Techniques du Traitement d'Images

Estimation du mouvement dans les séquences d'images

Master IA&D – TDI 2005

Antoine MANZANERA – ENSTA/LEI

Karlsruhe2.avi

TELESURVEILLANCE

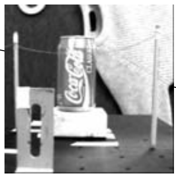
- Intrusion
- Trafic routier
- Poursuite de cibles

Large spectre d'applications

Mom_daughter_mp4.avi
+ Mom_daughter.avi

ROBOTIQUE

- Détection d'obstacles
- Ego-mouvement
- Guidage de missiles



Nasa_mp4.avi

- Compression
- Indexation

CODAGE

Plan du cours « Estimation de mouvement »

• Introduction : analyse du mouvement dans une séquence d'images

• Calcul du mouvement apparent

• Interprétation du flot optique

• Méthodes fréquentielles (transformations globales)

• Méthodes par appariement (flot optique)

• Application : codage séquences *mpeg2*

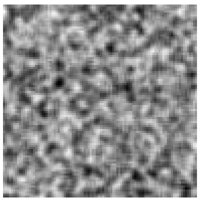
• Méthodes différentielles (flot optique)

• Approches multi-échelles

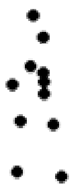
• Approches par pré-segmentation

Mouvement et perception

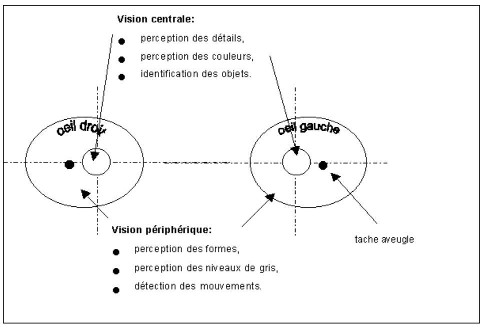
En vision, le mouvement constitue une source d'information perceptuelle très importante.



la forme par le mvt

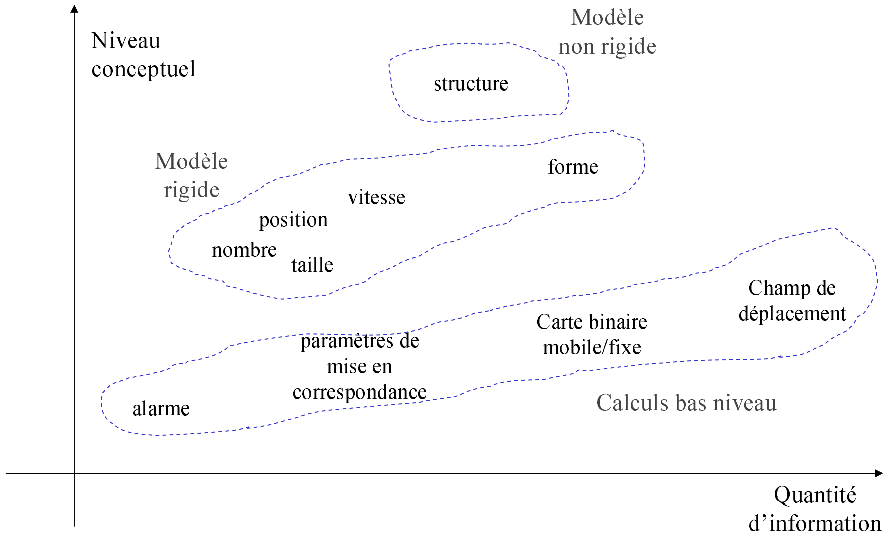


la structure par le mvt



vision périphérique des vertébrés

Quelle information extraire ?



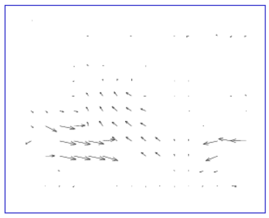
Traitement d'images

On distingue en général 3 tâches de traitement d'images :

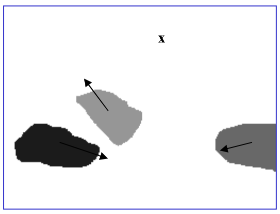
Taxi_mp4.avi

Niveau conceptuel

DETECTION



ESTIMATION

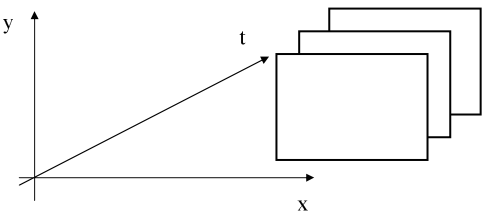


SEGMENTATION

Mouvement et problèmes inverses

PROBLEME INVERSE : Extraire des caractéristiques *physiques* (i.e. *objectives*) de l'environnement…

MAL POSE : …à partir d'une information qui ne permet de les recouvrir que partiellement.

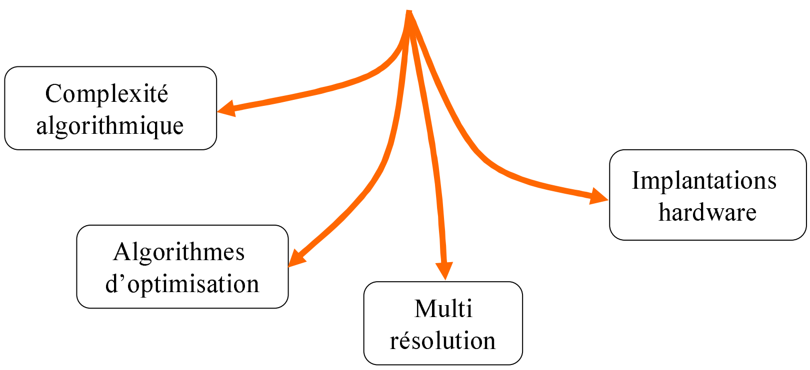


$$\underbrace{I\,(\underbrace{x\,,\,y}_{\text{espace}}\,,\,\underbrace{t}_{\text{temps}})}_{\textit{discrets}}$$

REGULARISATION $\longrightarrow$ Exploiter des connaissances *a priori* (i.e. *subjectives*)

Notions algorithmiques

En général le ***temps réel*** est ***exigé*** (même pour la compression dans le cas d'une transmission à la volée). Le temps de calcul est donc une préoccupation fondamentale dans le traitement des séquences d'images.



Calcul du mouvement apparent

(1) Le calcul d'un mouvement apparent *global* (mise en correspondance) entre deux images correspond à l'estimation des paramètres d'une transformation affectant *tous* les points de l'image : translation, rotation, homothétie, affinité,...

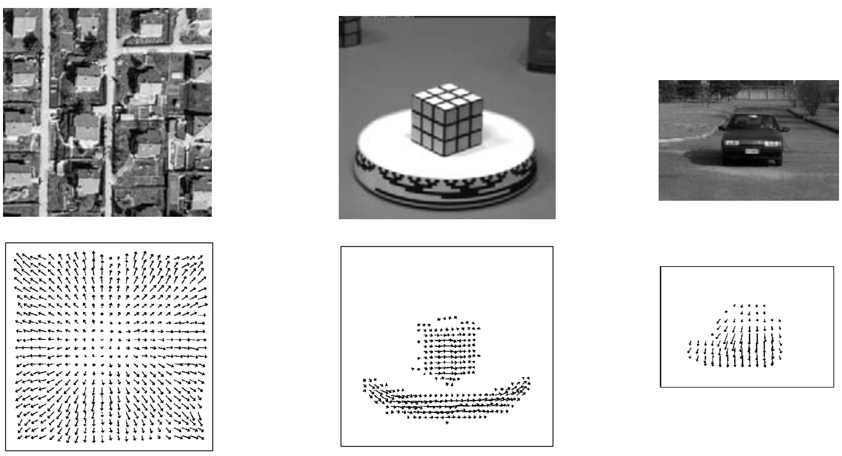
(2) Le calcul du mouvement apparent *local* consiste à associer à chaque pixel (x,y,t) de I un vecteur (v_x^t, v_y^t) représentant la *vitesse apparente* du pixel (x,y) à l'instant t.

⟶ Calcul du *flot optique* (= Champ de mouvement apparent)

Idéalement : le vecteur (v_x^t, v_y^t) représente la projection sur le plan image du vecteur vitesse (V_x^t, V_y^t, V_z^t) des objets de la scène par rapport au repère image (O,x,y,z) (grandeur objective).

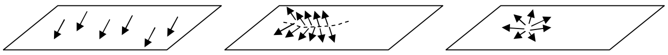
On le calcule à partir des variations temporelles de la fonction I(x,y,t).

Exemples de flot optique



Source : Pierre Kornprobst - INRIA

Quelles informations peut fournir le flot optique ?



mouvement dominant

« discontinuités », singularités

contours **foyer d'expansion (FOE)**

*** temps avant collision**

*** profondeur**

*** orientation de surface**

*** ego-mouvement**

niveau conceptuel

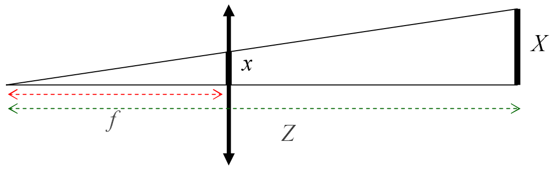
Distortion perspective

On note *(X,Y,Z)* les coordonnées d'un point de la scène.

(x,y) les coordonnées du même point projeté dans l'image

f la distance focale de la caméra

Distortion perspective (modèle sténopé) :



$$x = f\frac{X}{Z}$$

$$y = f\frac{Y}{Z}$$

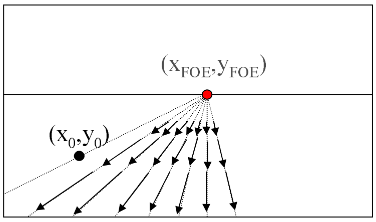
Le foyer d'expansion (FOE)

Lors d'un déplacement de la caméra dans une scène statique, les directions de vitesse des points projetés sur le plan image converge vers un point du plan projectif appelé FOE.

Si la caméra se déplace à la vitesse (-X',-Y',-Z'), alors tous les points de la scène sont à la même vitesse (X',Y',Z') (avec X' = dX/dt ; Y' = dY/dt ; Z' = dZ/dt).

Soit (X_0,Y_0,Z_0) un point de la scène. Après un temps t, il est projeté sur l'image au point (x_t,y_t), avec :

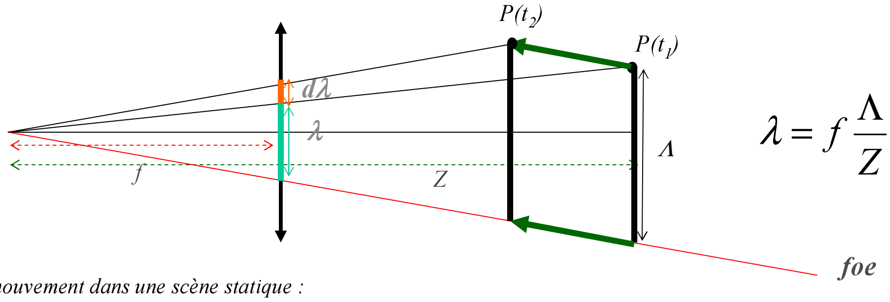
$$(x_t, y_t) = \left(f\frac{X_0 + tX'}{Z_0 + tZ'}, f\frac{Y_0 + tY'}{Z_0 + tZ'} \right)$$



$$(x_{FOE}, y_{FOE}) = \lim_{t \to -\infty} (x_t, y_t)$$

$$= \left(f\frac{X'}{Z'}, f\frac{Y'}{Z'} \right)$$

Temps avant collision



$$\lambda = f\frac{\Lambda}{Z}$$

mouvement dans une scène statique :

$$\lambda' = -\frac{f\Lambda}{Z^2}Z'$$

et donc

$$\boxed{\frac{\lambda}{\lambda'} = -\frac{Z}{Z'}}$$

temps avant collision

Profondeur et coordonnées scène

• Connaissant la profondeur d'un point de la scène on peut déterminer celle de tout point qui a la même vitesse :

$$Z_2 = \frac{Z_1 \lambda_2 \lambda_1'}{\lambda_2' \lambda_1}$$

• Calcul des coordonnées scène :

$$\begin{cases} Z = \dfrac{Z'\lambda}{\lambda'} \\ Y = \dfrac{yZ'\lambda}{f\lambda'} \\ X = \dfrac{xZ'\lambda}{f\lambda'} \end{cases}$$

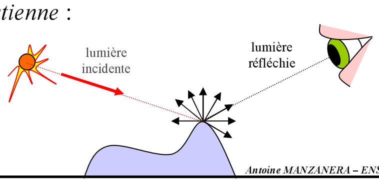
Calcul du flot optique : limites et contraintes

(1) On suppose :

MOUVEMENT ⇔ VARIATION D'INTENSITE

(⇐) On suppose que l'observation d'une variation d'intensité dans l'image traduit nécessairement l'existence d'un mouvement dans la scène. Cela correspond à une hypothèse *d'éclairement constant* sur la scène.

(⇒) C'est la variation de l'intensité dans l'image qui doit permettre de retrouver le mouvement apparent de l'objet dans la scène. On suppose donc que la lumière réfléchie par un point de la scène reste constante indépendamment du mouvement relatif scène / caméra. L'hypothèse sous-jacente est que les objets sont à surface *lambertienne* :



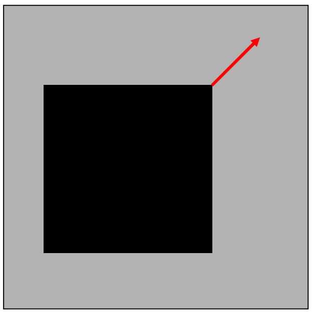
L'intensité de la lumière réfléchie est la même dans toutes les directions.

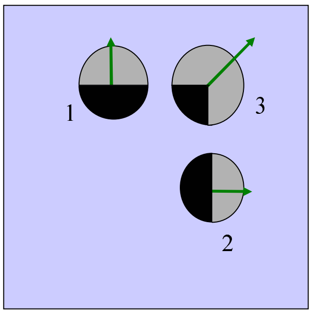
Calcul du flot optique : limites et contraintes

(2) ***PROBLEME DE L'OUVERTURE***

On n' « accède » au mouvement apparent d'un point que grâce à un calcul effectué *dans un voisinage borné* de ce point.

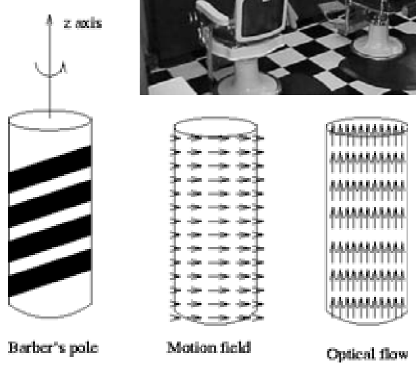
→ On ne peut calculer que la composante du mouvement *dans la direction du gradient* (i.e. *perpendiculaire au contour*).





Problème de l'ouverture

Un exemple classique de manifestation du problème de l'ouverture : *L'enseigne de barbier*.



Techniques fréquentielles

Les techniques fréquentielles d'estimation du mouvement entre deux images sont fondées sur l'équivalence translation/déphasage de la transformée de Fourier :

Rappel : l'expression d'une image dans le domaine fréquentiel consiste à décomposer la fonction bidimensionnelle en sommes de sinusoïdes complexes :

$$I(x,y) = \frac{1}{wh}\sum_{u=0}^{w-1}\sum_{v=0}^{h-1} F(u,v)e^{2i\pi(ux+vy)/wh}$$

Transformée de Fourier discrète inverse

↕

Les coefficients des différentes sinusoïdes sont calculés par la transformée de Fourier :

$$F(u,v) = \sum_{x=0}^{w-1}\sum_{y=0}^{h-1} I(x,y)e^{-2i\pi(ux+vy)/wh}$$

Transformée de Fourier discrète directe

Notation (module,phase) : $F(u,v) = \|F(u,v)\| e^{i\varphi_F(u,v)}$

La propriété de translation/déphasage dit que si F est la transformée de Fourier de I :

Alors la TF de I translatée de ($-\delta x,-\delta y$), est G, avec :

$$I(x,y) \xrightarrow{\text{TF}} F(u,v)$$

$$I(x+\delta x, y+\delta y) \xrightarrow{\text{TF}} G(u,v) = F(u,v)e^{2i\pi(u\delta x+v\delta y)/wh}$$

Soit : $\|G(u,v)\| = \|F(u,v)\|$ et : $\varphi_G(u,v) = \varphi_F(u,v) + 2\pi(u\delta x + v\delta y)/wh$

Le déphasage entre F et G vaut donc : $\Delta\phi(u,v) = 2\pi(u\delta x + v\delta y)/wh$

Il suffit donc en théorie de considérer ce déphasage pour 2 couples (u,v) pour calculer ($\delta x, \delta y$), mais cette technique est sensible au bruit et aux changement d'illumination qui induisent des variations dans les basses fréquences.

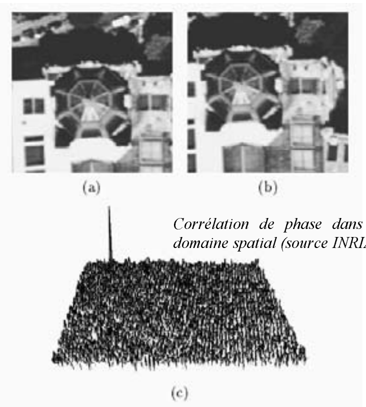
On utilise plutôt la technique de corrélation de phase.

Corrélation de phase

La technique de corrélation de phase exploite une conséquence directe de la propriété de translation/déphasage. Si F est la TF de I et G la TF de I translatée de ($-\delta x,-\delta y$), alors le déphasage entre F et G est égal à leur spectre de puissance croisé normalisé (SPCN), i.e. :

$$\frac{F^*(u,v)G(u,v)}{\|F^*(u,v)G(u,v)\|} = e^{2i\pi(u\delta x+v\delta y)/wh}$$

La TF inverse du SPCN est donc égale à la fonction de Dirac du vecteur de translation : $\delta_{(\delta x,\delta y)}(x,y)$



Corrélation de phase dans le domaine spatial (source INRIA)

La technique de corrélation de phase consiste donc à :

1. Calculer les TF de $I(x,y,t)$ et $I(x,y,t+1)$, soit F_1 et F_2
2. Calculer χ le SPCN de $F1$ et F_2
3. Calculer D la TF inverse de χ
4. Rechercher le maximum de D

Avantages et inconvénients

+ Robuste car toutes les fréquences contribuent au calcul

+ Relativement rapide grâce au calcul de la FFT

- En pratique limité à un déplacement global sur toute l'image

Invariant de Fourier-Mellin

L'utilisation de la transformée de *Fourier-Mellin* permet de calculer les paramètres d'une similitude (*rotation et homothétie*) comme un *vecteur de translation* de manière analogue au cas précédent, grâce à une représentation log-polaire de l'espace des fréquences $(u,v) \rightarrow (\theta, \log \rho)$:

Soit g l'image transformée de f par une rotation d'angle α, une homothétie de rapport σ, et une translation de vecteur (x_0,y_0) :

$$g(x,y) = f\big(\sigma(\cos\alpha x + \sin\alpha y) - x_0, \sigma(-\sin\alpha x + \cos\alpha y) - y_0\big)$$

Les amplitudes des transformées de Fourier de f et g sont liées par la relation suivante :

$$\|G(u,v)\| = \frac{1}{\sigma^2}\left\|F\left(\tfrac{1}{\sigma}(u\cos\alpha + v\sin\alpha), \tfrac{1}{\sigma}(-u\sin\alpha + v\cos\alpha)\right)\right\|$$

donc l'amplitude :
- ne dépend pas de la translation (x_0,y_0).
- subit une rotation d'angle α.
- subit une modification d'échelle d'un facteur $1/\sigma$.

En passant les fréquences en coordonnées polaires :

$$F_p(\theta,\rho) = \|F(\rho\cos\theta, \rho\sin\theta)\|; 0 \le \theta \le 2\pi, 0 \le \rho < \infty$$
$$G_p(\theta,\rho) = \|G(\rho\cos\theta, \rho\sin\theta)\|; 0 \le \theta \le 2\pi, 0 \le \rho < \infty$$

on obtient :

$$G_p(\theta,\rho) = \frac{1}{\sigma^2} F_p\left(\theta - \alpha, \frac{\rho}{\sigma}\right)$$

Enfin, en passant la coordonnée radiale au logarithme :

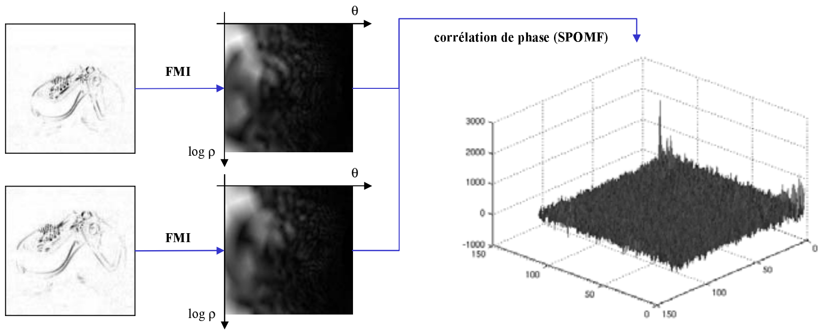
$$r = \log\rho \qquad F_{lp}(\theta,r) = F_p(\theta,\rho)$$
$$s = \log\sigma \qquad G_{lp}(\theta,r) = G_p(\theta,\rho)$$

on obtient :

$$\boxed{G_{lp}(\theta,r) = \frac{1}{\sigma^2} F_{lp}(\theta - \alpha, r - s)}$$

Donc une similitude dans l'espace image se traduit par une *translation* dans l'espace des *fréquences log-polaires*.

Invariant de Fourier-Mellin



Un exemple d'utilisation de la transformée de Fourier-Mellin : calcul de la position de la tête des Robots Aibo dans l'image par corrélation de phase des invariants de Fourier-Mellin. (FMI-SPOMF : Fourier-Mellin Invariant Symmetric Phase Only Matched Filtering) : ***J.C. Baillie et M. Nottale*** 2004.

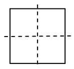
⚠ L'information de phase de l'image originale est perdue dans la FMI. Le FMI-SPOMF revient à chercher la meilleure (rotation, homothétie) qui mette en correspondance 2 spectres d'amplitude. *On ne retrouve donc pas les paramètres de translation entre les deux images, et de plus l'information de forme portée par la phase n'existe plus.*

Pour compléter cette transformation, on peut appliquer une corrélation de phase classique sur le couple d'image de départ, après avoir appliqué sur l'une des images la transformation (rotation, homothétie) fournie par le FMI-SPOMF.

Notons enfin, que comme pour la corrélation de phase, cette méthode est utilisée en pratique pour estimer des transformations *globales*, car elle utilise la contribution de tout le spectre (ou au moins une large partie), ce qui implique une étendue spatiale importante des pixels utilisés pour l'estimation de chaque transformation.

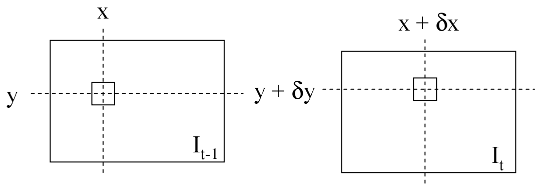
Technique par appariement

Mesure d'appariement :

On considère $B \subset \mathbf{Z}^2$ voisinage de l'origine : 

$$A^t_{(x,y)}(\delta x, \delta y) = \sum_{(b_1,b_2)\in B} \left(I(x+b_1, y+b_2, t-1) - I(x+\delta x+b_1, y+\delta y+b_2, t)\right)^2$$

C'est la somme des différences au carrée de 2 blocs :



typiquement : B de taille 9×9, 15×15...

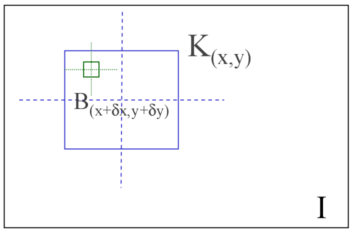
Solution du flot optique :

$$(v^t_x, v^t_x) = \arg \min_{(\delta x, \delta y)\in K} A^t_{(x,y)}(\delta x, \delta y)$$

$K \subsetneq I$

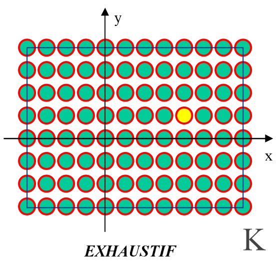
pour limiter le temps de calcul

+ stratégies d'optimisation…

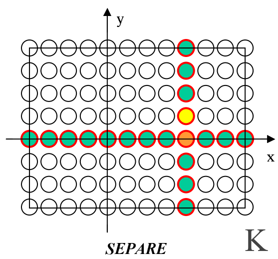


Appariement : stratégies d'optimisation

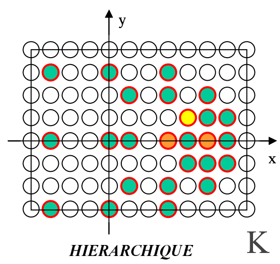
Le principe des stratégies d'optimisation est de parcourir « intelligemment » l'espace de recherche K pour minimiser la complexité de calcul de l'optimum d'appariement :



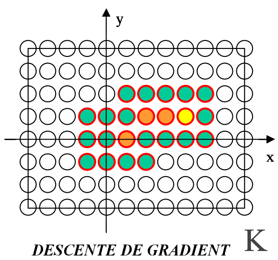
EXHAUSTIF



SEPARE



HIERARCHIQUE



DESCENTE DE GRADIENT

Pour chaque point de K, on ne calcule la somme des différences que tant qu'on demeure inférieur à un seuil *T*. Le meilleur déplacement correspond à celui pour lequel le plus de points ont été examinés.

SEQUENTIEL

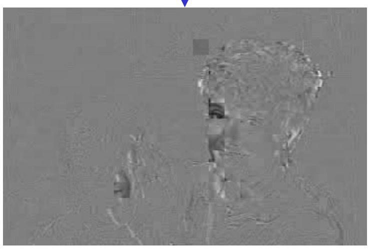
En l'absence d'hypothèse sur la fonction d'appariement (séparabilité, convexité...) dans le domaine K, seul le parcours exhaustif garantit le calcul d'un optimum *global* sur K.

Exemple : norme mpeg2 de codage video

Flot optique sous-résolu

- Découpage en blocs 16×16 ou 32×32.
- Hypothèse de même déplacement pour tous les pixels du blocs.
- Codage : déplacements + erreurs comprimées.
  - Compression statique *JPEG* :
    - Transformée en cosinus discrète
    - Quantification
    - Codage entropique

Trame prédite



Erreurs de prédiction

+ démo soft
VcDemo
TU Delft

Norme mpeg4 de codage video

OBJETS VIDEO

Segmentation au sens du mouvement :

✓ Codage de la forme

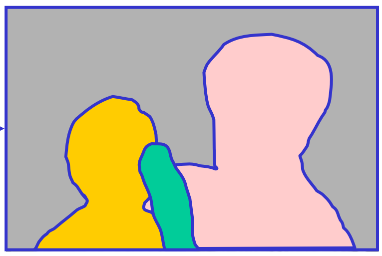
Calcul du masque des objets mobiles

✓ Codage du mouvement

Estimations de mouvement indépendantes

✓ Codage du fond

- statique
- panoramique (mode *sprite*)



Techniques différentielles

HORN & SCHUNCK 1981

Exploitation directe de la contrainte (1) :

$$I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$$

Formule de Taylor

$$= I(x,y,t) + \frac{\partial I}{\partial x}\cdot\delta x + \frac{\partial I}{\partial y}\cdot\delta y + \frac{\partial I}{\partial t}\cdot\delta t + \ldots$$

Termes d'ordres supérieures négligés

SOIT : $\boxed{\nabla I \cdot v + \dfrac{\partial I}{\partial t} = 0}$

avec

$$\begin{cases} \nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}) & \text{gradient spatial} \\ v = (v_x^t, v_y^t) & \text{inconnues} \\ \frac{\partial I}{\partial t} & \text{gradient temporel} \end{cases}$$

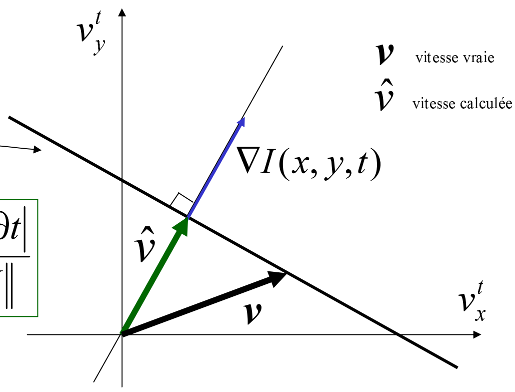
Equation de contrainte du mouvement apparent (ECMA)

ou : équation du flot optique

Interprétation de l'ECMA

(1) En faisant l'hypothèse d'une certaine *régularité* du champ et de *petits déplacements*, les *changements temporels* dans l'image sont *équivalents* (au premier ordre) au produit scalaire des *changements spatiaux* et de la *vitesse apparente*.

(2) Droite de contrainte :



pb d'ouverture

≡

sous-contrainte (1 eq°, 2 inc.)

$$\boxed{\|\hat{v}\| = \frac{|\partial I/\partial t|}{\|\nabla I\|}}$$

Résolution de l'ECMA

HORN & SCHUNCK 1981

Résolution de l'ECMA par *ajout d'une contrainte de régularité*.

⇔ *Régularisation* du pb mal posé par hypothèse de champ *lisse* de déplacement.

→ Minimisation d'une fonction de coût :

$$C^t_{(x,y)}(v^t_x, v^t_y) = \underbrace{\left(\nabla I \cdot v + \frac{\partial I}{\partial t}\right)^2}_{\text{ECMA}} + \lambda \underbrace{\left[\left(\frac{\partial v^t_x}{\partial x}\right)^2 + \left(\frac{\partial v^t_x}{\partial y}\right)^2 + \left(\frac{\partial v^t_y}{\partial x}\right)^2 + \left(\frac{\partial v^t_y}{\partial y}\right)^2\right]}_{\text{REGULARISATION}}$$

facteur de pondération

Résolution de l'ECMA

HORN & SCHUNCK 1981

1. Minimisation d'une fonctionnelle quadratique :

$$C_{(x,y)}(u,v) = \left(I_x u + I_y v + I_t\right)^2 + \lambda\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right)$$

avec :
- (u,v) composantes du champ (inconnues) à calculer
- $\{I_x, I_y, I_t\}$ dérivées partielles de l'image
- $\{u_x, u_y, v_x, v_y\}$ dérivées partielles des composantes du champ

→ Annulation des dérivées premières $\partial/\partial u(\ldots) = 0$; $\partial/\partial v(\ldots) = 0$

Equations d'Euler-Lagrange de minimisation de C :

$$\begin{cases} 2\left(I_x u + I_y v + I_t\right)I_x + 2\lambda\left[\left(\frac{\partial}{\partial u}\frac{\partial u}{\partial x}\right)\frac{\partial u}{\partial x} + \left(\frac{\partial}{\partial u}\frac{\partial u}{\partial y}\right)\frac{\partial u}{\partial y}\right] = 0 \\ 2\left(I_x u + I_y v + I_t\right)I_y + 2\lambda\left[\left(\frac{\partial}{\partial v}\frac{\partial v}{\partial x}\right)\frac{\partial v}{\partial x} + \left(\frac{\partial}{\partial v}\frac{\partial v}{\partial y}\right)\frac{\partial v}{\partial y}\right] = 0 \end{cases}$$

soit : $\begin{cases} \left(I_x u + I_y v + I_t\right)I_x + \lambda\Delta u = 0 \\ \left(I_x u + I_y v + I_t\right)I_y + \lambda\Delta v = 0 \end{cases}$ avec : $\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \\ \Delta v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \end{cases}$ laplaciens de u et v

Résolution de l'ECMA

2. Approximation du Laplacien : $\nabla^2 f = f - \tilde{f}$ $\quad \tilde{f}$ = moyenne de f dans un certain voisinage

soit : $\begin{cases} (v-\tilde{v})\left(I_x^2+I_y^2+\lambda\right) = -I_y\left(I_x\tilde{u}+I_y\tilde{v}+I_t\right) \\ (u-\tilde{u})\left(I_x^2+I_y^2+\lambda\right) = -I_x\left(I_x\tilde{u}+I_y\tilde{v}+I_t\right) \end{cases}$

3. Schéma itératif de résolution :

Méthode de Gauss-Seidel : $\begin{cases} u^k = \tilde{u}^{k-1} - I_x\left(I_x\tilde{u}^{k-1}+I_y\tilde{v}^{k-1}+I_t\right)\cdot\left(I_x^2+I_y^2+\lambda\right)^{-1} \\ v^k = \tilde{v}^{k-1} - I_y\left(I_x\tilde{u}^{k-1}+I_y\tilde{v}^{k-1}+I_t\right)\cdot\left(I_x^2+I_y^2+\lambda\right)^{-1} \end{cases}$

Soit, en reprenant les notations originales :

ALGORITHME DE HORN & SCHUNCK

avec :

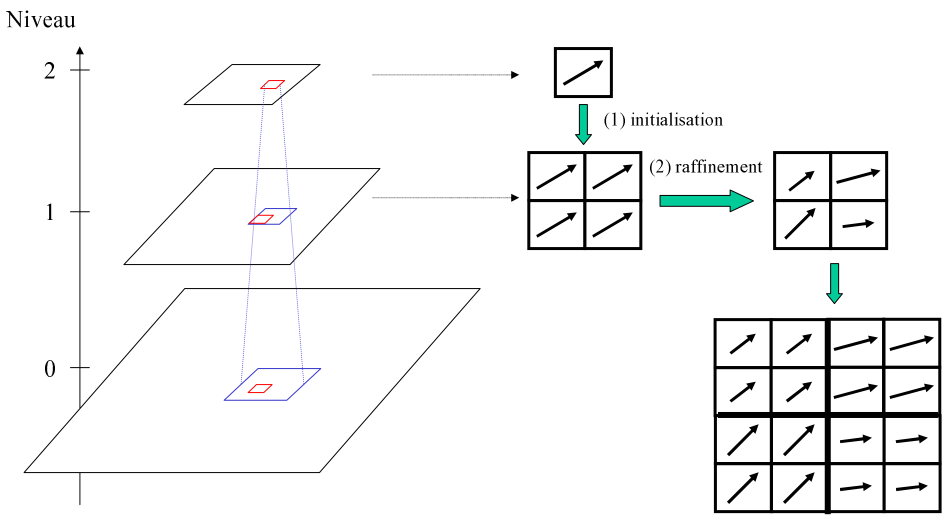
$$\begin{cases} N = \dfrac{\partial I}{\partial x}\tilde{v}_x^{k-1} + \dfrac{\partial I}{\partial y}\tilde{v}_y^{k-1} + \dfrac{\partial I}{\partial t} \\ D = \lambda + \left(\dfrac{\partial I}{\partial x}\right)^2 + \left(\dfrac{\partial I}{\partial y}\right)^2 \end{cases}$$

- Initialisation : $v_x^0 = 0 \quad v_y^0 = 0$
- Répéter jusqu'à convergence :

$$v_x^k = \tilde{v}_x^{k-1} - \frac{\partial I}{\partial x}\frac{N}{D}$$

$$v_y^k = \tilde{v}_y^{k-1} - \frac{\partial I}{\partial y}\frac{N}{D}$$

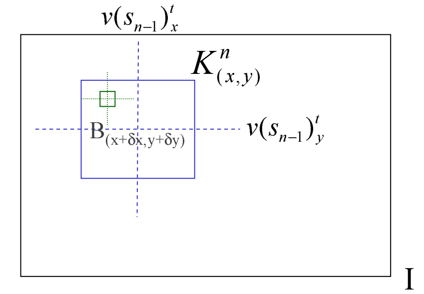
Estimation multi-échelles



Principe du calcul multi-échelles

ALGORITHME PAR CORRELATION

$$\left(v(s_n)_x^t, v(s_n)_y^t\right) = \arg \max_{(\delta x, \delta y) \in K^n} C^t_{(x,y)}(\delta x, \delta y)$$



avec : $K^n_{(x,y)} = \left[v(s_{n-1})_x^t \pm \Delta_1 \right] \times \left[v(s_{n-1})_y^t \pm \Delta_2 \right]$

ALGORITHME DE HORN & SCHUNCK

- Initialisation : $v(s_0)_x^0 = 0 \; v(s_0)_y^0 = 0$
- Répéter jusqu'à convergence :

$$v(s_0)_x^k = \tilde{v}(s_0)_x^{k-1} - \frac{\partial I}{\partial x} \frac{N}{D}$$

$$v(s_0)_y^k = \tilde{v}(s_0)_y^{k-1} - \frac{\partial I}{\partial y} \frac{N}{D}$$

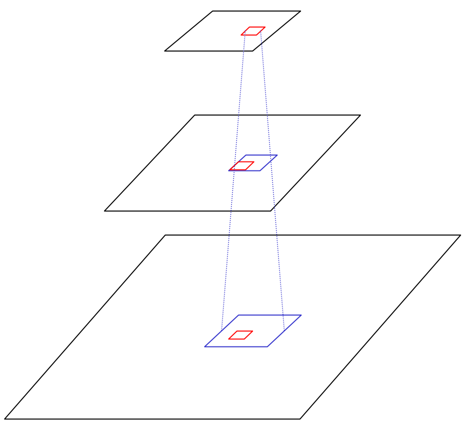
- Init. : $v(s_1)_x^0 = v(s_0)_x^\infty \; v(s_1)_y^0 = v(s_0)_y^\infty$
- Répéter jusqu'à convergence :

$$v(s_1)_x^k = \tilde{v}(s_1)_x^{k-1} - \frac{\partial I}{\partial x} \frac{N}{D}$$

$$v(s_1)_y^k = \tilde{v}(s_1)_y^{k-1} - \frac{\partial I}{\partial y} \frac{N}{D}$$

…/…

Avantages des calculs multi-échelles

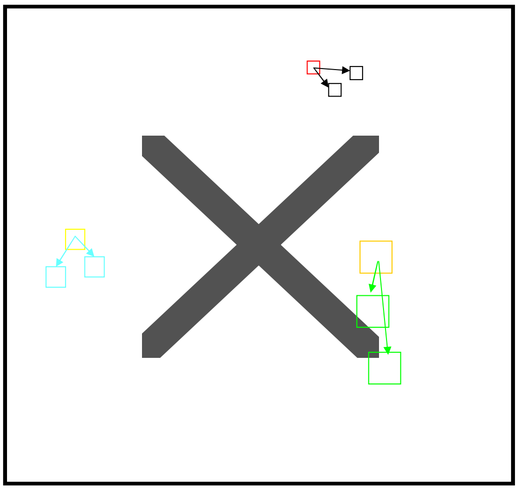


- ✓ Agrandir le domaine des vitesses calculables
- ✓ Diminuer le temps de calcul
- ✓ Prévoir les inconsistances
- ✓ Régulariser de manière implicite
- ✓ Résoudre les problèmes d'aliasing temporel

Calculs multi-échelles et aliasing temporel

L'aliasing temporel correspond à l'indécidabilité du problème de mise en correspondance de motifs dans le cas de structures périodiques lorsque l'amplitude du déplacement est supérieure à celle des motifs.

Ce problème se rencontre à différentes échelles dans l'image :



Calculs multi-échelles et aliasing temporel

Pyramide Gaussienne

Approches par pré-segmentation

Il existe classiquement 2 approches de l'analyse du mouvement :

Approche « *traiter après* » — luminance → • différence • calcul du flot optique → extraction de caractéristiques → ANALYSE

PIXEL-BASED METHOD

OU

Approche « *traiter avant* » — luminance → extraction de caractéristiques → mise en correspondance → ANALYSE

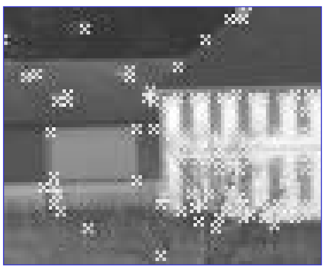
FEATURE-BASED METHOD

Approches par pré-segmentation

Dans le cadre du calcul du flot optique, on cherche un compromis entre les deux approches « pixel-based » et « feature-based ».

En effet, pour l'extraction de certaines caractéristiques du flot optique, il est intéressant d'avoir un champ dense (pixel-based), mais une pré-segmentation permet de différencier le calcul spatialement (feature-based), pour éviter l'apparition et la propagation des erreurs.

Exemple de pré-segmentation : ***Calcul des points d'intérêt***



Exemple de calcul de points d'intérêt

ESPACE D'ECHELLE GAUSSIEN

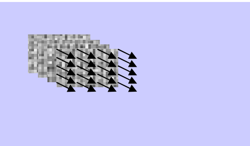
ESPACE D'ECHELLE MORPHOLOGIQUE

Points d'intérêt et flot optique

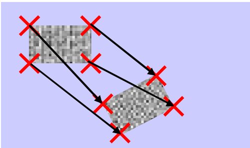
Lien avec le flot optique : *RESTREINDRE*, ou *DIFFERENCIER* le calcul du flot optique pour les points d'intérêt.

Lien avec la vision biologique :

2 types de perception du mouvement bien distincts chez les primates :



petits mouvements → analyse du flux optique



grands mouvements → mise en correspondance de points caractéristiques

Points d'intérêt et flot optique

Avantages ***algorithmiques*** de la pré-segmentation par point d'intérêt :

✓ problème bien posé au voisinage des points d'intérêt

$\longleftrightarrow$ *consistance du flot*

✓ réduction du temps de calcul

$\longleftrightarrow$ *diminution du domaine de définition*

✓ optimisation convexe

$\longleftrightarrow$ *algorithmique gloutonne, implantation hardware, ...*

Bibliographie du chapitre

• *B.K.P Horn & B. Schunck 1974 « **Determining Optical Flow** » Artificial Intelligence **23 185-203***

• *D.H. Ballard & C.M Brown 1982 « **Computer Vision** » Prentice Hall* **(Ch. 3, Ch. 7)**

• *T.S Huang 1983 « **Image sequences processing and dynamic scene analysis**» Springer Verlag NATO ASI Series*

• *J.M. Jolion & A. Rosenfeld 1994 « **A pyramid framework for early vision** » Kluwer Academic Publishers **Dordrecht, NL***

• *R. Jain, R. Kasturi, & B. Schunck 1995 « **Machine Vision** » McGraw-Hill Inc.* **(Ch.14)**