DIRO
IFT 6150

TRAITEMENT D'IMAGES

ANALYSE DU MOUVEMENT

Max Mignotte

Département d'Informatique et de Recherche Opérationnelle.
Http : //www.iro.umontreal.ca/∼mignotte/ift6150
E-mail : mignotte@iro.umontreal.ca

ANALYSE DU MOUVEMENT

SOMMAIRE

ANALYSE DU MOUVEMENT
PRÉSENTATION

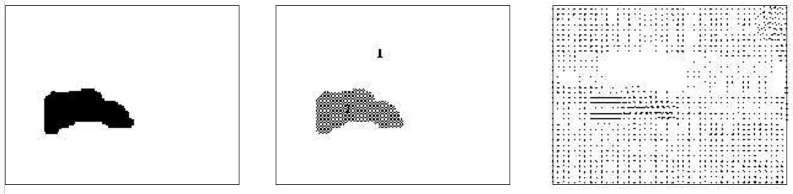
▶ Recherche active à l'heure actuelle en raison de son importance dans de nombreuses applications :

Télé-surveillance, compression pour les télécommunications ou l'archivage, diagnostique médical, météorologie, robotique mobile, . . .

Trois tâches

1. *La détection* (de zones mobiles dans le plan image) (cf. Fig. b)
2. *La segmentation* (en régions cohérentes au sens du mouvement) (cf. Fig. c)
3. *L'estimation* (des vecteurs vitesses, en chaque pixel ou pour chaque objet) (cf. Fig. d)

(a)



(b) (c) (d)

ANALYSE DU MOUVEMENT
DÉTECTION DU MOUVEMENT (1)

▶ Le but de la *détection* du mouvement est d'associer à chaque pixel une étiquette (ou classe) binaire indiquant l'appartenance ou la non-appartenance du point considéré à une zone en mouvement

En faisant l'hypothèse d'une caméra fixe + éclairement quasi-constant de la scène observée, on peut extraire une information portant sur la variation temporelle de l'intensité lumineuse I du pixel :

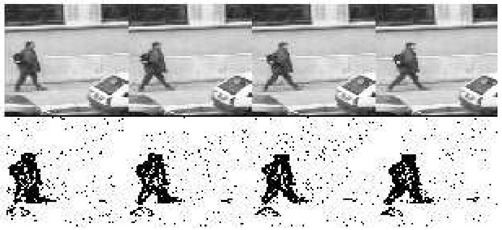
$$\boxed{o_s = |I(x,y,t) - I(x,y,t-1)|}$$

Dans le cas idéal :

- $o_s = 0$ pour un pixel du fond fixe
- $o_s \neq 0$ s'il y a eu mouvement d'un objet d'intensité non uniforme

▶ Cependant cette observation est peu robuste et sensible au bruit d'acquisition, légère variations d'éclairement, . . . Un simple seuillage ne permet pas d'extraire les zones mobiles

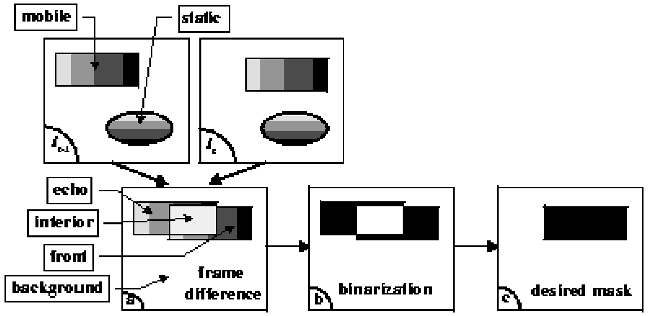
Exemples



ANALYSE DU MOUVEMENT
DÉTECTION DU MOUVEMENT (2)

On distingue typiquement quatre zones dans le champ d'observation

1. Le fond fixe
2. La zone d'écho (zone de fond découverte par l'objet à l'instant t)
3. La zone de glissement de l'objet sur lui-même
4. La zone de recouvrement (zone de fond recouverte par l'objet à l'instant t)



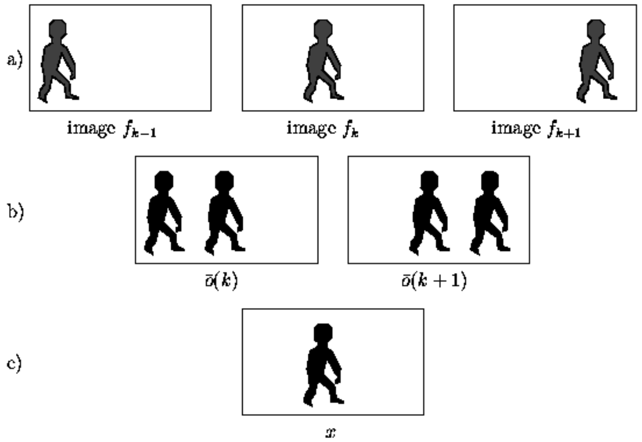
Il faut d'une part éliminer l'écho, d'autre part reconstruire l'intérieur des objets mobiles (zone de glissement)

Techniques simples mais peu robustes

- La technique de simple différence par rapport à une image censée ne contenir que le fond fixe de la scène
- La technique du ET logique entre cartes binaires de chgt temporels (qui ne donne de bon résultats que si le mouvement est assez rapide pour qu'il n'y ait pas de recouvrement entre deux positions successives de l'objet)

ANALYSE DU MOUVEMENT
DÉTECTION DU MOUVEMENT (3)

Exemple de technique du ET logique

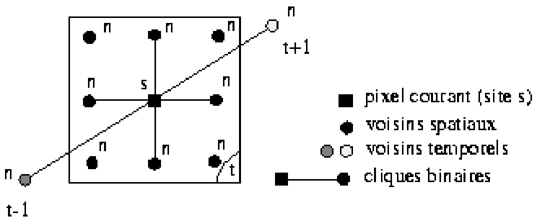


▶ C'est pourquoi on fait appel à une approche statistique probabiliste pour régulariser le problème mal posé. Le principe consiste à introduire une modélisation *a priori* du champ d'étiquettes, à l'aide de de la théorie des champs aléatoires de Markov

▶ On contraint alors la solution vers la configuration la plus probable du champ des étiquettes étant donné les observations et le modèle *a priori*

Notation

S	L'image à l'instant courant t
s	Un pixel (*site*) quelconque de S
η_s	Un voisinage spatio-temporel de s
n	N'importe quel voisin de s (spatial ou temp.)
$\mathcal{C}$	L'ensemble des cliques binaires $c = (s, n)$ constituant les voisinages η_s
$O = \{O_s, s \in S\}$.	Le champ aléatoire des observations O
$L = \{L_s, s \in S\}$ ..	Le champ aléatoire des étiquettes
$o = \{o_s, s \in S\}$ ..	Une réalisation particulière du champ d'observations O à l'instant t
$l = \{l_s, s \in S\}$ ...	Une réalisation particulière du champ d'étiquettes L à l'instant t
Ω	L'ensemble des configurations possibles l du champ aléatoire L



▶ Les intéractions spatiales et temporelles entre étiquettes sont modélisées par un MRF qui constitue le modèle *a priori* du champ des étiquettes

$$\begin{aligned} \forall\, l \in \Omega \qquad P_L(l) &> 0 \\ P_{L_s\,|\,L_n}(l_s\,|\,l_n,\ n \neq s, n \in S) &= P_{L_s\,|\,L_n}(l_s\,|\,l_n,\ n \in \eta_s) \end{aligned}$$

Un MRF étant équivalent à une distribution de Gibbs, on peut exprimer la probabilité *a priori* du champ d'étiquettes à l'aide d'une fonction d'énergie

$$P_L(l) = \frac{1}{Z} \exp(-U_m(l))$$

où Z est une constante de normalisation
et U_m une fonction d'énergie associée au modèle *a priori*

Le champ d'étiquettes le plus probable relativement aux observations est obtenu par le critère du Maximum A Posteriori (MAP). A l'aide du théorème de Bayes, on montre que

$$\arg\max_l\ P_{L|O}(l|o) \ \Leftrightarrow\ \arg\min_l\ \big(U_m(l) + U_\alpha(o,l)\big)$$

$U_m(l)$ est un terme d'énergie qui assure la régularisation de la solution. Il s'exprime

$$U_m(l) = \sum_{c \in C} V_c(l_s, l_n)$$

Ces potentiels élémentaires sont définis en fonction du problème à traiter. On peut par exemple prendre

$$V_c(l_s, l_n) = \begin{cases} -\beta & \text{si} \quad l_s = l_n \\ +\beta & \text{si} \quad l_s \neq l_n \end{cases}$$

où $\beta > 0$ prend une des trois valeurs β_s, β_p, ou β_f selon la clique considérée (spatiale, passée ou future)

DÉTECTION DU MOUVEMENT
RÉGULARISATION PAR APPROCHE MARKOVIENNE

▶ Ces potentiels energétiques favorisent un étiquetage homogène puisque des étiquettes semblables sur des pixels voisins induisent une moindre contribution énergétique, donc une configuration plus favorable

En pratique, on prend $\beta_f > \beta_p$ pour éliminer les zones d'écho du mouvement

$U_a(0, l)$ est l'énergie d'adéquation qui assure une bonne attache aux données. Elle traduit le lien entre observations et étiquettes et s'exprime

$$U_a(o, l) = \frac{1}{2\sigma^2} \sum_{s \in S} [o_s - \Psi(l_s)]^2$$

où σ^2 est la variance des observations et Ψ est la fonction d'attache aux observations. Pour la détection de mouvement on peut par exemple prendre une fonction simple du type

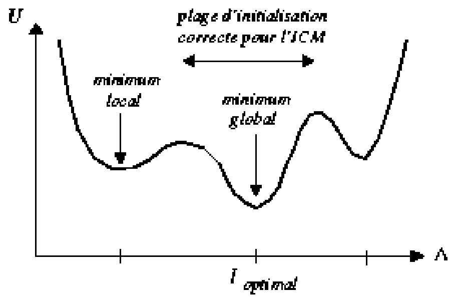
$$\Psi(l_s) = \begin{cases} 0 & \text{si} \quad l_s = b \\ \alpha > 0 & \text{sinon} \end{cases}$$

où α correspond à la valeur moyenne des observations non nulles

Algorithme de Relaxation spatiale

- Les algorithmes de relaxation stochastiques, du type Recuit Simulé ▶ convergence vers le minimum global de la fonction d'énergie mais au prix d'un coût de calcul prohibitif
- Les algorithmes de relaxation déterministes, du type ICM ▶ convergence rapide mais ne garantissant pas de converger vers le minimum global de la fonction d'énergie. ▶ Requiert une bonne configuration initiale du champ d'étiquettes

DÉTECTION DU MOUVEMENT
RÉGULARISATION PAR APPROCHE MARKOVIENNE



Algorithme ICM

| $\hat{l}^{[0]}$ Configuration initiale

(0) Initialisation : on choisit une configuration initiale $\hat{l}^{[0]}$ grossièrement estimée par binarisation du champ correspondant d'observation

(1) Calcul de $\hat{l}^{[k+1]}$ à partir de $\hat{l}^{[k]}$

$l \leftarrow \hat{l}^{[k]}$

1. On balaie l'ensemble des sites s (selon une stratégie de visite de sites) et
   - En chaque site s, on calcule $U_m(l_s) + U_\alpha(o_s, l_s)$ pour chacune des classes possibles
   - On sélectionne la classe l_s qui correspond à l'energie minimale

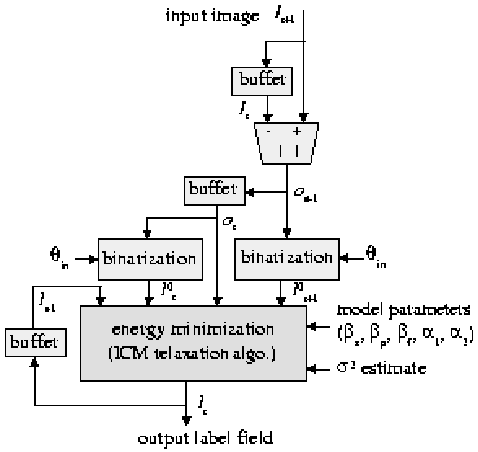
   $$\hat{l}_s = \arg\min_{l_s} \Big(U_m(l_s) + U_\alpha(o_s, l_s) \Big)$$

   $\hat{l}^{[k+1]} \leftarrow l$
2. Retour en (1) jusqu'à la réalisation d'un critère de convergence. Généralement,

   si $\quad \hat{l}_s^{[k+1]} \not\approx \hat{l}_s^{[k]} \quad k \leftarrow k+1 \quad$ et retour en (1)

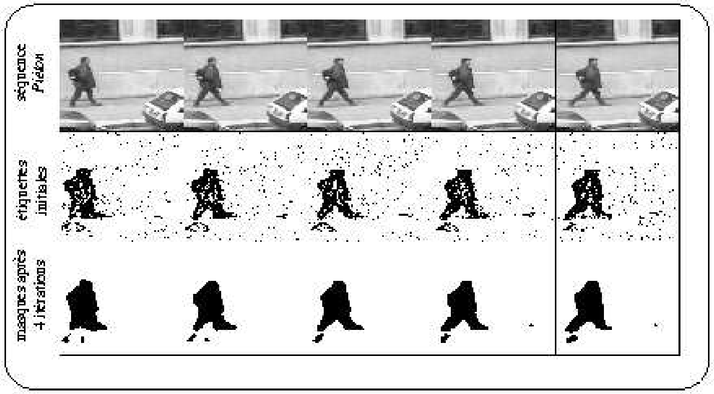
DÉTECTION DU MOUVEMENT
RÉGULARISATION PAR APPROCHE MARKOVIENNE

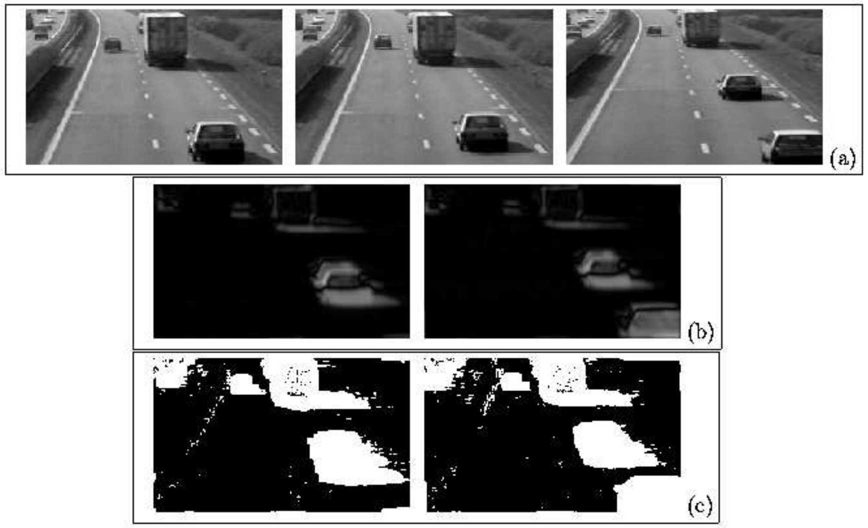
Synoptique de l'algorithme de détection du mouvement



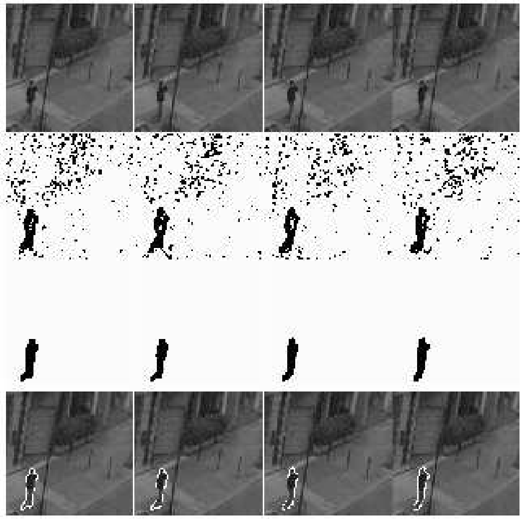
Étant donné le voisinage spatio-temporel utilisé pour détecter le mouvement, voisinage qui inclut un voisin futur et un voisin passé pour chaque pixel, le résultat de traitement est obtenu avec un retard d'une image par rapport à l'acquisition, puisqu'il faut disposer de l'image en $t+1$ pour estimer le champ d'étiquettes en t

DÉTECTION DU MOUVEMENT
EXEMPLES





DÉTECTION DU MOUVEMENT
EXEMPLES



Nb_itération	Temps cpu /image
4	5s