

# Estimation de Mouvement

**Patrick Clarysse**

**Contact: [patrick.clarysse@creatis.insa-lyon.fr](mailto:patrick.clarysse@creatis.insa-lyon.fr)**

**Document: <http://www.creatis.insa-lyon.fr/~clarysse>**

# **Estimation / Analyse de Mouvement - Sommaire -**

**1- Introduction**

**2- Définitions**

**3- Détection de Mouvement**

**4- Estimation d'un champ dense**

– **Approches différentielles**

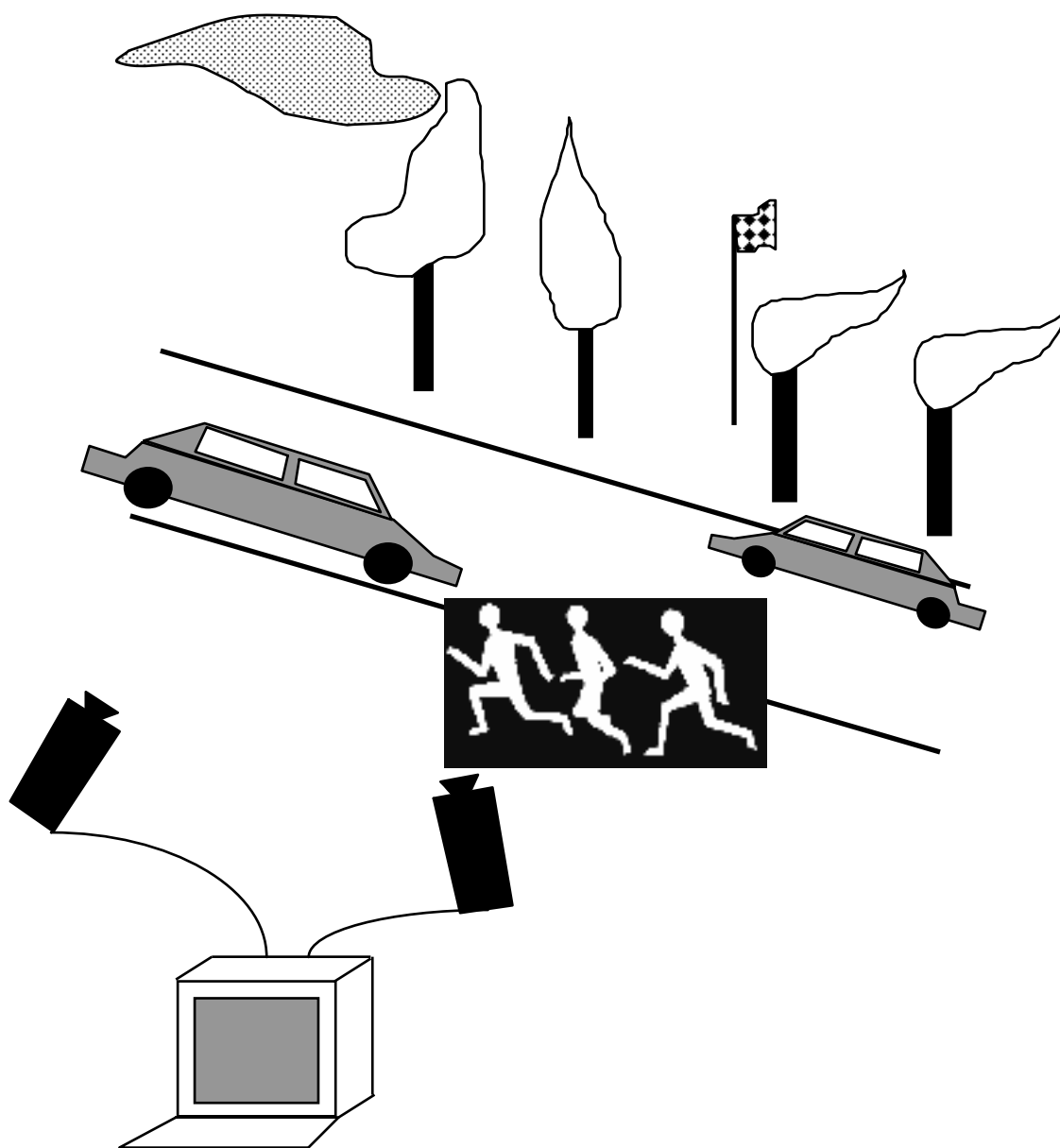
– **Mise en correspondance**

– **Techniques spatio-temporelles**

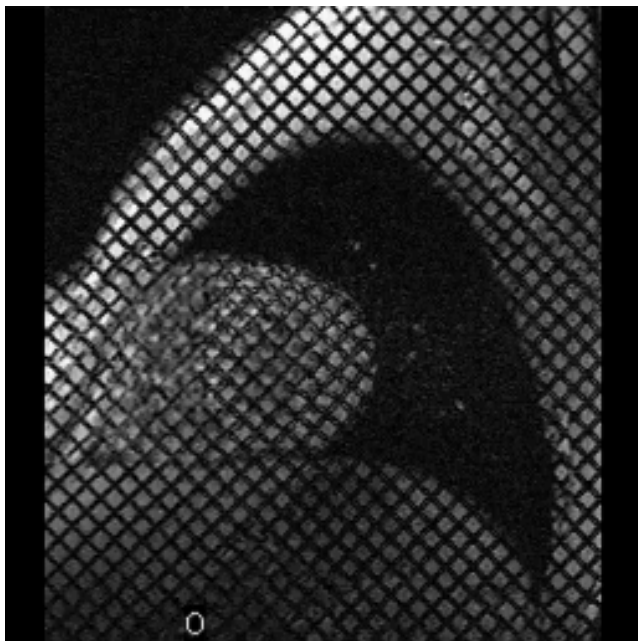
**5- Compléments**

**6- Conclusion**

# 1- Introduction



Scène 3D



## **1.1 Objectifs**

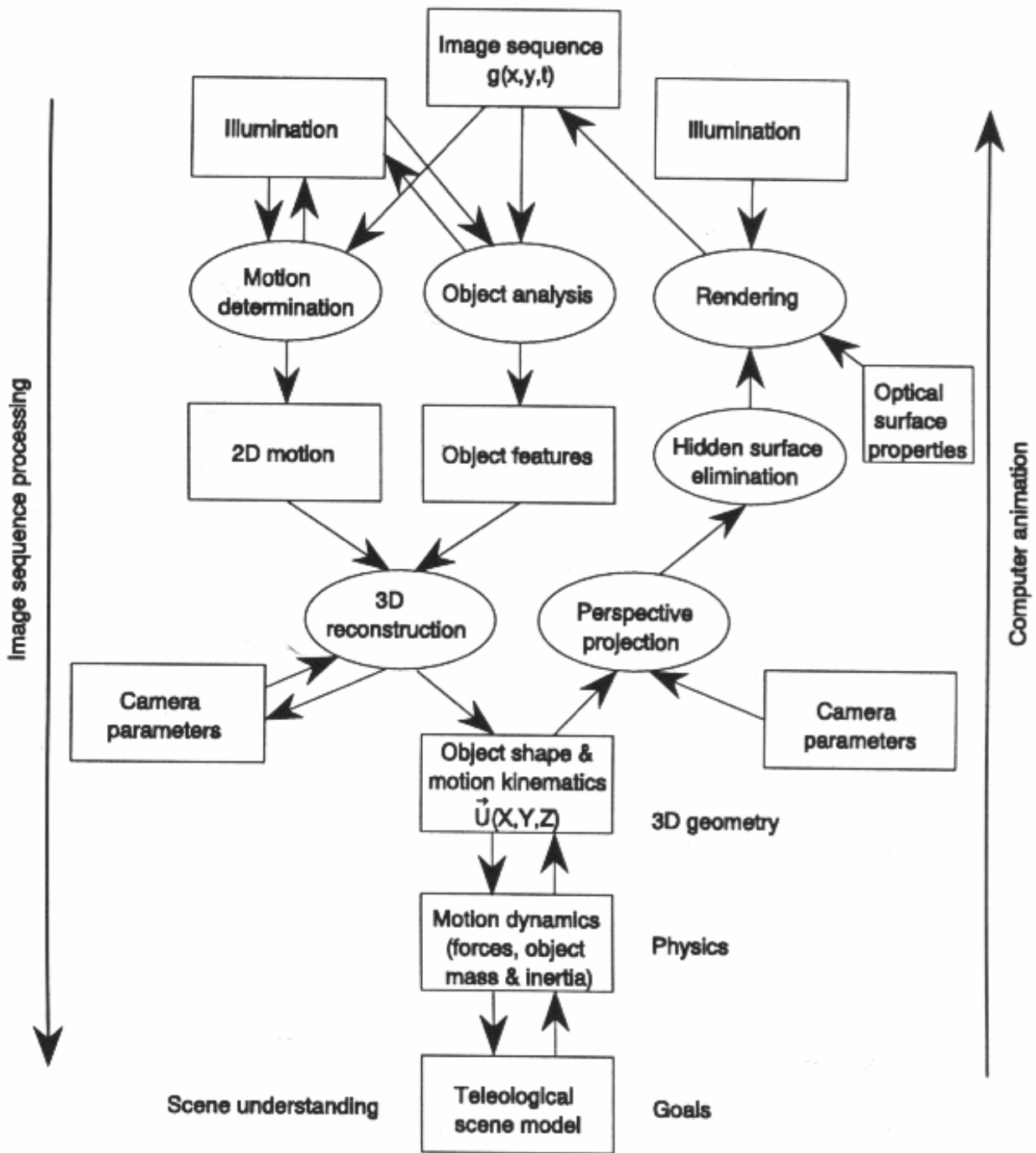
### **Interprétation spatiale et temporelle des séquences d'images**

- Détection**
- Estimation**
- Segmentation**
- Etiquetage**
- Analyse**

## **1.2 Domaines d'application**

### **Surveillance, observation, inspection:**

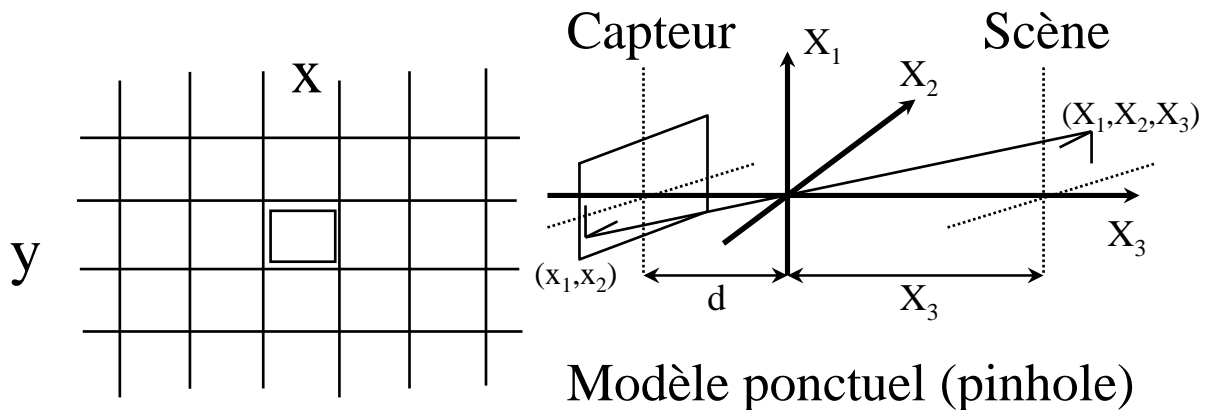
- Industrie**
- Applications militaires**
- Météorologie**
- Robotique (Interprétation de scènes)**
- Imagerie médicale (cœur et vaisseaux)**



D'après B. Jähne, Springer Verlag, 1993

## 1.3 Acquisition

- **Optique géométrique**
- **Quantification**



### - Proj. perspective (modèle ponctuel)

$$(X_1, X_2, X_3) \xrightarrow{P} \left( x_1 = \frac{d \cdot X_1}{X_3}, x_2 = \frac{d \cdot X_2}{X_3} \right)$$

### - Transfo. scène-image: $SPR_z R_y R_x T$

- influence sur la qualité des images
- ⇒ nécessité de disposer d'un système d'acquisition cohérent et performant
- ⇒ modélisation du processus d'acquisition

### – Variétés des sources:

- optique, X, infra-rouge, micro-ondes, ...

## 1.4 Description du mouvement

→ **Cinématique**: description de la trajectoire des objets (indépendamment des causes)

→ **Dynamique**: décrit comment les forces agissant sur un objet modifient son mouvement (+ de 2 images)

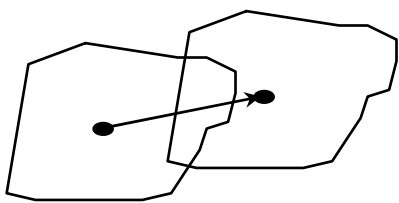
### • Cinématique

$$X = (X_1, X_2, X_3) \mapsto x = (x_1, x_2) \mapsto \hat{x} = (\hat{x}_1, \hat{x}_2) = \frac{x}{d}$$

$$V = (V_1, V_2, V_3) \mapsto v = (v_1, v_2) \mapsto \hat{v} = (\hat{v}_1, \hat{v}_2) = \frac{dx}{dt}$$

**Séquence d'images**: on mesure le déplacement  $d$ , on estime la vitesse par  $v = d/\Delta t$

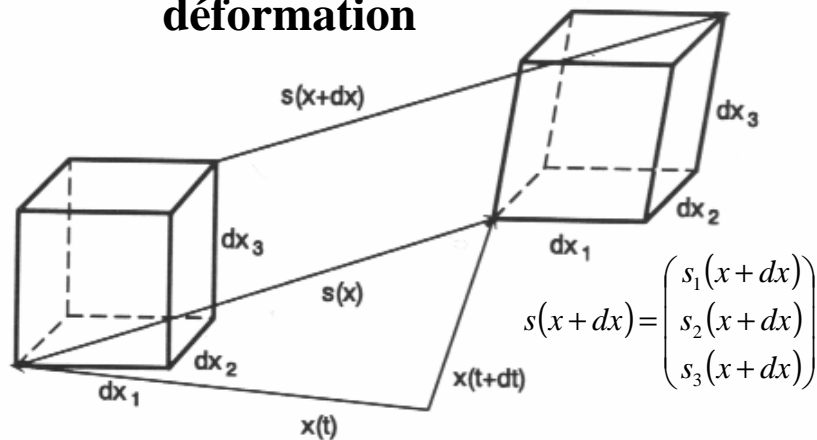
– **Mvt. d'un point de masse** (translation)


$$\begin{cases} \hat{v}_1 = \frac{1}{X_3} (V_1 - \hat{x}_1 V_3) \\ \hat{v}_2 = \frac{1}{X_3} (V_2 - \hat{x}_2 V_3) \end{cases}$$

Si  $v=0$ ,  $V$  n'est pas nécessairement nul  
⇒ pb. sous déterminé

## – Objets déformables

### déformation



$$s_i(x+dx) = s_i(x) + \frac{\partial s_i}{\partial x_1} dx_1 + \frac{\partial s_i}{\partial x_2} dx_2 + \frac{\partial s_i}{\partial x_3} dx_3, i = 1, 2, 3$$

$$A_{ij} = \frac{\partial s_i}{\partial x_j}, j = 1, 2, 3$$

$$s(x+dx) = s(x) + A dx = s(x) + (A^s + A^a) dx$$

*En petits déplacements:*

-  $A^s = 1/2(A + A^T)$ , déformation (6 d.l.)

$$A_{ii}^s = \frac{\partial s_i}{\partial x_i} \text{ (changement de longueur)}$$

$$A_{ij}^s = \frac{1}{2} \left( \frac{\partial s_i}{\partial x_j} + \frac{\partial s_j}{\partial x_i} \right) \text{ (cisaillement)}$$

-  $A^a = 1/2(A - A^T)$ , rotation (3 d.l.)

$$A_{ii}^a = 0, A_{ij}^a = \frac{1}{2} \left( \frac{\partial s_i}{\partial x_j} - \frac{\partial s_j}{\partial x_i} \right) \rightarrow \text{rotation d'axe } \omega$$

- Dilatation volumique

$$\theta_v = \frac{dV' - dV}{dV} = \text{trace}(A^s)$$

- En résumé:

$$s(x + dx) = s(x) \quad + \quad \omega \times dx \quad + \quad A^s dx$$

translation      rotation      déformation

- **Dynamique**

→ **Introduction des lois physiques de mouvement:**

$$a = \frac{\partial v}{\partial t} = \frac{F}{m}$$

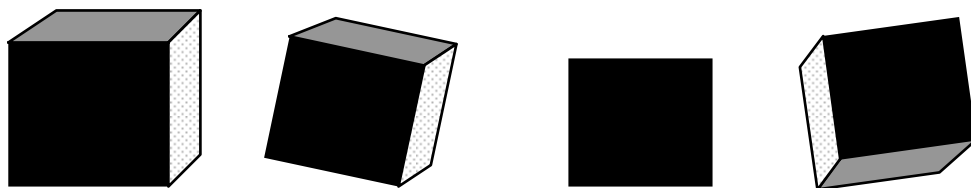
+ 2 images

→ prise en compte explicite de la dimension temporelle, de l'inertie...

## 1.5 Les difficultés

### – non spécifiques

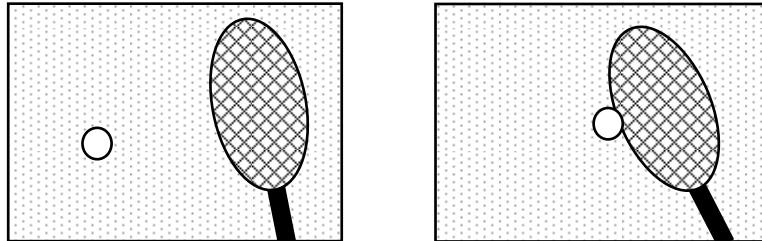
- bruit: capteur, quantification...
- projection 3D->2D: ambiguïté d'interprétation de la forme 3D, du mouvement, occultations



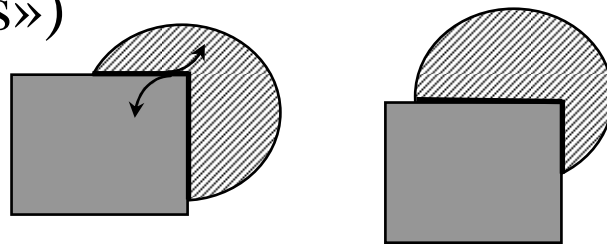
- caractéristiques des objets: forme, taille, texture
- caractéristiques de l'acquisition: optique (éclairage, ombrages), RX (superpositions)...

## – spécifiques

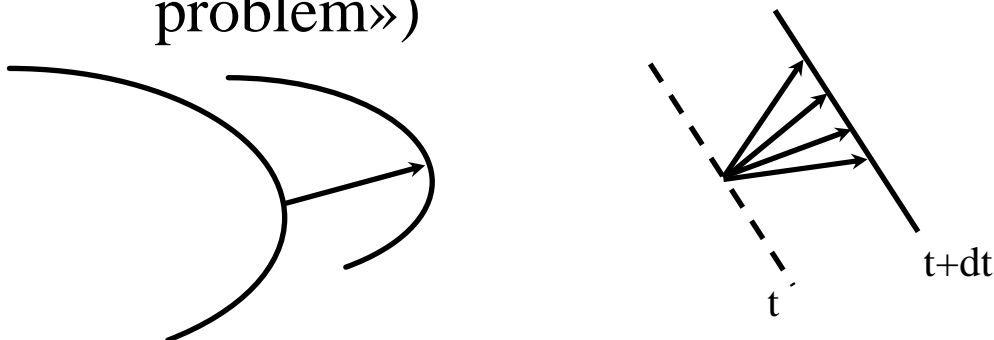
- vitesses propres et relatives des objets (cadence d'acquisition)



- quantité d'information et traitement en temps réel
- objets déformables: on ne cherche pas forcément le même objet
- prise de vue fixe ou mobile
- frontières d'occultation («boundary occlusions»)

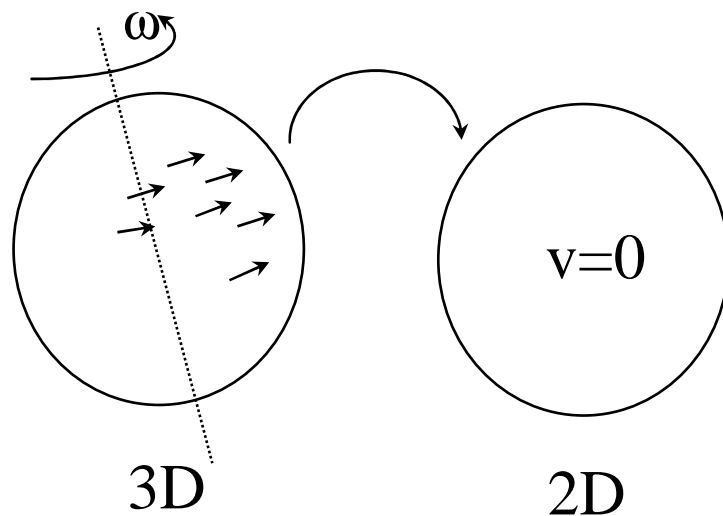


- problème d'ouverture («aperture problem»)



## 2-Définitions

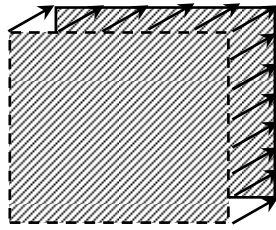
- **Mouvement apparent**  
→ **Mouvement perçu grâce au changement de la distribution spatiale des intensités lumineuses**  
≠ **de la projection dans le plan image du mouvement 3D**



- **Champ des vitesses apparentes = flux optique («optical flow»)**

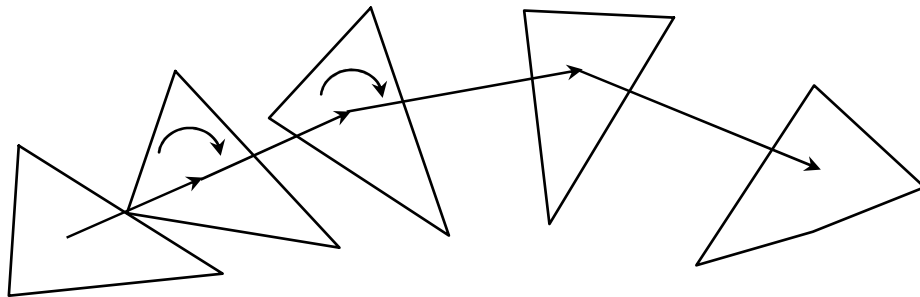
- **2 approches liées à 2 mécanismes psychophysologiques de la perception du mouvement:**

- **Déplacements de petite amplitude («short range»)**



→ Réponse aux changements de l'intensité lumineuse

- **Déplacements de grande amplitude («long range»)**



→ Extraction d'indices visuels dans l'image courante, mise en correspondance dans les images suivantes

# 3- Détection du mouvement apparent

- hypothèses
  - invariance d'illumination
  - caméra fixe
  - aucune connaissance *a priori* sur le mouvement des objets
  - aucune connaissance *a priori* sur la nature des objets

## 3.1 Différence entre images

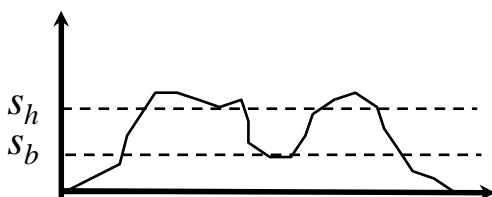
Changement temporel validé si:

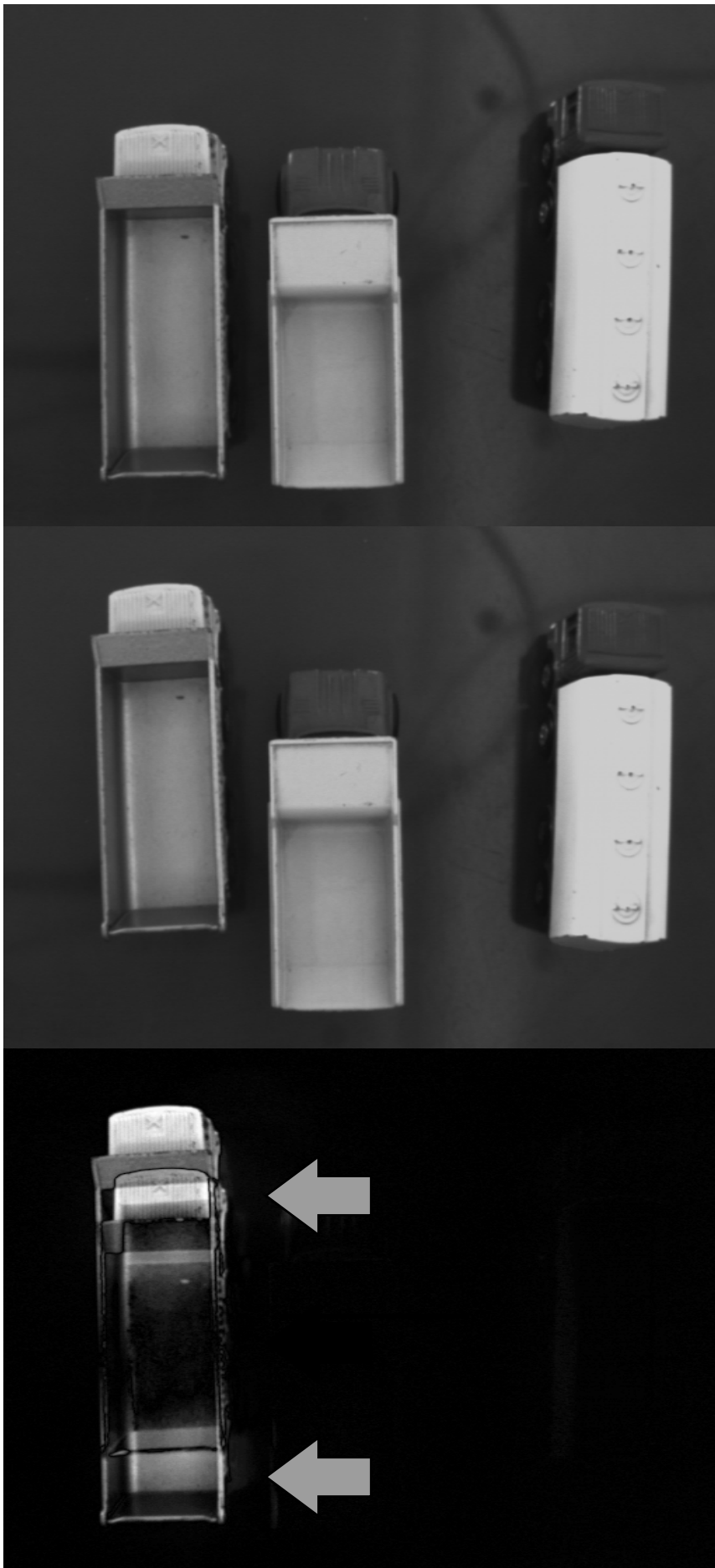
$$|I(x, y, t) - I(x, y, t - dt)| \geq \text{seuil}$$

⇒ très sommaire et sensible au bruit

$$M_t(V_{x,y}) - M_{t-dt}(V_{x,y}) + \text{seuil hystérésis}$$

$M$ : métrique (statistiques),  $V$ : voisinage





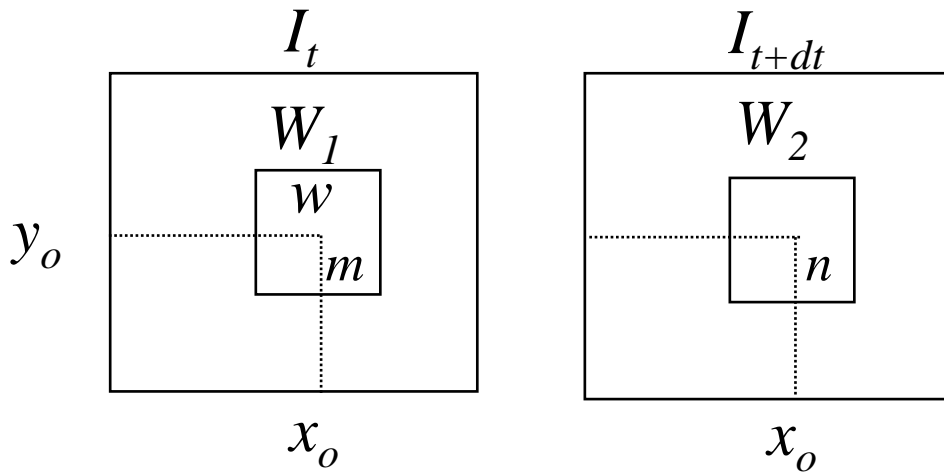
### 3.2 Test de vraisemblance [Hsu 84]

– **Modélisation**

**distribution d'intensité** = fonction polynomiale  
(cst/linéaire/quadratique)

+ bruit Gaussien  $N(0, \sigma)$

– **Test d'hypothèse**



$H_0$ : même dist.  $N(\mu, \sigma)$  pour  $W_1$  et  $W_2$

➤ *Zone stationnaire*

$H_1$ : dist. différentes pour  $W_1$  et  $W_2$ :  $N(\mu_1, \sigma)$  et  $N(\mu_2, \sigma)$

➤ *Zone mobile*

Critère de décision :

$$\frac{p(w|H_1)}{p(w|H_0)} > \underbrace{\theta \frac{P_0}{P_1}}_{\text{Seuil}}$$

Prob. a priori

– **Test de vraisemblance**

→ si  $H_1$  vraie, la fonction de vraisemblance sera (Loi de Gauss):

$$L(H_1) = \left( \frac{1}{2\pi\sigma^2} \right)^{\binom{m+n}{2}} \exp \left\{ \frac{1}{2\sigma^2} \left[ \sum_{i=1}^m (w_i^1 - \mu_1)^2 + \sum_{i=1}^n (w_i^2 - \mu_2)^2 \right] \right\}$$

$$\frac{\partial L}{\partial \mu_1}, \frac{\partial L}{\partial \mu_2}, \frac{\partial L}{\partial \sigma} \rightarrow 0$$

$$\Rightarrow \left\{ \begin{array}{l} \hat{\mu}_1 = \frac{1}{m} \sum_{i=1}^m w_i^1, \hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^n w_i^2 \\ \hat{\sigma}^2 = \frac{1}{m+n} \left[ \sum_{i=1}^m (w_i^1 - \hat{\mu}_1)^2 + \sum_{i=1}^n (w_i^2 - \hat{\mu}_2)^2 \right] \end{array} \right.$$

→ si  $H_0$  vraie, la fonction de vraisemblance sera  $L(H_0)$  :

$$\Rightarrow \left\{ \begin{array}{l} \hat{\mu}_0 = \frac{1}{m+n} \left[ \sum_{i=1}^m w_i^1 + \sum_{i=1}^n w_i^2 \right] \\ \hat{\sigma}_0^2 = \frac{1}{m+n} \left[ \sum_{i=1}^m (w_i^1 - \hat{\mu}_0)^2 + \sum_{i=1}^n (w_i^2 - \hat{\mu}_0)^2 \right] \end{array} \right.$$

→ **Maximisation du rapport de vraisemblance:**

$$\frac{L(H_1)}{L(H_0)} > 1 \quad \rightarrow (H_1)$$

$$\frac{L(H_1)}{L(H_0)} < 1 \quad \rightarrow (H_0)$$

- **Ne peut prendre en compte une variation d'illumination**
- **Changement d'illumination global**
  - Transformation linéaire d'intensité

$$\hat{I}_t(w) = aI_t(w) + b$$

such that  $\hat{\mu}_t = \mu_{t+dt}$  and  $\hat{\sigma}_t^2 = \sigma_{t+dt}^2$

- **Alors**

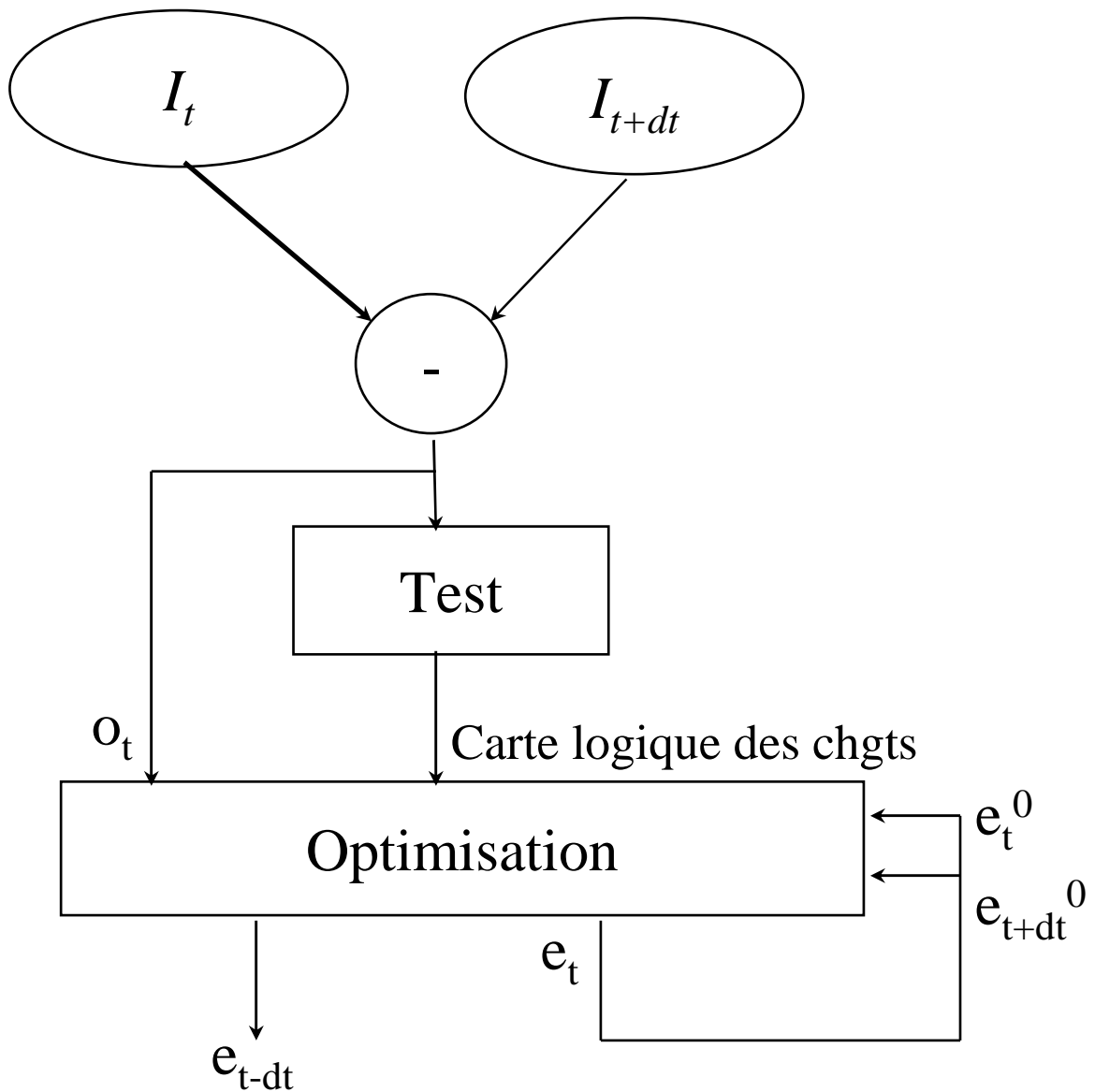
$$\hat{I}_t(w) = \frac{\sigma_{t+dt}}{\sigma_t} (I_t(w) - \mu_t) + \mu_{t+dt}$$

- **Changement locaux : comparaison de gradients d'intensité**

$$\frac{1}{N} \sum_{w \in W} \left\| \nabla I_t(w) - \nabla I_{t+dt}(w) \right\| \begin{matrix} > \\ < \end{matrix} \theta$$

### 3.3 Approche basée sur un modèle

→ **Détection des variations d'intensité et extraction des masques d'objets mobiles**  
[Bouthemy 90].



- **Maximum a posteriori et champ de Markov**

- Soient  $e$  les étiquettes de mouvement,  $o$  les observations.
- On cherche (maximum a posteriori):

$$\max_e p(E = e, O = o)$$

- **Modélisation: champ de Markov**

$$\begin{aligned} p(E = e, O = o) &= \frac{1}{Z} \exp(-U(e, o)) \\ &= p(O = o | E = e) p(E = e) \end{aligned}$$

⇒ minimiser la fonctionnelle d'énergie globale  $U$

$$p(O = o|E = e) = \prod_{image} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma} \varepsilon^2(\mathbf{x}, e)\right)$$

$$\begin{aligned} \log(p(O = o|E = e)) &= \sum_{image} \log\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma} \varepsilon^2(\mathbf{x}, e)\right)\right) \\ &= C_1 - \frac{1}{2\sigma} \sum_{image} \varepsilon^2(\mathbf{x}, e) \end{aligned}$$

$$P(E = e) = \frac{1}{Z} \exp(-U(e))$$

$$\log(P(E = e)) = C_2 - U(e)$$

$$U(e) = \sum_{c \in C} V_c(e)$$

au total, la maximisation de  $P(O = o|E = e)P(E = e)$

conduit à la minimisation de :

$$C + U(e) + \frac{1}{2\sigma} \varepsilon^2(\mathbf{x}, e)$$

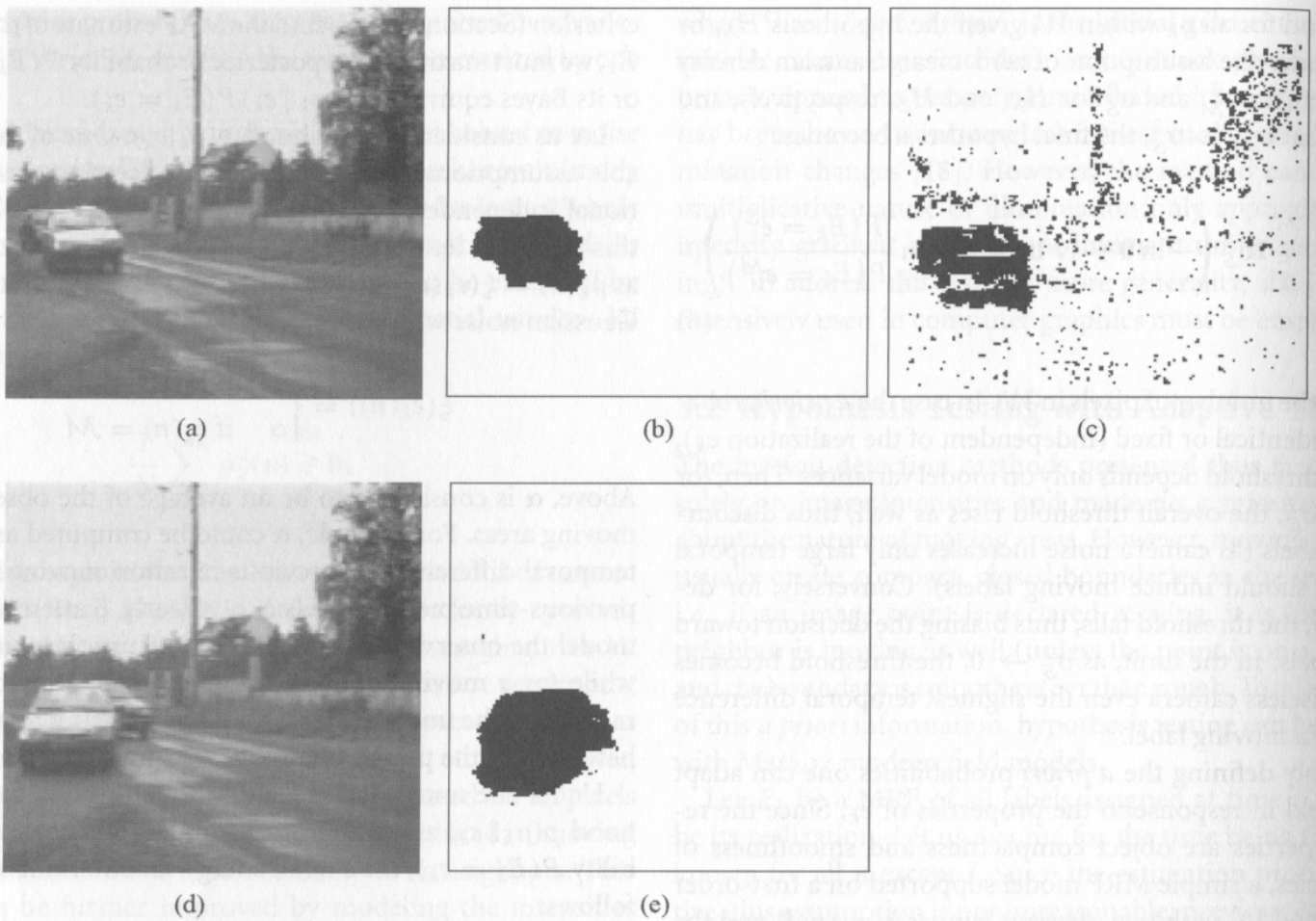
$$U(e_{t-dt}, e_t, o_t, \bar{o}_t) = U_{et}(e_{t-dt}, e_t, o_t) + U_T(e_{t-dt}, e_t, \bar{o}_t) + U_e(e_t)$$

$U_{et}$  : Consistence de l'étiquetage/variations temporelles

$U_T$  : Propriétés temporelles des masques d'objets mobiles

$U_e$  : Propriétés spatiales du (Plan de l'Étiquette)

# Exemple



**FIGURE 1** Motion detection results for two frames of a road traffic sequence of images: (a) frame 137; (b) MAP-detected motion for frame 137; (c) motion in frame 137 detected by the fixed-threshold hypothesis test; (d) frame 143; (e) MAP-detected motion for frame 143. (From Bouthemy and Lalande [6]. Reproduced with permission of the SPIE.)

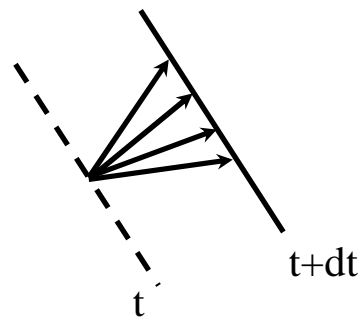
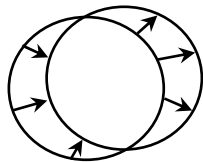
*From Handbook of Image and Video Processing, 2000*

# 4- Estimation d'un champ dense

- **En général:**

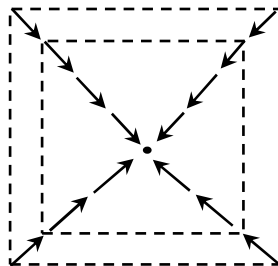
- le flux optique est différent de la projection dans le plan image du mouvement 3D

- Pb. d'ouverture



- **flux optiques exacts**

- Translation selon l'axe de vue



- Rotation dans un plan parallèle au plan image

- **Les approches principales:**
  - **Approches différentielles**
  - **Mise en correspondance**
  - **Techniques spatio-temporelles**

## 4.1 Approches différentielles

→ **Approches classiques sur 2 images**

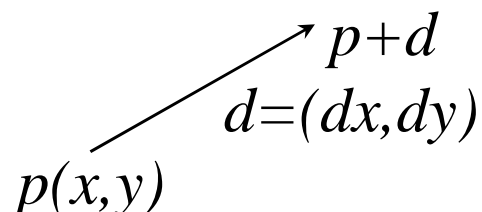
→ **Estimation de petits déplacements**

– **Hypothèse:**

**Niveau de gris constant au cours du mouvement:**

$$I(x+dx, y+dy, t+dt) = I(x, y, t)$$

⇒ **Illumination cste, pas d'ombrages**



A diagram illustrating a displacement vector. It shows a point labeled  $p(x, y)$  at the bottom left. An arrow points from this point to a point labeled  $p+d$  at the top right. The arrow is labeled with the vector  $d = (dx, dy)$ .

– **Displaced Frame Difference (DFD)**

$$DFD(p, d, t) = I(p + d, t + dt) - I(p, t)$$

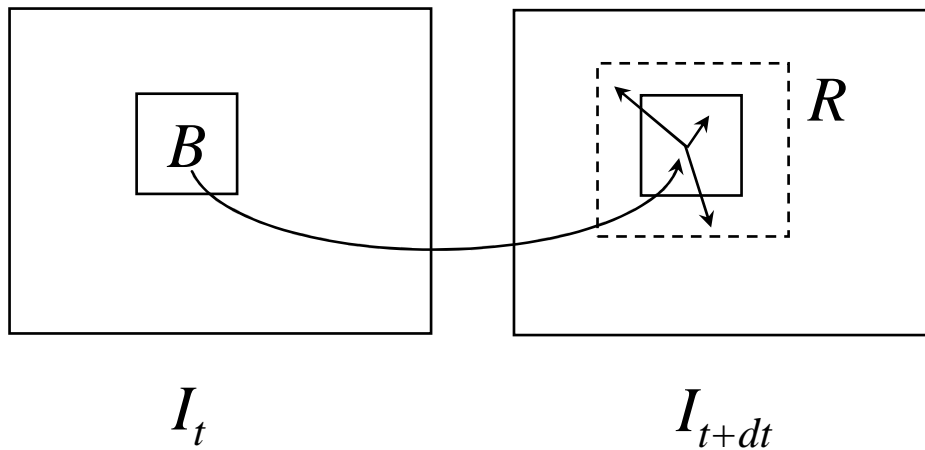
⇒ Minimisation de la  $DFD$ :  $DFD \rightarrow 0$

- **Minimisation directe de la DFD**

- **Similarité de blocs**

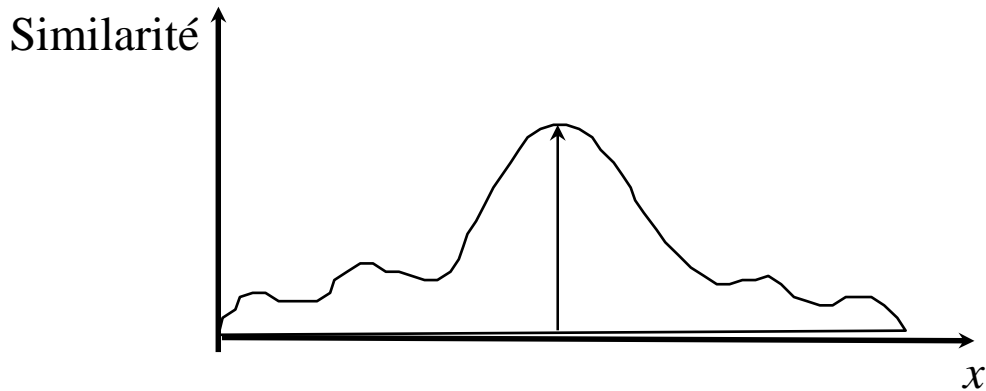
- Principe: Recherche pour tout point de  $I_t$  de son correspondant dans l'image  $I_{t+dt}$
    - Hypothèse: les pixels voisins subissent le même mouvement (~translation)

=> maximiser la ressemblance ou minimiser la différence



$R$ : zone de Recherche

- Critères de (di)similarité



$$DQ(x, y, u, v) = \sum_{i=1}^I \sum_{j=1}^J \left( I_{t+dt}(x+u+i, y+v+j) - I_t(x+i, y+j) \right)^2$$

$$DA(x, y, u, v) = \sum_{i=1}^I \sum_{j=1}^J \left| I_{t+dt}(x+u+i, y+v+j) - I_t(x+i, y+j) \right|$$

Corrélation simple:

$$C(x, y, u, v) = \sum_{i=1}^I \sum_{j=1}^J \left( I_{t+dt}(x+u+i, y+v+j) I_t(x+i, y+j) \right)$$

Corrélation centrée:

$$C(x, y, u, v) = \frac{\sum_{i=1}^I \sum_{j=1}^J \left[ I_{t+dt}(x+u+i, y+v+j) - \bar{I}(x+u, y+v) \right] \left[ I_t(x+i, y+j) - \bar{I}_t(x, y) \right]}{\sum_{i=1}^I \sum_{j=1}^J \left[ I_{t+dt}(x+u+i, y+v+j) - \bar{I}(x+u, y+v) \right] \left[ I_t(x+i, y+j) - \bar{I}_t(x, y) \right]}$$

Corrélation normée par la variance:

$$C(x, y, u, v) = \frac{\sum_{i=1}^I \sum_{j=1}^J \left[ I_{t+dt}(x+u+i, y+v+j) - \bar{I}(x+u, y+v) \right] \left[ I_t(x+i, y+j) - \bar{I}_t(x, y) \right]}{\sqrt{\sigma^2(I_{t+dt}(x+u, y+v))} \sqrt{\sigma^2(I_t(x, y))}}$$

## → **Remarques:**

- Il faut que l'image ait des détails
- la corrélation directe ne s'affranchit pas des conditions d'éclairement
- la corrélation normée par la variance est moins sensible au bruit. Elle s'affranchit des petites déformations de l'objet mais elle est coûteuse.

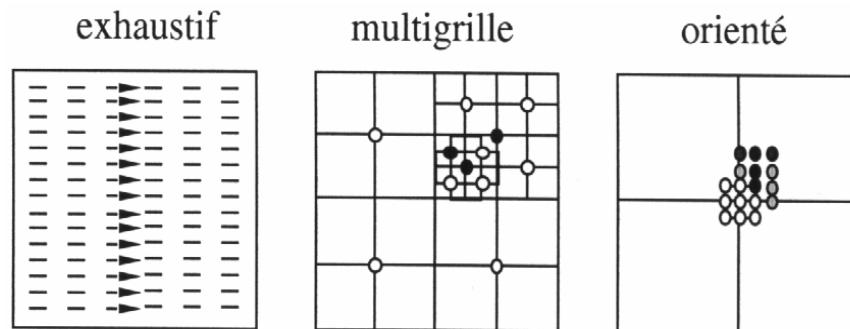
## → **Stratégies de recherche**

- Taille des fenêtres = compromis
  - $R$  grand  $\Rightarrow$  +temps, +confusion
  - $R$  petit  $\Rightarrow$  petit déplacements
  - $B$  grand  $\Rightarrow$  +temps, risque objets diff.
  - $B$  petit  $\Rightarrow$  pas discriminant
- La mise en correspondance peut être limitée aux zones susceptibles de contenir des motifs discriminants (contours ou structures plus complexes)

$\Rightarrow$  Champ épars

- $DQ$  ou  $DA$  sont cumulatifs  $\Rightarrow$  on peut suspendre le calcul au delà d'un certain seuil à déterminer
- Techniques hiérarchiques:

- 1) Évaluation d'un grand déplacement
- 2) Raffinement de l'estimation



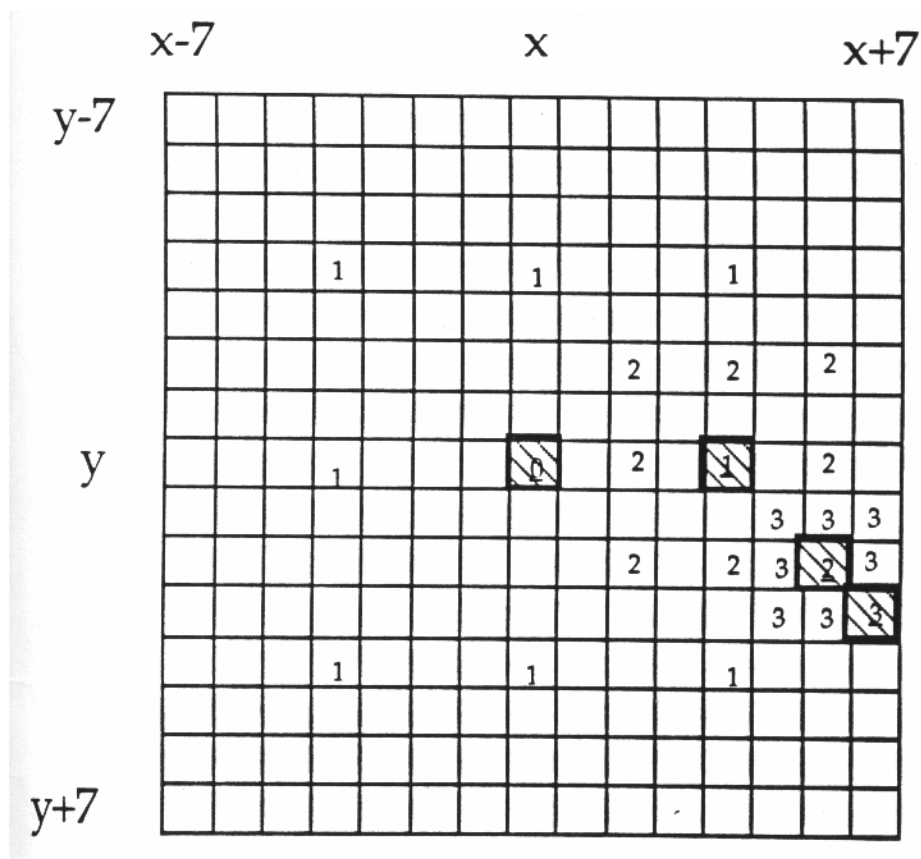
(D'après [Orkisz 97])

## – Performances

- + simplicité, détection des blocs sans correspondant (occultations)
- temps calcul, mvts. localement translationels

## Thomas & Bierling [Thomas 89]:

- sous-échantillonnage des fenêtres
- Technique du log D-step
- et de nombreuses variantes



O: Point de départ de la recherche

1,2,3: Points de recherche aux étapes 1,2,3

1,2,3: Points de meilleure correspondance aux étapes

1,2,3

(D'après [Devevey 92])

- Algorithme de Li et Salari :  
Successive Elimination Algorithm  
(SEA)
    - **Pour la mesure DA, on peut établir un encadrement qui permet d'éliminer d'emblée un nombre significatif de candidats lors de la recherche du minimum**
- ⇒ Le résultat est identique à la recherche exhaustive en un temps nettement inférieur
- ⇒ Démonstration

[Li & Salari, IEEE IP, 1995]

# Algo de mise en correspondance de blocs standard

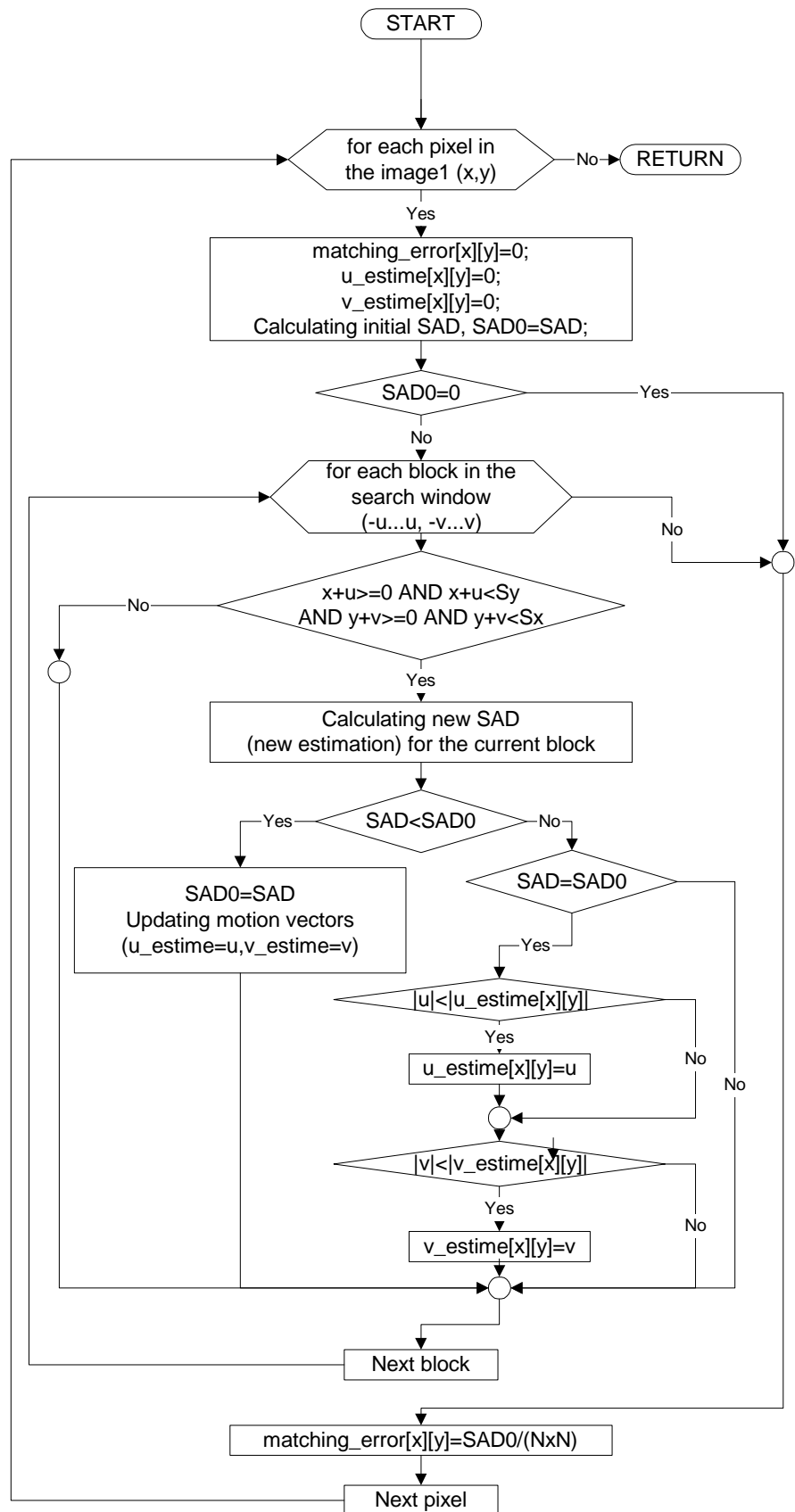
## Estimation using Block Matching (SAD)

**image size:**  $S_x \times S_y$   
**maximal displacement:**  $d\_max$   
**search window size:**  
 $(2 \cdot d\_max + 1) \times (2 \cdot d\_max + 1)$   
**block size:**  $N \times N$   
**motion vector components:**  
 $u\_estime, v\_estime$   
**error image:**  $matching\_error$

Eliminates the positions where the initial estimation is 0. Means that the minimum value for MAD was obtained and the best match was obtained for the current pixel

Eliminates the search positions which are not in the image.

Condition for minimizing the error



# Algo de Li et Salari

## Estimation using SEA (detailed)

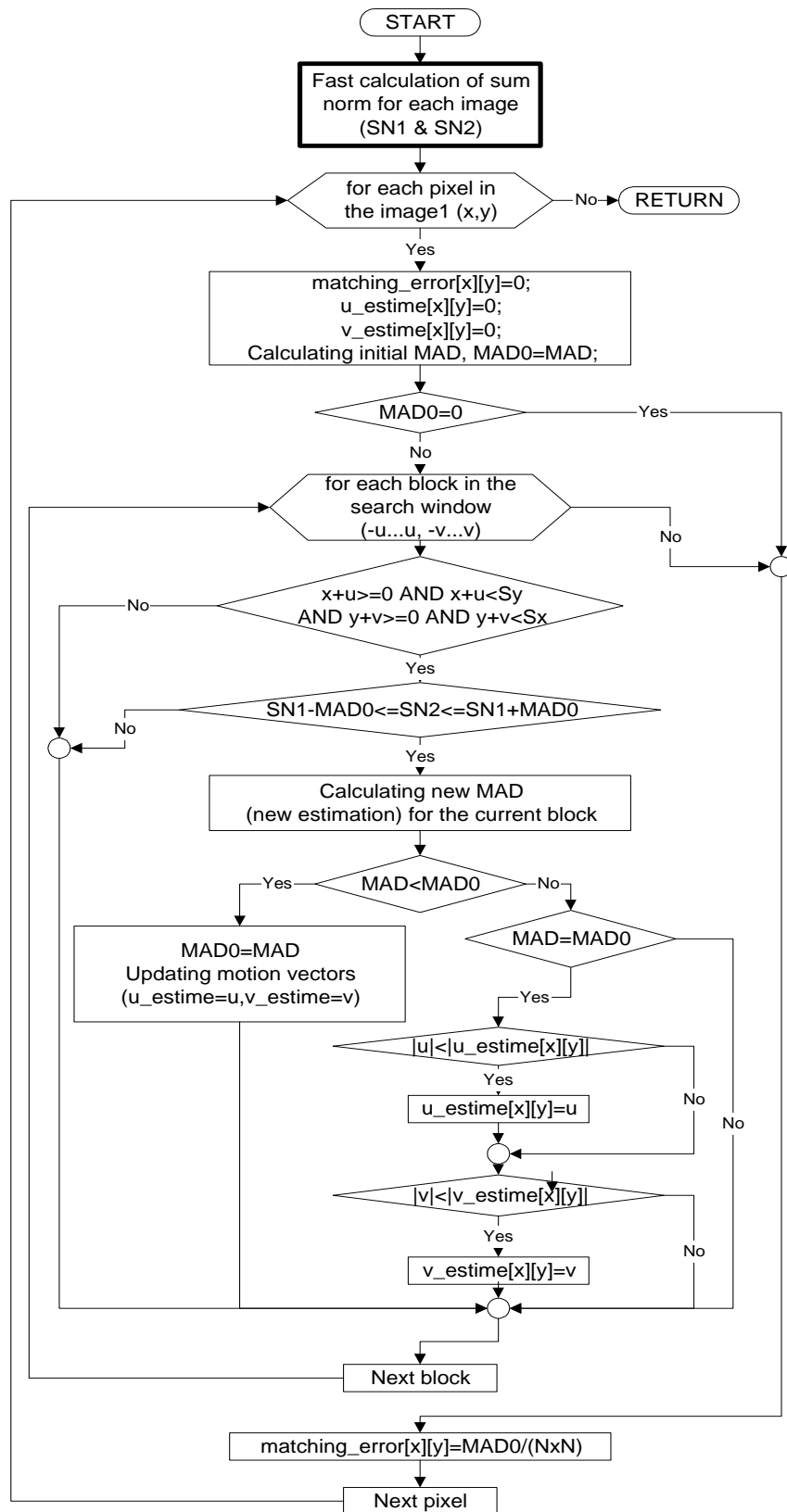
**image size:**  $S_x \times S_y$   
**maximal displacement:**  $d_{max}$   
**search window size:**  $(2 \cdot d_{max} + 1) \times (2 \cdot d_{max} + 1)$   
**block size:**  $N \times N$   
**motion vector components:**  $u_{estimate}, v_{estimate}$   
**error image:**  $matching\_error$   
**sum norms images:**  $SN1, SN2$

Eliminates the positions where the initial estimation is 0. Means that the minimum value for MAD was obtained and the best match was obtained for the current pixel

Eliminates the search positions which are not in the image.

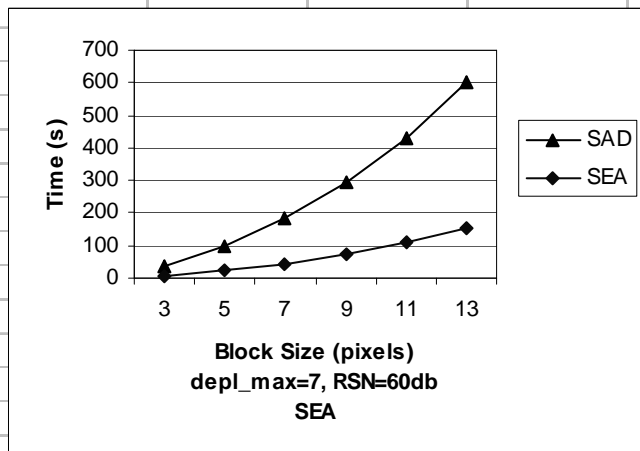
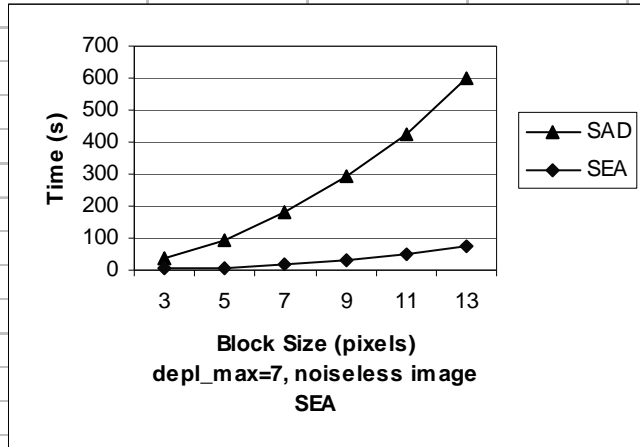
SEA main condition.

Condition for minimizing the error



# Exemple de performances

Ffd7	SAD	SEA
3	38,981774	3,383198
5	96,812794	8,832980
7	182,559364	18,265936
9	295,838166	32,215378
11	428,016212	50,931296
13	598,692718	73,213738
Ffd7_Bruit60	SAD	SEA
3	38,981774	7,383038
5	96,812794	22,649094
7	182,559364	44,931536
9	295,838166	73,947042
11	428,016212	110,245590
13	598,692718	151,660600



- **Minimisation directe de la DFD**

- **Descente de gradient**

- **Algo. récursif de descente [Netravali 79]**

$$\begin{aligned}\hat{d}_{k+1} &= \hat{d}_k - \frac{1}{2} \varepsilon \nabla \left[ DFD^2(p_k, \hat{d}_k, t_k) \right] \\ &= \hat{d}_k - \varepsilon DFD(p_k, \hat{d}_k, t_k) \nabla I(p_k - d_k, t_k - dt)\end{aligned}$$

$$\text{où } \nabla I(p_k - d_k, t_k - dt) = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} I(p_k - d_k, t_k - dt)$$

- **Développement au 1er ordre DFD**  
[HORN 81]

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + O^2$$

En négligeant les termes d'ordre supérieur:

$$DFD(p, d, t) = I_x dx + I_y dy + I_t dt$$

$DFD = 0$ , donne l'équation de contrainte du mouvement ou équation de contrainte du flux optique :

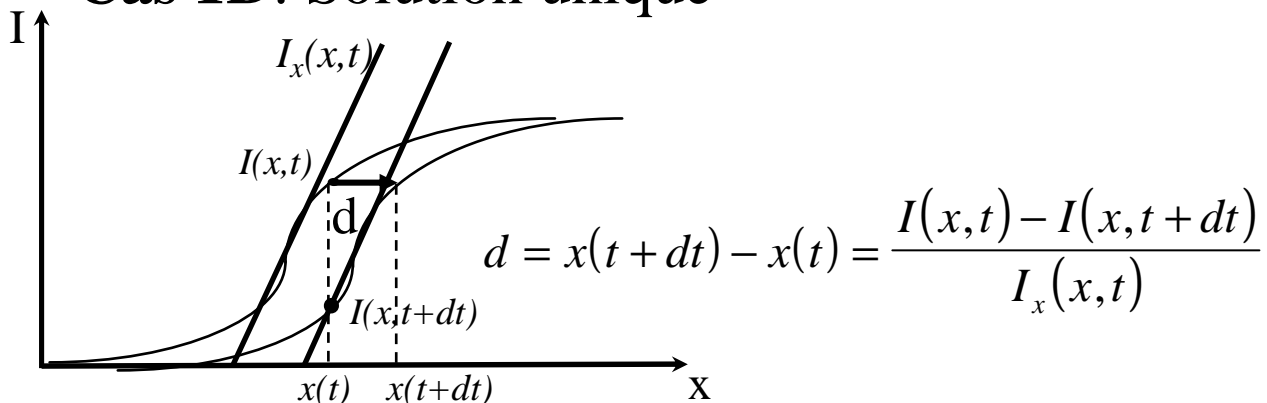
$$I_t + \nabla I \cdot \mathbf{v} = 0, \quad \mathbf{v} = \begin{pmatrix} u = \frac{dx}{dt} \\ v = \frac{dy}{dt} \end{pmatrix} \quad (1)$$

(Image flow constraint equation or

Brightness Change Constraint Equation (BCCE))

– **Hypothèse:** invariance de la luminance  $\Rightarrow$  pas de variation de la source lumineuse, pas d'ombrage

– **Cas 1D:** Solution unique



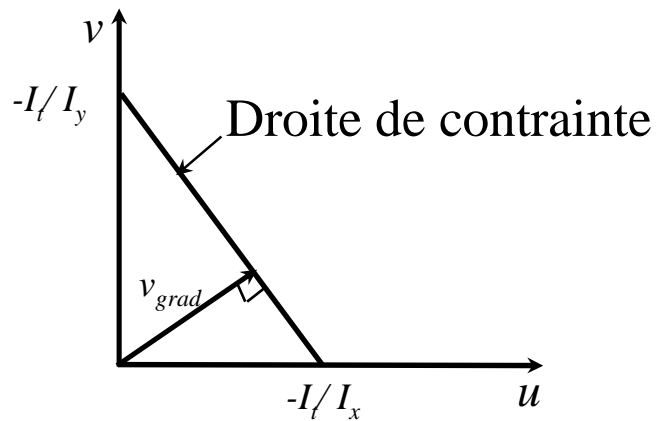
– **Cas 2D: 1eq. 2 inconnues**

→ **Proj. du vecteur vitesse sur le gradient spatial de  $I$**

$$\mathbf{v} = (u, v)$$

$$I_t + I_x u + I_y v = 0 \quad (2)$$

$$v_{grad} = \frac{I_t}{\|\nabla I\|}$$



⇒ Il faut une contrainte supplémentaire

– **Analogie dynamique des fluides**

→ **Equation de continuité** qui exprime la conservation de masse dans l'écoulement d'un fluide de densité  $\rho$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \underbrace{\frac{\partial \rho}{\partial t} + (\mathbf{v} \cdot \nabla) \rho + \rho (\nabla \cdot \mathbf{v})}_{\text{ECM}} = 0 \quad (3)$$

La variation temporelle de la densité est compensée par la divergence du flux de densité  $\mathbf{v}\rho$

→ **Cas du flux optique:  $\rho \rightarrow I$**

traduit la variation temporelle de  $I$  par un gradient d'intensité mobile

traduit une variation même dans les régions de luminance cste si la divergence du flux n'est pas nulle

La luminance ne s'accumule pas comme la masse,  $I \nabla \cdot \mathbf{v} \Rightarrow 0$

## – Optimisations locales

- *Méthode multi-entrées*

Si on considère les entrées  $e_1, e_2, \dots, e_n$ , on obtient un système à  $n$  équations, 2 inconnues (surdéterminé)

$$e_t^i + \nabla e_i \mathbf{v} = 0, i = 1..n \quad (4)$$

⇒ Chaque pixel conserve l'intensité et d'autres caractéristiques

– Ex:

- images spectrales
- images RGB [Wohn 83]

– Souvent les composantes sont corrélées et n'apportent aucune info. supplémentaire.

- *Méthode multi-opérateurs  $O_i(I)$*

– spatiaux: variance locale, contraste, entropie...[Mitiche 87]

– différentiels: dérivées partielles de  $I$

$$O_t^i + \nabla O_i \mathbf{v} = 0, i = 1..n \quad (5)$$

- *Contrainte locale: exp. analytique*

– mvt. localement constant

– décomposition rotation / extension / cisaillement

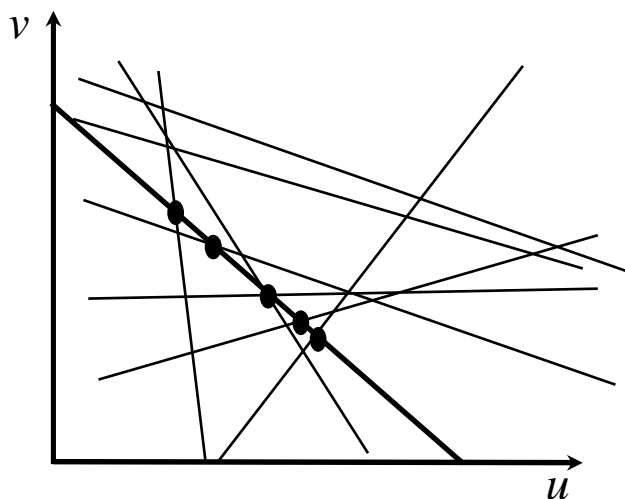
- **Regroupement des droites de contraintes** «*constraint line clustering*»

**Idéalement:** les droites de contrainte des pts d'un même motif s'intersecte en 1 seul pt.

**En pratique:** intersections multiples (bruit, non respect des hypothèses)

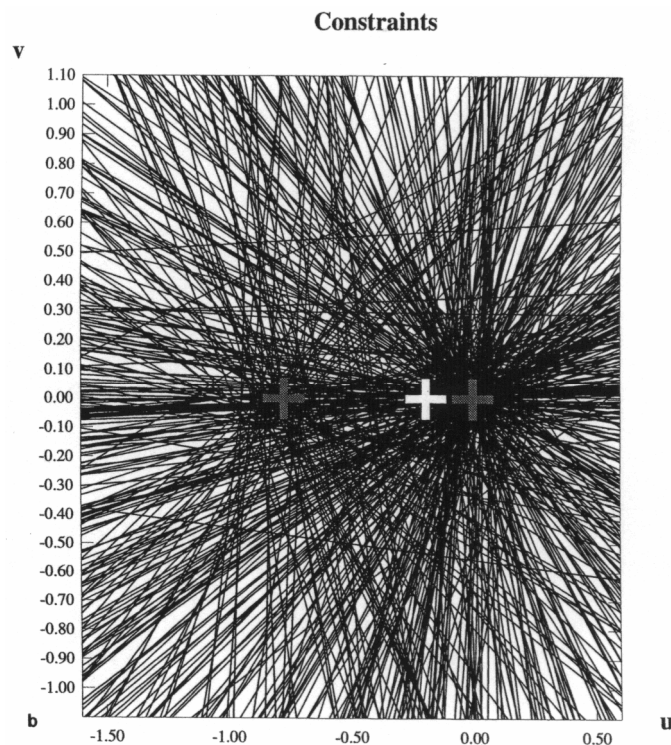
→ Travail sur un voisinage d'un point

→ Indétermination dans les zones homogènes et au niveau des frontières d'objets en mouvement





**Séquence** : une personne se déplace derrière une plante, la caméra est fixe



Un certain nombre de droites de contrainte et la solution aux moindres carrés (croix blanche)  
[Black & Anandan, CVIU, 1996]

## – Optimisation globale

→ Introduction d'une contrainte spatiale sur le flux: «**contrainte de lissage**»

⇒ **Les points voisins ont une vitesse similaire et le flux varie continûment**

## Minimisation de la fonctionnelle globale [Horn & Schunk 81]:

$$\iint \left[ (\mathbf{v} \nabla I + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right] dx dy \quad (6)$$

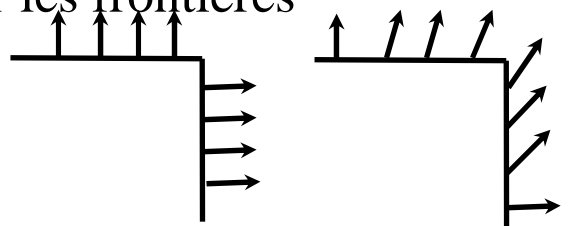
eq. flot optique

lissage

## = Régularisation

- Les calculs peuvent être limités aux zones où le gradient est élevé
  - détection de contours
  - estimation locale de la vitesse
  - propagation de la vitesse sur le contour [Hildreth 84]

⇒ Hyp. fautive sur les frontières d'occultation



- **Régularisation**
- **Minimisation**

• **Résolution (Horn & Schunk)**

$$f(u, v, u_x, u_y, v_x, v_y) = (uI_x + vI_y + I_t)^2 + \alpha^2 (u_x^2 + u_y^2 + v_x^2 + v_y^2) \quad (8)$$

Les équations d'Euler - Lagrange sont :

$$\begin{cases} I_x^2 u + I_x I_y v + I_x I_t - \alpha^2 \Delta u = 0 \\ I_y^2 v + I_x I_y u + I_y I_t - \alpha^2 \Delta v = 0 \end{cases} \quad (9)$$

où  $\Delta u = \nabla^2 u$  est le Laplacien de  $u$

Discrétisation:

• Gradient = moyenne de 4 mesures:

$$I_x = \frac{1}{4} \left( \begin{aligned} &I_{i,j+1,t} - I_{i,j,t} + I_{i+1,j+1,t} - I_{i+1,j,t} \\ &+ I_{i,j+1,t+1} - I_{i,j,t+1} + I_{i+1,j+1,t+1} - I_{i+1,j,t+1} \end{aligned} \right)$$

• Laplacien

$$\nabla^2 u \approx 3(\bar{u}_{i,j,k} - u_{i,j,k})$$

$$\begin{aligned} \bar{u}_{i,j,k} = &\frac{1}{6} \left\{ u_{i-1,j,k} + u_{i,j+1,k} + u_{i+1,j,k} + u_{i,j-1,k} \right\} \\ &+ \frac{1}{12} \left\{ u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j+1,k} + u_{i+1,j-1,k} \right\} \end{aligned}$$

dans (9):

$$\begin{cases} (\alpha^2 + I_x^2)u + I_x I_y v = (\alpha^2 \bar{u} - I_x I_t) \\ I_x I_y u + (\alpha^2 + I_y^2)v = (\alpha^2 \bar{v} - I_y I_t) \end{cases} \quad (10)$$

ou enfin

$$\begin{cases} (\alpha^2 + I_x^2 + I_y^2)u = (\alpha^2 + I_y^2)\bar{u} - I_x I_y \bar{v} - I_x I_t \\ (\alpha^2 + I_x^2 + I_y^2)v = (\alpha^2 + I_x^2)\bar{v} - I_x I_y \bar{u} - I_y I_t \end{cases} \quad (11)$$

→ Grande matrice creuse (nbre lignes, colonnes = 2 fois le nombre de pixels de l'image)

⇒ méthode itérative de Gauss-Seidel:

$$u^{n+1} = \frac{\bar{u}^n - I_x [I_x \bar{u}^n + I_y \bar{v}^n + I_t]}{(\alpha^2 + I_x^2 + I_y^2)}$$

Conditions aux bords: recopie des vitesses

$\alpha$  intervient lorsque le gradient est faible, il doit être du même ordre de grandeur que le bruit sur

$$E_x^2 + E_y^2$$

• **Lissage dans le temps [Black 91],[Singh 92]**

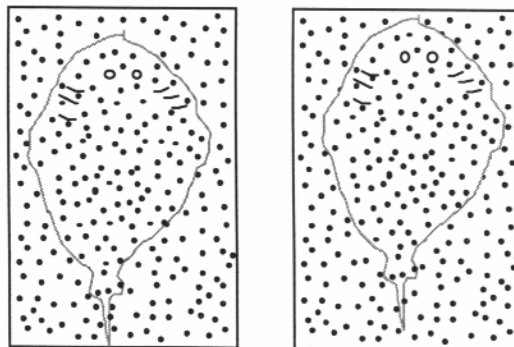
$v(t) \rightarrow v(t+dt)$ : prédiction + incertitude sur les estimations par filtrage de Kalman

- ***Autres contraintes de lissage***

- Laplacien de  $\mathbf{v}$  [Horn 81]
- Contrainte de divergence nulle [Fitzpatrick 88], [Song 91] (extension 3D).

$$I\nabla\mathbf{v} = 0$$

- lissage orienté [Nagel 87]
- Les frontières de mvt. coïncident avec les contours photométriques
- hyp. pas tj. vraie:



**zone homogène:** lissage isotrope

**contour:** lissage selon la tangente au contour

$$\iint \left[ (\mathbf{v}\nabla I + I_t)^2 + \lambda \text{trace}(\nabla\mathbf{v}^T W \nabla\mathbf{v}) \right] dx dy \quad (7)$$

$$W = \frac{1}{\|\nabla I\|^2 + 2\gamma} \begin{pmatrix} I_y^2 + \gamma & -I_x I_y \\ -I_x I_y & I_x^2 + \gamma \end{pmatrix}, \gamma = cte$$

## 4.2 Mise en correspondance

### → Grands déplacements

#### – Principe

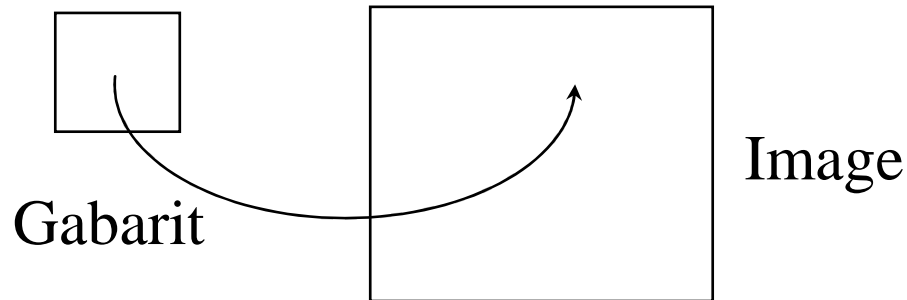
**Identifier les structures dans les images  
et les correspondances entre ces  
structures au cours du temps**

- Structures iconiques = segment binaire ou sa représentation binaire = gabarit («template») à localiser

#### **Aggarwal &al.**

- token = structures caractéristiques de l'image issues d'un traitement bas niveau (segments de bord + orientation, contraste, étendue ou blobs à ndg) dont les avantages sont:
  - invariance / illumination, perspective
  - peu nombreuxmais champ éparse.

## – Structures iconiques



→ On retrouve les différentes techniques vues dans le cadre de la minimisation directe de la DFD:

- fonctions de similarité
- seuils de similarité
- heuristiques de recherche et stratégies multi-résolution

→ et les pbs:

- info. partielles
- changement de l'orientation des gabarits
- hyp. de mouvement translationnel non-occulté

## – **Tokens: 3 approches**

- 1) Identification d'une structure de tokens et recherche d'une structure correspondante dans une autre image
- 2) Identification de 2 structures de tokens indépendamment et détermination des sous-structures correspondantes
  - théorie des graphes: établissement d'une structure relationnelle pour chaque ensemble de tokens, construction d'un graphe d'association
  - étiquetage itératif: ensemble d'hypothèses concurrentes avec prise en compte du voisinage
- 3) Idem avec mise en correspondance réalisée par minimisation d'une fonctionnelle

## – **Poursuite de Tokens**

Processus de suivi définit par:

- 1) Modèle de représentation des tokens (coord. de points, longueurs, orientations)

- 2) Modèle cinématique de l'évolution des tokens (vitesse cte, accélération cte, affine)
- 3) Un ensemble de relations entre modèle et image
- 4) un filtre temporel (Kalman)

Les avantages du suivi sont:

- trajectoires explicites des tokens
- rôle prédictif
- effet de lissage sur les paramètres du modèle
- gestion des incertitudes

**Les structures poursuivies:**

- points
- segments de contours
  - pas de groupement de tokens se déplaçant de façon cohérente
  - sensible aux occultations
- régions (+récent)

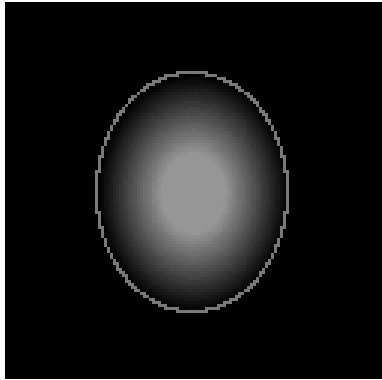
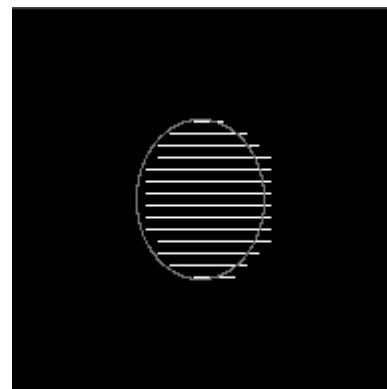
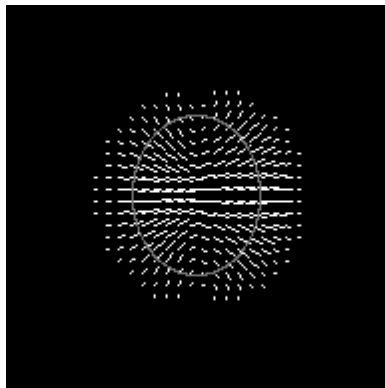


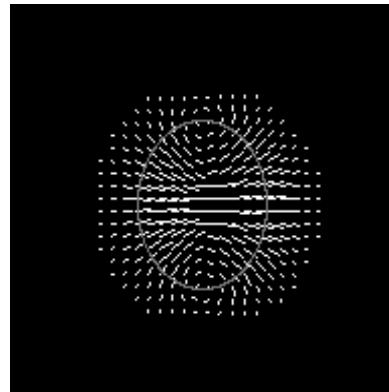
Image originale



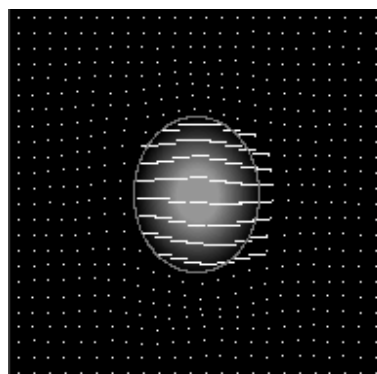
Mouvement de translation



flux optique: divergence



lissage



Multi-échelle souple [Planat 99]