

# Chapitre 4 : le mouvement

---

## 4.1. Introduction

L'analyse de mouvement dans les vidéos s'est révélé être un outil indispensable pour les applications aussi diverses que la vidéo surveillance, l'image médicale, la robotique, ...

En effet, les zones de mouvement d'une séquence d'images correspondent souvent à des événements sur lesquels un système de vision doit se focaliser. L'analyse de mouvement englobe un certain nombre de problématiques

- La détection des objets en mouvement, c.à.d. la détection d'un ensemble de régions d'intérêts en mouvement dans la scène observée
- Le suivi de primitives ou de régions, dont le but est de déterminer la position de chaque primitive ou région dans l'image à chaque instant

## 4.2. La détection de mouvement

C'est une première étape pour des outils automatiques de vision par ordinateur. Le but est soit de uniquement détecter, soit de détecter et reconnaître, soit de détecter et suivre des objets pour, par exemple, analyser le comportement ou la trajectoire de ces objets.

- **La détection par soustraction du fond :**

L'approche la plus évidente consiste à prendre comme modèle de fond (arrière-plan), une des images de la séquence ne contenant pas d'objet mobile. Notons  $I$ , l'ensemble des images de la séquence,  $I(t)$  une image de cette séquence à un instant  $t$  et  $I(x,y,t)$  un pixel composant l'image de cette séquence à un instant  $t$ . ainsi on compare au modèle de fond notée  $BG$ , l'ensemble des images de la séquence

$$|I(x,y,t) - BG(x,y,t)| > th$$

En appliquant un seuil  $th$ , on obtient une séquence d'images contenant les pixels appartenant au premier plan ou aux régions mobiles de la scène

Il existe d'autres variantes de cette méthode. Une de ces méthodes consiste à considérer simplement l'image précédente dans la séquence comme fond pour l'image courant

$$|I(x,y,t) - I(x,y,t-1)| > th$$

Une autre méthode utilisée comme modèle de fond la moyenne des  $n$  images précédent l'image courante

$$BG = \frac{1}{n} \sum_{k=0}^{n-1} I(x,y,t-k) \quad \text{et}$$

$$|I(x,y,t) - BG(x,y,t)| > th$$

- Cependant ces méthodes exigent que la camera soit parfaitement fixe
- La luminosité de la scène change et varie dans le temps, l'image peut être entachée du bruit, un mouvement récurrent et intéressant peut être détecté en faux positif

## 4.3. Pursuit de cible (tracking)

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

Nous considérons que les méthodes de suivi peuvent être divisées en 3 classes :

- Méthodes de suivi de région : elles extraient les objets d'intérêts de la scène en utilisant des méthodes de segmentation.
- Méthodes de suivi à partir des mesures de vitesse
- Méthodes de suivi de caractéristiques : elles suivent certaines caractéristiques de l'objet : points, contours, couleurs, texture....

# Chapitre 4 : le mouvement

## 4.3.1. Méthode de mise en correspondance de motif

Le principe de cette méthode repose sur une mesure de similarité entre 2 images. Cette mesure de similarité est soit une distance statistique soit une mesure de corrélation.

La corrélation de signaux  $x_k$  et  $y_k$  est :  $R_{xy}(\tau) = \frac{1}{N} \sum_{k=1}^N x_k y_{k+\tau}$

Le coefficient de corrélation vaut 1 si  $x_k$  et  $y_{k+\tau}$  se superposent parfaitement et est inférieur à 1 sinon, il s'écrit :

$$\rho_{xy} = \frac{\sum_{k=1}^N (x_k - \bar{x})(y_{k+\tau} - \bar{y})}{\sqrt{\sum_{k=1}^N (x_k - \bar{x})^2 \sum_{k=1}^N (y_k - \bar{y})^2}}$$

Avec

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k \quad , \quad \bar{y} = \frac{1}{N} \sum_{k=1}^N y_k$$

L'utilisation de la corrélation pour la poursuite consiste à retrouver le décalage  $(u, v)$  donnant la meilleure corrélation entre le signal motif et le signal image courante. Le motif étant toujours d'une dimension inférieure à celle de l'image, la procédure la plus simple consiste à extraire de l'image courante des imagettes de même taille que le motif sélectionné et de calculer la corrélation.

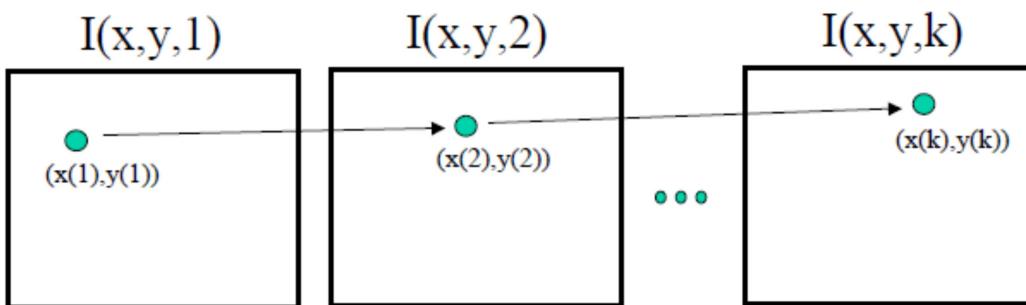
En traitement d'image, la mesure de corrélation 2D est notée

$$\text{ZNCC}(u, v) = \frac{\sum_{i,j} (I_1(i, j) - \bar{I}_1)(I_2(i + u, j + v) - \bar{I}_2)}{\sqrt{\sum_{i,j} (I_1(i, j) - \bar{I}_1)^2 \sum_{i,j} (I_2(i, j) - \bar{I}_2)^2}}$$

Les valeurs de  $u$  et  $v$  qui maximisent cette fonction de corrélation doivent être retenues comme mesure du décalage entre le motif et l'image

## 4.3.2. Le flot optique :

La méthode du flot optique consiste à supposer que le motif est une fonction discrète à 2 dimensions produite par la discrétisation de l'illumination locale.



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 éoulé entre deux acquisitions d'images

Si le motif a toujours la même répartition d'illumination, alors

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

Le flop optique suppose que :

1. Le motif a toujours la même répartition d'illumination (la cible apparaît la même pour chaque instant)

## Chapitre 4 : le mouvement

---

### 2. Des petits déplacements

### 3. La cohérence spatiale : les points se déplacent de façon similaire à leurs voisins

$$1 \rightarrow I(x, y, t - 1) = I(x + u(x, y), y + v(x, y), t)$$

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

$$\begin{aligned} I(x, y, t - 1) &\approx I(x, y, t) + I_x \cdot u(x, y) + I_y \cdot v(x, y) \\ I_x u + I_y v + I_t &\approx 0 \end{aligned}$$

$I_x, I_y$  : dérivées spatiales,  $I_t$  : dérivée temporelle

Elle peut être écrite sous la forme

$$\nabla I(u, v) + I_t = 0$$

Ce qui donne une seule équation avec 2 inconnus  $u, v$ .

Pour obtenir plus d'équations : la cohérence spatiale  $3 \rightarrow$  les pixels voisins ont le même déplacement  $(u, v)$ . Si nous utilisons une fenêtre de  $5 \times 5$ , cela nous donne 25 équations par pixels :

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_{25}) & I_y(p_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_{25}) \end{bmatrix}$$
$$\begin{matrix} A & d & b \\ 25 \times 2 & 2 \times 1 & 25 \times 1 \end{matrix}$$

La solution de ce problème est donnée par les moindres carrés

$$\begin{matrix} A & d = b \\ 25 \times 2 & 2 \times 1 & 25 \times 1 \end{matrix} \longrightarrow \text{minimize } \|Ad - b\|^2$$

$$(A^T A) \begin{matrix} d \\ 2 \times 1 \end{matrix} = A^T b \begin{matrix} 2 \times 1 \\ 2 \times 1 \end{matrix}$$

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$
$$\begin{matrix} A^T A & A^T b \\ 2 \times 2 & 2 \times 1 \end{matrix}$$

La solution sera

$$d = (A^T A)^{-1} \cdot A^T \cdot b$$