

POURSUITE DE CIBLE.

1• QU'EST-CE QUE LA POURSUITE DE CIBLE ?

Lorsqu'un objet bouge devant une caméra ou lorsque la caméra bouge, il se produit une variation de l'éclairage du capteur d'image (réteine) modifiant la distribution des niveaux de gris (voir Figure 1).

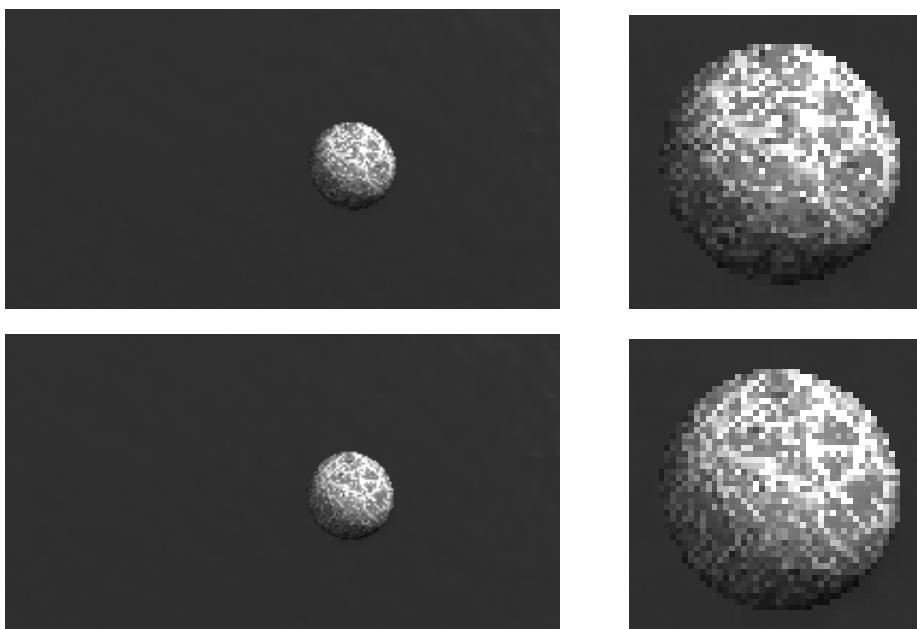


Figure 1 : deux images consécutives de la séquence "Boule de marbre" et détails. Cette modification des niveaux de gris porte le nom de "mouvement apparent". Le mouvement apparent est l'effet que produit un mouvement réel sur la distribution de l'éclairage projetée.

La poursuite de cible est une technique permettant de trouver, dans une séquence d'images, la position d'une zone prédefinie correspondant à un motif particulier dont on souhaite connaître la position.

Les utilisations de la poursuite de cible sont multiples :

- accrochage d'un objet virtuel dans une image,
- caractérisation du fonctionnement d'un organe en imagerie médicale,
- vidéo surveillance automatique,
- ...

Pour réaliser une poursuite de cible, il faut être capable de retrouver le mouvement projeté de la cible à partir des variations des niveaux de gris de la séquence d'images. Plusieurs méthodes peuvent être utilisées pour ce faire qui vous sont détaillées dans le paragraphe suivant.

2• MÉTHODES DE MISE EN CORRESPONDANCE DE MOTIF.

2.1• Principe.

Le principe de la mise en correspondance de motif repose sur une mesure de similarité entre deux images (ici il s'agira d'images extraites, l'une étant le motif à rechercher et l'autre étant une sous-image extraite de l'image courante ayant les mêmes dimensions que le motif). Ces mesures de similarités sont basées soit sur des distances statistiques, soit sur des mesures de corrélation.

2.2• Corrélation.

La corrélation R des motifs de deux images $I_1(x, y)$ et $I_2(x, y)$ est donnée par :

$$R = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I_1(x, y) I_2(x, y),$$

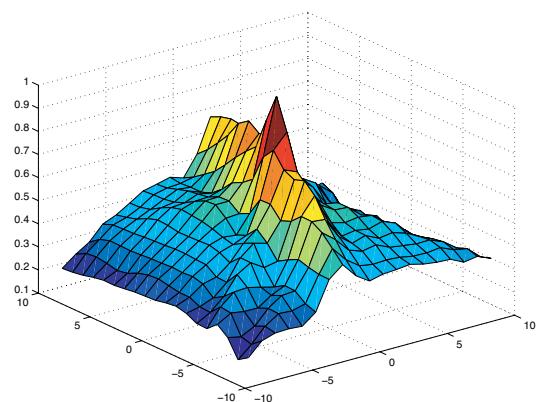
où N_x, N_y est la dimension commune des deux images.

Pour s'affranchir des problèmes de variation lumineuses dus aux ajustements automatiques des capteurs d'images modernes, on lui préfère généralement sa version centré et normalisée, aussi appelée "coefficient de corrélation de Pearson" :

$$\rho = \frac{1}{N_x N_y} \frac{\sum_{x=1}^{N_x} \sum_{y=1}^{N_y} (I_1(x, y) - \mu_1)(I_2(x, y) - \mu_2)}{\sigma_1 \sigma_2},$$

où μ_1 est la moyenne de l'image $I_1(x, y)$: $\mu_1 = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I_1(x, y)$, μ_2 la moyenne de l'image $I_2(x, y)$, σ_1 l'écart-type de l'image $I_1(x, y)$:

$$\sigma_1^2 = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} (I_1(x, y) - \mu_1)^2, \text{ et } \sigma_2 \text{ l'écart-type de l'image } I_2(x, y).$$



L'utilisation de la corrélation pour la poursuite de motif consiste à rechercher, sur l'image courante, le sous-motif dont la corrélation est maximale avec le motif à poursuivre. Il faut donc faire subir à l'image courante un ensemble de transformations possibles et retenir celle qui permet d'avoir la meilleure corrélation.

2.3• Distances statistiques.

Il est aussi possible de mesurer une distance entre les deux motifs (les deux sous-

images) à comparer.

Ces distances portent le nom de SAD (somme absolue des distances) :

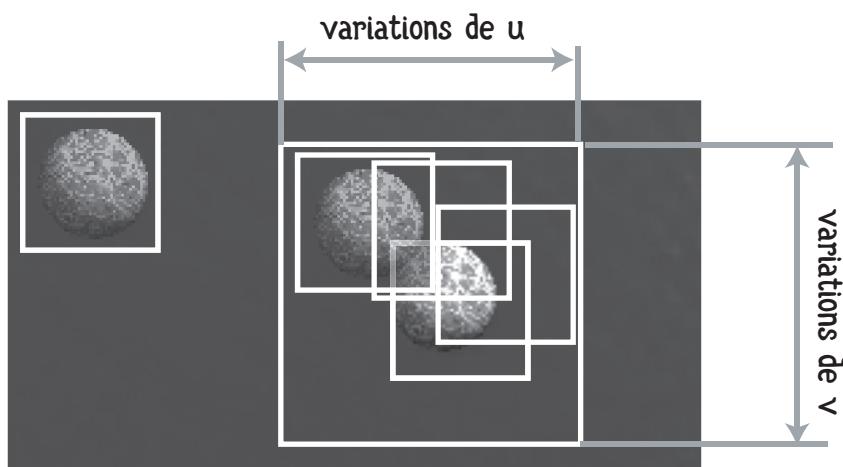
$$SAD = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} |I_1(x, y) - I_2(x, y)|$$

et SSD (somme des écarts quadratiques : summ of square distances) :

$$SSD = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} |I_1(x, y) - I_2(x, y)|^2$$

Il existe pour ces deux distances des versions normalisées centrées comme pour la corrélation.

2.4• D'un point de vue pratique.



En pratique, on suppose que le mouvement du motif entre deux images ne peut excéder un certain rayon. On va donc utiliser cette hypothèse pour réduire l'espace de recherche. Si on appelle $M(x, y)$ le motif recherché dans l'image courante et $(\delta x, \delta y)$ la position du motif dans l'image précédente, le principe consiste à trouver les valeurs de (u, v) les plus probables au sens où le motif extrait de l'image courante en $(\delta x + u, \delta y + v)$ soit le plus corrélé ou le plus proche du motif initial ($M(x, y)$).

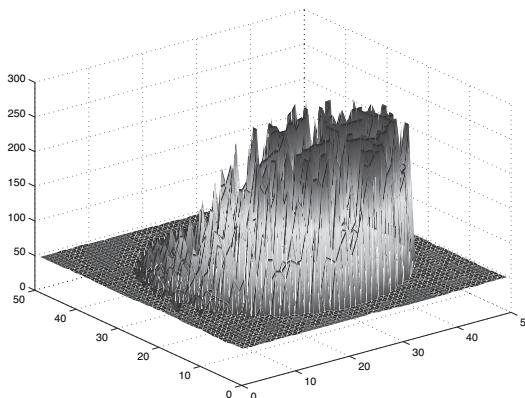
Il vous faut donc créer un tableau des valeurs de corrélation (ou de distance) pour chaque valeur de u et v et prendre, pour chaque image de la séquence, celle qui maximise la corrélation ou minimise la distance.

Naturellement, plus le motif extrait est grand, et plus l'espace de recherche est important, plus le temps de calcul sera long.

3• FLOT OPTIQUE.

La méthode du flot optique consiste à supposer que le motif est une fonction discrète à deux dimensions produite par la discréétisation (par le capteur d'images) de l'illumination locale.

3.1• Rappels.



Soit $I(x,y,t)$ la répartition de l'illumination sur la fenêtre courante. Si le motif est correctement recalé, $I(x,y,t)$ est aussi la fonction d'illumination du motif.

A l'instant suivant le motif s'est déplacé de (dx, dy) . Soit dt le laps de temps écoulé entre deux acquisitions d'images.

Si le motif a toujours la même répartition d'illumination, alors $I(x+dx, y+dy, t+dt) = I(x, y, t)$.

La méthode du flot optique consiste à faire une expansion au premier ordre de cette équation :

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

ou O^2 regroupe les termes d'ordre supérieurs.

Et comme la conservation de l'illumination impose $I(x+dx, y+dy, t+dt) = I(x, y, t)$, il vient :

$$\frac{\partial}{\partial x} I(x, y, t) dx + \frac{\partial}{\partial y} I(x, y, t) dy + \frac{\partial}{\partial t} I(x, y, t) dt \approx 0$$

ou encore :

$$\frac{\partial}{\partial x} I(x, y, t) \frac{dx}{dt} + \frac{\partial}{\partial y} I(x, y, t) \frac{dy}{dt} + \frac{\partial}{\partial t} I(x, y, t) \approx 0$$

$\frac{\partial}{\partial t} I(x, y, t)$, c'est la variation de l'illumination au point (x, y) entre les deux instant t et $t+dt$. C'est donc (approximativement) $\frac{\partial}{\partial t} I(x, y, t) = I_t(x, y, t) = I(x, y, t+dt) - I(x, y, t)$.

Quant aux quantités $I_x(x, y) = \frac{\partial}{\partial x} I(x, y, t)$ et $I_y(x, y) = \frac{\partial}{\partial y} I(x, y, t)$, ce sont des valeurs qui peuvent être calculées a priori sur l'image grâce à un opérateur que l'on nomme "extracteur de gradient". L'extracteur de gradient permet d'obtenir, à partir d'une image $I(x, y, t)$, les valeurs estimées des deux images $I_x(x, y)$ et $I_y(x, y)$ en chaque point du motif.

1	2	3	4	5	6
7	...				
				35	36

Si on suppose que tous les pixels d'un motif sont numérotés comme ci-contre (on suppose ici que le motif est $(6, 6)$ et on a donc 36 pixels de coordonnées

(x_i, y_i) i de 1 à 36). Si on suppose que tous les points du motif subissent le même déplacement (u, v) , alors on dispose, pour identifier (u, v) de 36 équations de flot optique pour deux inconnues :

$$\begin{bmatrix} I_x(x_1, y_1) & I_y(x_1, y_1) \\ \dots & \dots \\ I_x(x_{36}, y_{36}) & I_y(x_{36}, y_{36}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(x_1, y_1, t) \\ \dots \\ I_t(x_{36}, y_{36}, t) \end{bmatrix}$$

Cette équation est de la forme : $A \begin{bmatrix} u \\ v \end{bmatrix} = B$. Elle peut être résolue par moindres carrés

en calculant A^\dagger la pseudo-inverse de A . Comme A ne varie pas, on peut précalculer A^\dagger .

Il est préférable de calculer A^\dagger sous son format explicite : $A^\dagger = (A^T A)^{-1} A^T$ tant pour des raisons de temps de calcul que pour vérifier que la pseudo inverse calculée correspond bien à une minimisation des moindres carrés.

La méthode de calcul est alors la suivante :

- à chaque image on extrait l'imagette qui se trouve juste sous la position du motif sous l'image courante.
- on calcule la différence des intensités entre les deux imagettes (le motif et l'image extraite)
- on remplit le vecteur B
- on multiplie B par A^\dagger et on obtient $\begin{bmatrix} u \\ v \end{bmatrix}$.
- on remet à jour la position de la fenêtre définissant le motif et on passe à l'image suivante.

Si on calcule le nombre d'opérations élémentaires que nécessite cette méthode, on s'aperçoit qu'elle est beaucoup plus rapide. Elle est donc avantageuse lorsqu'il faut réduire les temps de calculs. Par contre elle ne peut fonctionner que si l'approximation différentielle est valide, c'est à dire si le mouvement est faible entre les deux images. Inconvénient que ne présentent pas les méthodes de corrélation et de distance.

3.2. D'un point de vue pratique.

Le problème dans la méthode de flot optique c'est que, contrairement à ce qu'il se passe avec les méthodes statistiques, les valeurs (u, v) du déplacement estimé à chaque itération n'est pas entier. Ce qui veut dire qu'on doit estimer, pour l'image courante, la valeur d'illumination $I(x + \delta x, y + \delta y)$ pour des valeurs non-entières de $(\delta x, \delta y)$. Celà se fait par interpolation. Cette interpolation peut être déportée dans une fonction provocant la translation de l'image courante, ou bien être réalisée directement dans la boucle qui calcule I_t la variation temporelle des niveaux de gris du motif.

Point important, la méthode de suivi par flot optique étant une approximation au premier

ordre, il se peut qu'elle ne converge pas en une seule itération. Si votre motif "décroche" cela veut dire qu'il faut faire plusieurs itérations (5-10) de l'estimation de (u, v) avant de passer à l'image suivante de la séquence.

4• TRAVAIL DE PROGRAMMATION.

Vous disposez de deux fichiers, l'un portant le nom de PoursuiteCible.cpp et l'autre TraitementImage.h. Le premier vous donne un canevas du programme à réaliser (lecture d'une séquence d'images à partir du nom de la première image qui doit être préalablement compilé puis lancé par :

PoursuiteCible ./Ghost1/GITSOO1.bmp

(par exemple)

l'autre un ensemble de fonction que vous pouvez utiliser (inversion d'une matrice 2x2, calcul d'une transformation rigide sur une image, calcul du gradient d'une image, ...)

Vous avez la liberté de tester plutôt la méthode basée distance et la méthode basée flot optique. L'idéal serait de programmer les deux pour comparer leur performance. Les séquences nommées Ghost(1 à 7) sont des extraits du film Ghost in the Shell avec des travellings (sauf pour la séquence 7) ... sélectionnez un détail et voyez ce qui se passe.

Remarque générale.

Il vous est conseillé de programmer **proprement** c'est à dire en définissant tous vos paramètres sous forme de variable (avec des noms explicites) et en commentant le plus possible ce que vous écrivez.