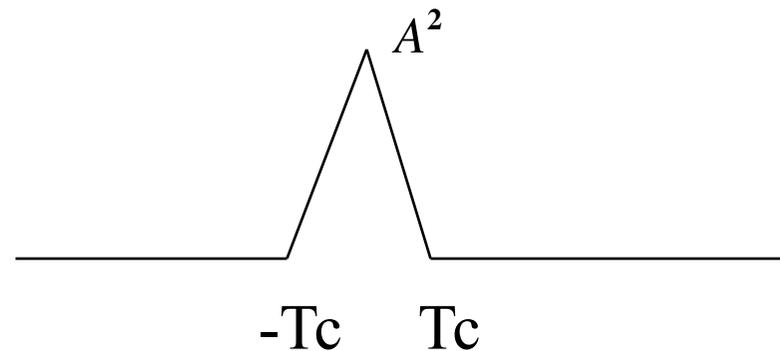
Code C/A



$$T_c = 977 \text{ nsec (300m)}$$

Période : 1ms (300 km) • Bande = 2 Mhz

Autocorrélation



FP pour le GPS : modélisation

Emission :

$$s_e(t) = A [C/A](t) \exp[2i\pi f_{L_1} t] + \varepsilon(t)$$

$$A = |A| D e^{i\varphi_0} \quad (D : \text{bit de nav})$$

Réception :

$$y(t) = A [C/A](t - \tau_0 - t\dot{\tau}_0) \exp[2i\pi\{f_1 t - f_{L_1} t \dot{\tau}_0 - \frac{1}{2} f_{L_1} t^2 \ddot{\tau}_0\}] + \varepsilon(t) = A g(t, \theta) + \varepsilon(t)$$

$$\theta = [\tau_0, \dot{\tau}_0, \ddot{\tau}_0] : \text{retards et dérivées (Doppler)}$$

FP pour le GPS : modélisation

Contexte : poursuite mono-satellite (C/A code du satellite poursuivi)

$$y(t) = A [C/A](t - \tau_0 - t\dot{\tau}_0) \exp[2i\pi\{f_L t - f_{L_1} t \dot{\tau}_0 - \frac{1}{2} f_{L_1} t^2 \ddot{\tau}_0\}] + \varepsilon(t) = A g(t, \theta) + \varepsilon(t)$$

$$Y_{1:M}(t) = [y(t_1), \dots, y(t_M)]^T$$

Vraisemblance

$$\log\left(\prod_{t=t_1}^{t_M} P(y(t) | \theta, A)\right) \approx \frac{1}{M} \left| \sum_{t=t_1}^{t_M} y(t) \bar{g}(t, \theta) \right|^2 - \left[\left| AM - \sum_{t=t_1}^{t_M} y(t) \bar{g}(t, \theta) \right|^2 \right]$$

↑
Corrélation

FP pour le GPS : modélisation

$\theta = [\tau_0, \dot{\tau}_0, \ddot{\tau}_0]$: retards et dérivées (Doppler)

Dépend des positions/vitesses du récepteur et des éphémérides du satellite

Vecteur d'état : $X_n = [x_n, y_n, z_n, vx, vy, vz]^T$

$$X_{n+1} = F_n X_n + U_{n,n+1}$$

FP pour le GPS : brouillage

Sans brouillage



Avec brouillage



Corrélation réplique/code pour une particule

FP pour le GPS : modélisation

$$y(t) = A [C/A](t - \tau_0 - t\dot{\tau}_0) \exp[2i\pi\{f_L t - f_{L_1} t \dot{\tau}_0 - \frac{1}{2} f_{L_1} t^2 \ddot{\tau}_0\}] + \varepsilon(t)$$

$$\tau_0 = \frac{1}{c} \|Z\| \quad ; \quad \dot{\tau}_0 = \frac{\langle \dot{Z}, Z \rangle}{c \|Z\|} \quad \ddot{\tau}_0 = \frac{\|Z\|^2 (\langle \ddot{Z}, Z \rangle + \|\dot{Z}\|^2) - \langle \dot{Z}, Z \rangle^2}{c \|Z\|^3}$$

$$\theta = [\tau_0, \dot{\tau}_0, \ddot{\tau}_0]^T = F(Z) = F(X, S, \dot{S}, \ddot{S}) \quad \text{fonction de l'état et des éphémérides}$$

|
/

Position relative Récep./Sat
Kepler

Pour chaque particule X^i on calcule θ^i et la corrélation associée

Inclure la dynamique dans la corrélation
(pistage avant détection, track before detect)

FP pour le GPS : modélisation

$$w_k^{(i)} \propto \exp \left\{ -\frac{1}{2\sigma_a^2(k)} \left| \sum_{t=t_1}^{t_M} C/A(t, \theta^{(i)}(k)) \bar{g}[t, \theta^{(i)}(k)] y_k(t) \right|^2 \right\}$$

$$X_k^{(i)} = [x_k^{(i)}, y_k^{(i)}, z_k^{(i)}, vx^{(i)}, vy^{(i)}, vz^{(i)}]^T \quad \theta^{(i)} = F(X^{(i)}, S, \dot{S}, \ddot{S})$$

Résultats du FP : gain de 20 dB

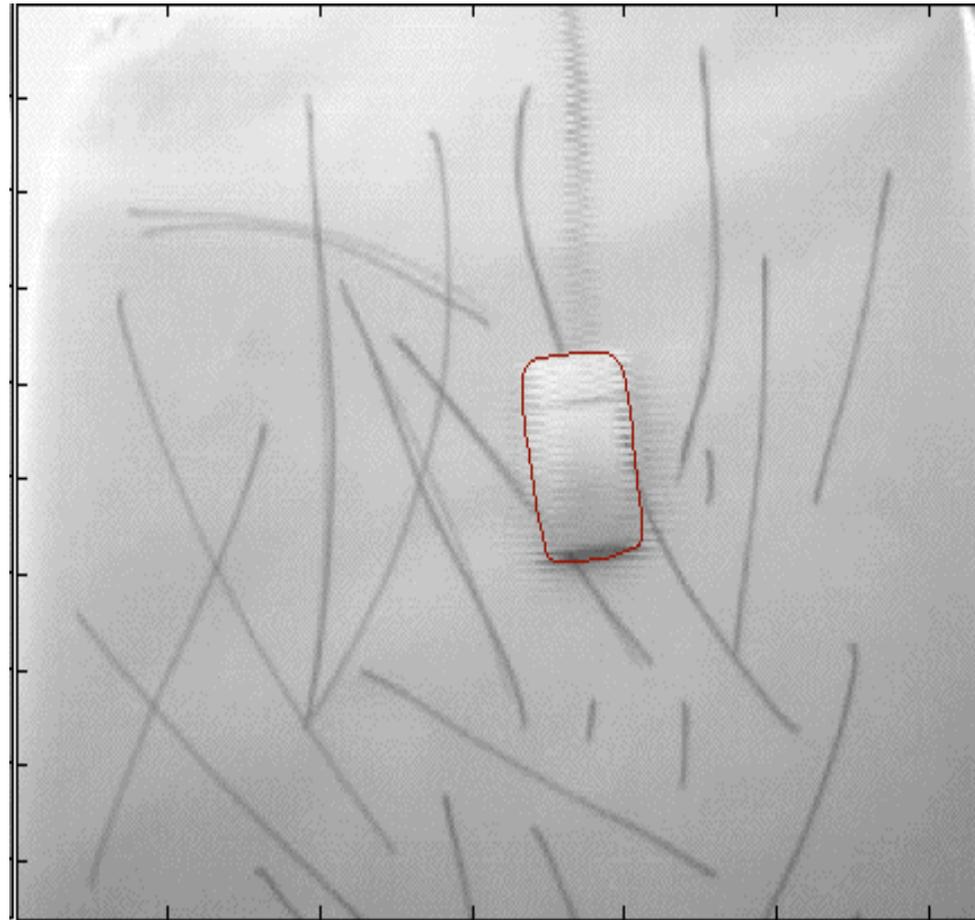
Suivi de contour dans une séquence d'images par FP



Suivi de contour dans une séquence d'images par FP



Suivi de contour dans une séquence d'images par FP



Suivi de contour dans une séquence d'images par FP

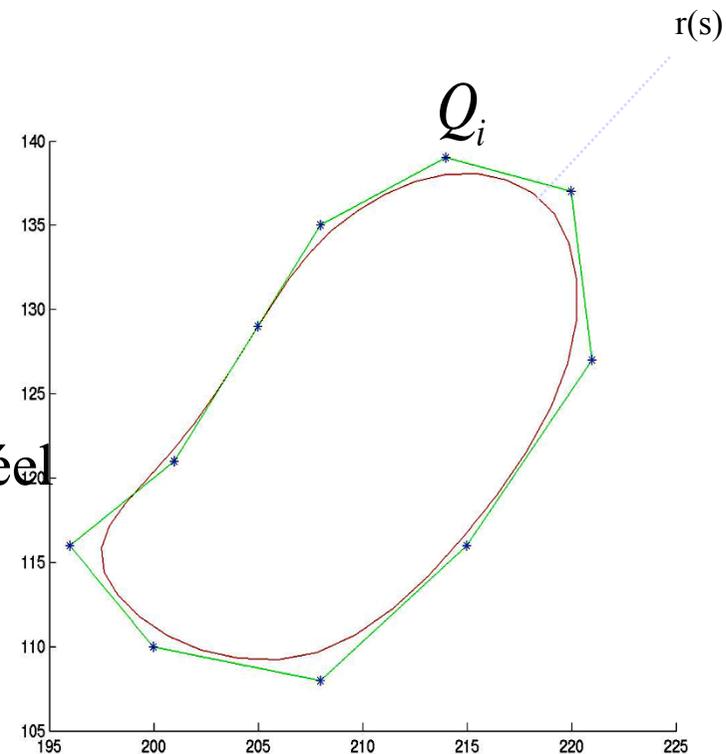
Contour apparent : - courbe B-spline (quadratique) périodique fermée

$$\forall s \in [0, L], r(s) = \begin{pmatrix} x(s) \\ y(s) \end{pmatrix} = \sum_{i=0}^{L-1} B_i(s) \cdot Q_i$$

$$\forall i > 0, B_i(s) = B_0(s-i) \text{ et } \sum_{i=0}^{L-1} B_i(s) = 1$$

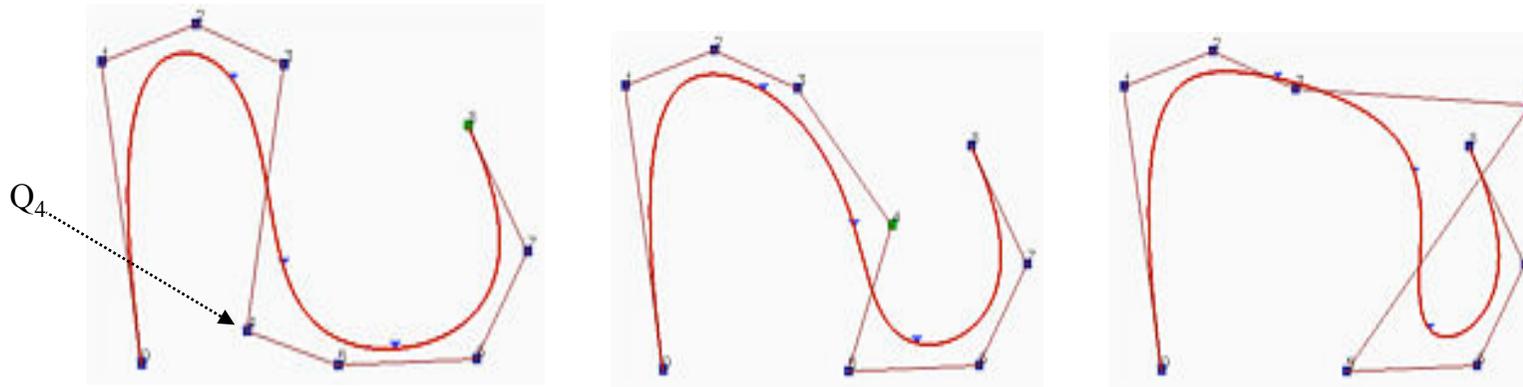
La courbe $r(s)$ approche le contour réel de l'objet semi-rigide suivi.

Points de contrôle : Q_i



Suivi de contour dans une séquence d'images par FP

- Déformation locale par translation d'un point de contrôle



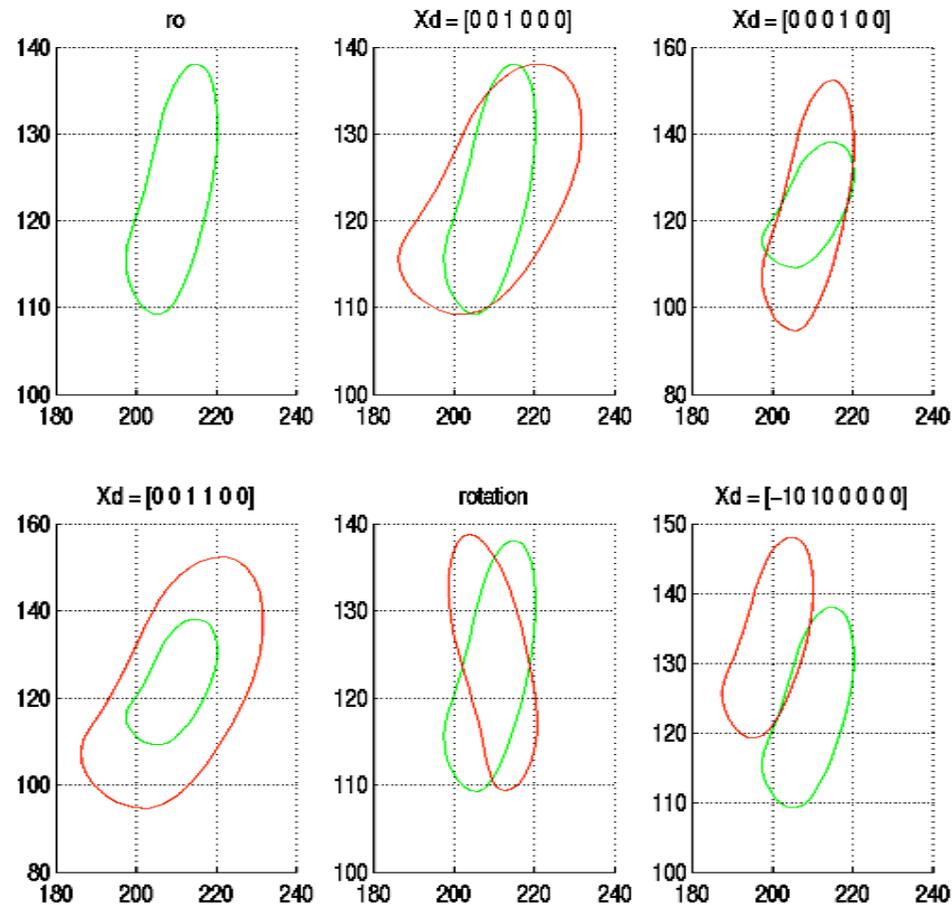
- Restriction des transformations : Affinités

$$Q_t = \begin{bmatrix} Q_t^x \\ Q_t^y \end{bmatrix} = W_t X_t + \begin{bmatrix} Q_{t-1}^x \\ Q_{t-1}^y \end{bmatrix}$$

X_t état de dimension 6 :
déplacement et déformation

Trans. aff. Q_i \longleftrightarrow Trans. aff. Courbe $r(s)$

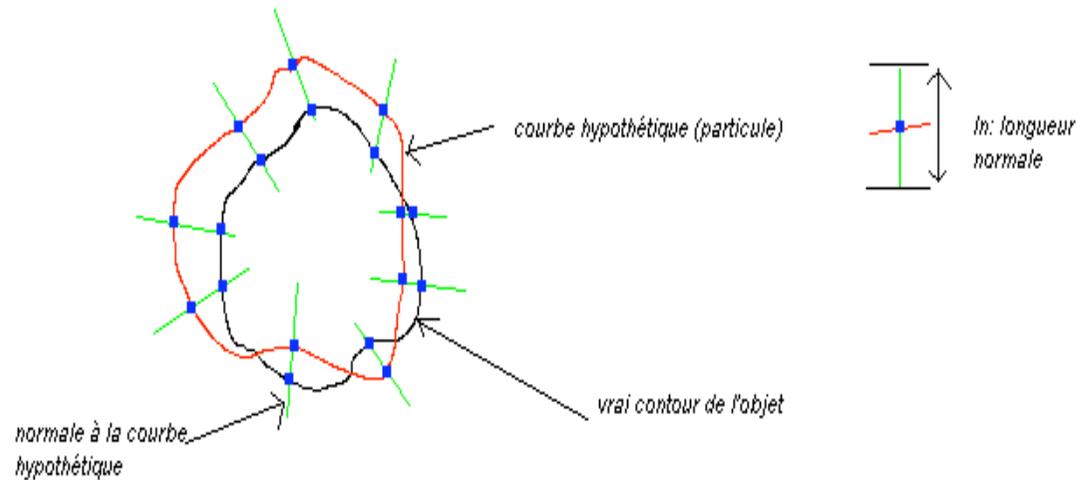
Suivi de contour dans une séquence d'images par FP



Suivi de contour dans une séquence d'images par FP

$X_{t/t-1}$ Particule - Courbe prédite
Contour réel - Mesure

} Distance entre 2 courbes - vraisemblance



$$d^2(r_1, r_2) = \frac{1}{L} \int (\langle (r_1(s) - r_2(s))^T, n_1(s) \rangle)^2 ds$$

Suivi de contour dans une séquence d'images par FP



Détection - Filtre de Canny

Points de contraste maximum sur la normale

Fausses alarmes

Robustesse : nouvelle formulation de la vraisemblance

Suivi de contour dans une séquence d'images par FP

Modélisation des FA par PDAF

$$P(Z/X) = \prod_{j=1}^M \left[\lambda^{-1} (1 - P_D) P_f(n(j)) + n(j)^{-1} P_D P_f(n(j) - 1) \frac{1}{\sqrt{2\pi\sigma}} \sum_{i=1}^{n(j)} \exp\left(-\frac{\|r_j - z_i\|^2}{2\sigma^2}\right) \right]$$

l : Longueur de la normale

P_D : Probabilité de détection du maxima de contraste

P_f : Probabilité d'avoir une fausse alarme

$n(j)$ = Nb de maxima locaux sur la normale

r_j : Point d'échantillonnage de la courbe prédite

z_i : Maximum local

Multimodalité, non-linéarité

Conclusion - Questions

- Le FP s'emploie quand on est sûr que le Kalman est inefficace
- Adapter l'algorithme pour chaque application