

# Analyse de la vidéo

## Chapitre 2.3 - Estimation du mouvement global

20 janvier 2014

# Plan de la présentation

- 1 Estimation du mouvement global (EMG) - Domaine fréquentiel
  - Corrélation de phase (translation)
  - Corrélation de phase (rotation, mise à échelle)
- 2 Estimation du mouvement global (EMG) - Domain spatial
  - Moindres carrés - estimation directe (intensité)
  - Moindres carrés - estimation indirecte (vecteurs de mouvement)
  - Moindres carrés - estimation indirecte (mouvement affine)

# Plan de chapitre

- 1 Estimation du mouvement global (EMG) - Domaine fréquentiel
  - Corrélation de phase (translation)
  - Corrélation de phase (rotation, mise à échelle)
- 2 Estimation du mouvement global (EMG) - Domain spatial
  - Moindres carrés - estimation directe (intensité)
  - Moindres carrés - estimation indirecte (vecteurs de mouvement)
  - Moindres carrés - estimation indirecte (mouvement affine)

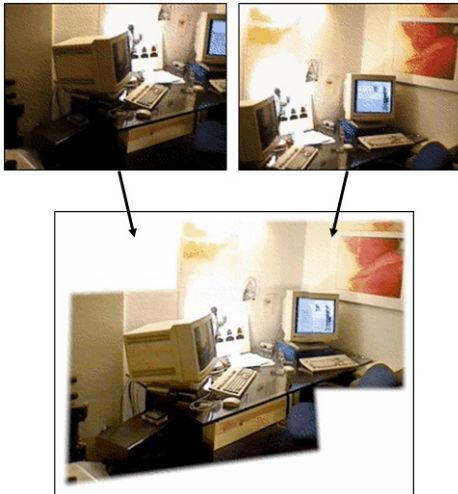
# Estimation du mouvement global

L'estimation du mouvement global peut être utilisé pour :

- Alignement d'images ;
- Réduction de la vibration de la caméra (*jittering*) ;
- Compensation de mouvement ;
- Mosaïque ;
- Segmentation vidéo ;
- ...

# Estimation du mouvement global

Exemple : Mosaïque

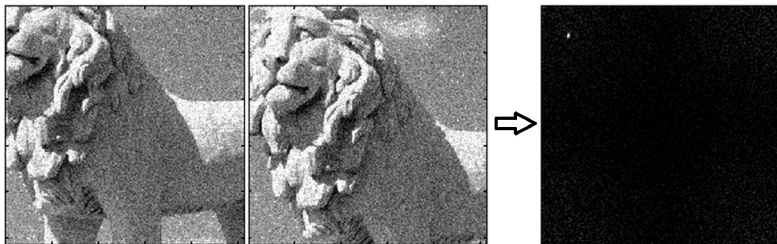


# Décalage par corrélation de phase

## Modèle de translation

Cette méthode utilise les propriétés de la transformation de Fourier ( $\mathcal{F}$ ) pour calculer le déplacement d'image.

La comparaison des images dans le domaine de Fourier nous permet d'extrapoler les paramètres de déplacement de l'image.



*Tiré de Wikipédia*

# Décalage par corrélation de phase

## Modèle de translation

On tente d'obtenir le vecteur de déplacement idéal par **corrélation**, c'est-à-dire en maximisant le produit<sup>1</sup> de deux images alignées :

$$E_{cc}(d_x, d_y) = \sum_{(x,y)} \left[ I(x, y, t_1) \cdot I(x + d_x, y + d_y, t_2) \right] \quad (1)$$

Quel est le problème avec cette opération ?

---

1. On dénote cette opération *cross-correlation* (*cc*)

# Décalage par corrélation de phase

## Modèle de translation

Par la constance d'illumination :

$$\begin{aligned}\mathcal{F}[I(x, y, t_1)] &= \mathcal{F}[I(x + d_x, y + d_y, t_2)] \\ \mathcal{I}(f_x, f_y) &= \mathcal{I}(f_x, f_y) \cdot e^{2\pi i(x + d_x, y + d_y)}\end{aligned}\tag{2}$$

Où le déplacement spatial est reflété par un décalage de phase dans le domaine fréquentiel.



# Décalage par corrélation de phase

## Modèle de translation

Par l'opération de *cross-correlation* en entre deux images :

$$\begin{aligned}
 E_{cc} &= \sum_{(x,y)} \left[ I(x, y, t_1) \cdot I(-x, -y, t_2) \right] \\
 \mathcal{F}[E_{cc}] &= \mathcal{F} \left[ I(-x, -y, t_1) \cdot I(x, y, t_2) \right] \\
 \mathcal{F}[E_{cc}] &= \mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_2)
 \end{aligned} \tag{3}$$

Où  $\mathcal{I}^*$  est le conjugué complexe de  $\mathcal{I}$ . Afin d'enlever l'influence d'un changement d'illumination global dans l'image, on normalise la cross-corrélation :

$$\mathcal{F}[E_{cc}] = \frac{\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_2)}{|\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_2)|} \tag{4}$$

# Décalage par corrélation de phase

## Modèle de translation

En combinant Eq.4 et Eq.2 :

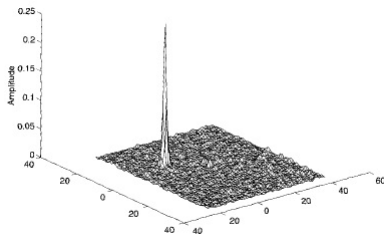
$$\begin{aligned}
 \mathcal{F}[E_{cc}(d_x, d_y)] &= \frac{\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_2)}{|\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_2)|} \\
 &= \frac{\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_1) \cdot |e^{2\pi i(x+d_x, y+d_y)}|}{|\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_1)| \cdot e^{2\pi i(x+d_x, y+d_y)}} \\
 &= \frac{\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_1)}{|\mathcal{I}^*(x, y, t_1) \cdot \mathcal{I}(x, y, t_1)|} \cdot e^{-2\pi i(x+d_x, y+d_y)} \\
 &= e^{-2\pi i(x+d_x, y+d_y)}
 \end{aligned} \tag{5}$$

# Décalage par corrélation de phase

## Modèle de translation

$$E_{cc}(d_x, d_y) = \delta(x + d_x, y + d_y) \quad (6)$$

L'expression de l'Eq.(6) est une impulsion Dirac située à l'emplacement  $(d_x, d_y)$ . Il est alors facile de cerner la solution.



Cette méthode est appelé **corrélation de phase**.

# Décalage par corrélation de phase

## Modèle de translation

L'Eq.6 est possible par les propriétés suivantes :

- Le déphasage spatial :  
$$\mathcal{F}[I(x + d_x, y + d_y)] = \mathcal{I}(f_x, f_y) \cdot e^{-2i\pi(d_x f_x + d_y f_y)}$$
- Le spectre de phase de  $\mathcal{I}(f_x, f_y) \cdot \mathcal{I}^*(f_x, f_y)$  est nul.
- La constance d'illumination, mais on peut la changer de façon globale
- Décalage rigide et translationnel.

# Décalage par corrélation de phase

## Modèle de translation

Avantages de cette méthode :

- Rapide ;
- Résultat très bon.

Désavantages de la méthode :

- Exige des images sans trop de bruits ou de changements d'illumination ;
- Adapté pour un mouvement global sans trop de mouvements locaux.
- Difficile à automatiser, car les mouvements globaux ne sont pas nécessairement qu'une translation.

# Décalage par corrélation de phase

## Algorithme de la corrélation de phase

**Entrées:**  $I(t), I(t+1)$

**Sorties:**  $d_x^*, d_y^*$  : Déplacement global de l'image

$\mathcal{I}(t) \leftarrow \mathcal{F}[I(t)], \quad \mathcal{I}(t+1) \leftarrow \mathcal{F}[I(t+1)];$   
 $\mathcal{I}^*(t+1) \leftarrow \text{Conjugué complexe de } \mathcal{I}(t+1);$

On calcule alors la phase pour tous les pixels :

$Im(t), Re(t) \leftarrow \mathcal{I}(t), \quad Im(t+1), Re(t+1) \leftarrow \mathcal{I}^*(t+1);$   
 $\theta(t) \leftarrow \text{atan2}\left(\frac{Im(t)}{Re(t)}\right), \quad \theta(t+1) \leftarrow \text{atan2}\left(\frac{Im(t+1)}{Re(t+1)}\right);$   
 $Re \leftarrow \cos(\theta(t) + \theta(t+1)), \quad Im \leftarrow \sin(\theta(t) + \theta(t+1));$   
 $E_{cc} \leftarrow \mathcal{F}^{-1}[Re, Im];$   
 $d_x^*, d_y^* \leftarrow \text{Emplacement du pixel maximum de } E_{cc};$   
**retourner**  $d_x^*, d_y^*$

# Décalage par corrélation de phase

## Modèle de rotation

Le modèle décrit précédemment ne permet que de décrire le mouvement de translation.

Cependant, dans le cas où les deux frames ne sont réunis que par *pure rotation* :

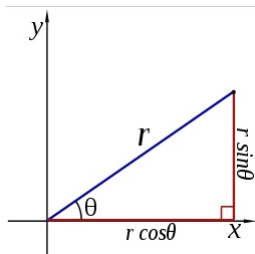
$$I(R(x, y), t_2) = I(x, y, t_1) \quad (7)$$

On peut retrouver une expression nous permettant d'estimer le mouvement global par un changement de base.

# Décalage par corrélation de phase

## Modèle de rotation

En exprimant les images en coordonnées polaires, on peut effectuer une translation dans l'espace polaire :



$$\begin{aligned}\bar{l}(r, \theta, t_1) &= l(x, y, t_1) \\ \bar{l}(r, \theta, t_2) &= l(x, y, t_2)\end{aligned}\quad (8)$$

$$\begin{aligned}\text{Où } r &= \sqrt{x^2 + y^2} \\ \theta &= \arctan\left(\frac{y}{x}\right)\end{aligned}$$

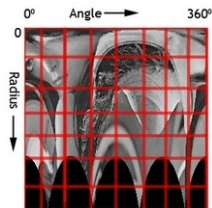
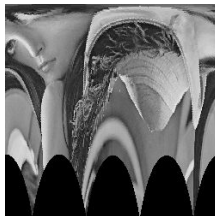
Et en supposant de petites rotations, on obtient :

$$\bar{l}(r, \theta + \hat{\theta}, t_2) = \bar{l}(r, \theta, t_1) \quad (9)$$



# Décalage par corrélation de phase

## Mappage en coordonnée polaire

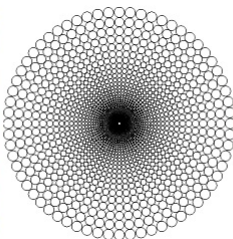


# Décalage par corrélation de phase

## Modèle de rotation et de changement d'échelle

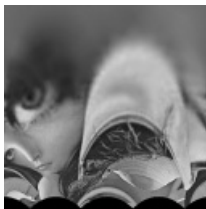
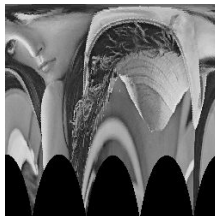
Toujours en effectuant un changement de coordonnées, cette fois-ci en **coordonnées log-polaires**, on peut représenter le **changement d'échelle**.

La différence avec les coordonnées polaires vient de l'espacement entre les cercles concentriques qui varient de façon exponentielle :



# Décalage par corrélation de phase

## Mappage en coordonnée log-polaire

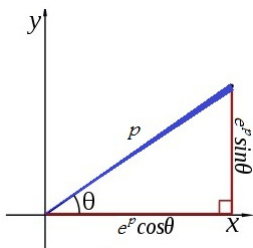


*Lena en coordonnée log-polaire*

# Décalage par corrélation de phase

## Modèle de rotation et de changement d'échelle

En exprimant les images en coordonnées log-polaires, on peut effectuer une translation dans l'espace log-polaire :



$$\begin{aligned}\bar{l}(\rho, \theta, t_1) &= l(x, y, t_1) \\ \bar{l}(\rho, \theta, t_2) &= l(x, y, t_2)\end{aligned}\quad (10)$$

$$\begin{aligned}\text{Où } \rho &= \log \sqrt{x^2 + y^2} \\ \theta &= \arctan\left(\frac{y}{x}\right)\end{aligned}$$

Et en supposant une petite rotation et un petit changement d'échelle, on obtient :

$$\bar{l}(\rho + \hat{\rho}, \theta + \hat{\theta}, t_2) = \bar{l}(\rho, \theta, t_1) \quad (11)$$

# Décalage par corrélation de phase

## Translation, Rotation et changement d'échelle

Le mapping **log-polaire** tente d'illustrer le changement d'échelle **rétinal** (de l'oeil humain).

Dans le cas où les trois transformations sont présentes, mais que la translation est relativement petite :

$$I(e^s \cdot R(x, y), t_2) = I(x, y, t_1) \quad (12)$$

Comme la focale est située sur le point central de l'image, il peut être difficile d'estimer le changement d'échelle réel dans les cas de **translation élevée**.

# Décalage par corrélation de phase

## Algorithme de la corrélation de phase (log-polaire)

Une approche séquentielle consiste à séparer en plusieurs étapes l'estimation des paramètres :

---

---

**Entrées:**  $I(t), I(t+1)$

**Sorties:**  $d_x^*, d_y^*, R, s$  : Déplacement global de l'image  $d^*$ , Rotation  $R$  et facteur d'agrandissement  $s$

$d_x^*, d_y^* \leftarrow$  Estimation par corrélation de phase avec  $I(t)$  et  $I(t+1)$ ;

$I^*(t) \leftarrow$  Image  $I(t)$  modifiée par  $d_x^*, d_y^*$ ;

$I_{log}^*(t), I_{log}(t+1) \leftarrow$  Mapping log-polaire des images  $I^*(t)$  et  $I(t+1)$ ;

$R, s \leftarrow$  Estimation par corrélation de phase avec  $I_{log}^*(t)$  et  $I_{log}(t+1)$ ;

**retourner**  $d_x^*, d_y^*, R, s$

---

# Plan de chapitre

- 1 Estimation du mouvement global (EMG) - Domaine fréquentiel
  - Corrélation de phase (translation)
  - Corrélation de phase (rotation, mise à échelle)
- 2 Estimation du mouvement global (EMG) - Domain spatial
  - Moindres carrés - estimation directe (intensité)
  - Moindres carrés - estimation indirecte (vecteurs de mouvement)
  - Moindres carrés - estimation indirecte (mouvement affine)

# Estimation paramétrique du mouvement

Soit  $T((x, y); \mathbf{p})$  la transformation du pixel  $(x, y)$  par le modèle paramétrique, où  $\mathbf{p} = (p_1, \dots, p_L)$  est l'ensemble des paramètres du modèle.

Dans le cas du modèle affine, par exemple, on aura :

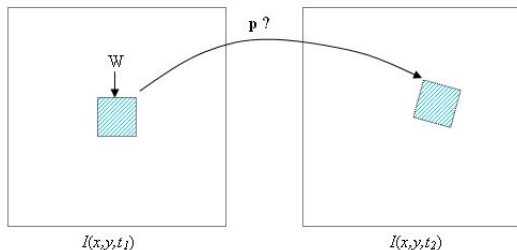
$$T((x, y); \mathbf{p}) = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + a_0 + a_1x + a_2y \\ y + b_0 + b_1x + b_2y \end{pmatrix}$$

et  $\mathbf{p} = (a_0, a_1, a_2, b_0, b_1, b_2)$ .



# Estimation paramétrique du mouvement

Pour estimer le mouvement pour une fenêtre  $W$ , il suffit d'estimer les paramètres  $\mathbf{p}$  du modèle. **La fenêtre  $W$  sera l'ensemble du frame au complet.**



# Estimation paramétrique du mouvement

Cette estimation est réalisée par la minimisation de la fonction suivante :

$$E(\mathbf{p}) = \sum_{(x,y) \in W} [I(T((x,y); \mathbf{p}), t_1) - I(x,y, t_2)]^2 \quad (13)$$

On peut minimiser cette fonction soit par un **algorithme d'optimisation** (descente de gradient), soit par **régression linéaire** ou par un **raffinement itératif**.

# Estimation paramétrique directe par raffinement itératif

Ayant une valeur courante  $\mathbf{p}^{(k)}$ , l'estimation de  $\Delta\mathbf{p}^{(k)}$  revient à minimiser la fonction suivante sur  $\Delta\mathbf{p}^{(k)}$  :

$$E(\Delta\mathbf{p}^{(k)}) = \sum_{(x,y) \in W} [I(T((x,y); \mathbf{p}^{(k)} + \Delta\mathbf{p}^{(k)}), t_1) - I(x, y, t_2)]^2 \quad (14)$$

Après le développement de Taylor à l'ordre 1 de la fonction (14), on obtient l'équation suivante :

$$E(\Delta\mathbf{p}^{(k)}) \simeq \sum_{(x,y) \in W} [I(T((x,y); \mathbf{p}^{(k)}), t_1) + \nabla I \frac{\partial T}{\partial \mathbf{p}} \Delta\mathbf{p}^{(k)} \quad (15)$$

$$- I(x, y, t_2)]^2 \quad (16)$$

# Estimation paramétrique directe

CRAYON

# Estimation paramétrique directe

Par manipulation directe, on obtient alors :

$$\Delta \mathbf{p}^{(k)} \simeq \frac{\sum_{(x,y) \in W} \left[ \nabla l \frac{\partial T}{\partial \mathbf{p}} \right] \left[ \nabla l \frac{\partial T}{\partial \mathbf{p}} \right]^T}{\sum_{(x,y) \in W} \left[ \nabla l \frac{\partial T}{\partial \mathbf{p}} \right]^T \left[ l(T((x,y); \mathbf{p}^{(k)}), t_1) - l(x,y, t_2) \right]} \quad (17)$$

# Algorithme de l'estimation paramétrique directe

**Entrées:**  $I(t), I(t+1)$

**Sorties:**  $\mathbf{p}^{(k)}$  : Paramètres du modèle de mouvement global de l'image

Recherche exhaustive ou descente du gradient :

$\mathbf{p}^0 \leftarrow$  Initialisation par recherche exhaustive;

Raffinement itératif :

**tant que**  $\Delta \mathbf{p}^{(k)} < \epsilon$  **faire**

$I \left( T \left( (x, y); \mathbf{p}^{(k)} \right), t \right) \leftarrow$  Image  $I(t)$  transformée par les paramètres  $\mathbf{p}^{(k)}$ ;

$\nabla I \leftarrow$  Calculer le gradient de l'image  $I \left( T \left( (x, y); \mathbf{p}^{(k)} \right), t \right)$ ;

$\frac{\partial T}{\partial \mathbf{p}} \leftarrow$  Calculer la matrice des dérivées partielles du modèle de transformation;

$\left[ \nabla I \frac{\partial T}{\partial \mathbf{p}} \right]$  et  $\left[ \left[ \nabla I \frac{\partial T}{\partial \mathbf{p}} \right] \left[ \nabla I \frac{\partial T}{\partial \mathbf{p}} \right]^T \right]^{-1} \leftarrow$  Calculer les termes de Eq.(17);

$\Delta \mathbf{p}^{(k)} \leftarrow$  Calculer Eq.(17).;

$\mathbf{p}^{(k+1)} \leftarrow \Delta \mathbf{p}^{(k)} + \mathbf{p}^{(k)}$ ;

$k \leftarrow k + 1$ ;

**retourner**  $\mathbf{p}^{(k)}$

## Estimation paramétrique indirecte (Translation)

À la différence de l'estimation paramétrique directe, on utilise les **vecteurs de mouvements précédemment calculés** afin d'y retrouver un **mouvement global commun**.

Supposons un mouvement translationnel et  $d^*$  le mouvement translationnel global commun  $d^*$ . L'expression de l'énergie à minimiser est :

$$E(\mathbf{d}^*) = \sum_{(x,y) \in W} [\mathbf{d}(x,y) - \mathbf{d}^*]^2 \quad (18)$$

On peut résoudre cette équation par **régression linéaire** ou **recherche exhaustive**.

# Algorithme de l'estimation paramétrique indirecte (Translation)

---

---

**Entrées:**  $I(t), I(t + 1)$

**Sorties:**  $\mathbf{d}^*$  : Paramètres du modèle de mouvement global de l'image

$\mathbf{d}_x, \mathbf{d}_y \leftarrow$  Estimation du mouvement (bloc fixe, H-S, ...);

$\mathbf{d}^* \leftarrow$  Minimisation de l'Eq.18;

**retourner  $\mathbf{d}^*$**

---

Si on a déjà accès aux vecteurs de déplacement, une **compensation du mouvement** est très simple à appliquer à l'aide de cette méthode.



## Estimation paramétrique indirecte (modèle affine)

Supposons un mouvement plus complexe tel un modèle affine.

$$T((x, y); \mathbf{p}) = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + a_0 + a_1 x + a_2 y \\ y + b_0 + b_1 x + b_2 y \end{pmatrix}$$

et  $\mathbf{p} = (a_0, a_1, a_2, b_0, b_1, b_2)$ .

L'expression de l'énergie à minimiser est :

$$E(\mathbf{p}^k) = \sum_{(x,y) \in W} [\mathbf{T}((x, y); \mathbf{p}^k) - \mathbf{d}(x, y)]^2 \quad (19)$$

On peut résoudre cette équation par **régression linéaire**.

# Estimation paramétrique indirecte (modèle affine)

CRAYON

# Estimation paramétrique indirecte (affine)

Algorithme de l'estimation paramétrique indirecte par régression linéaire

**Entrées:**  $I(t), I(t+1)$

**Sorties:**  $\mathbf{p}$  : Paramètres du modèle de mouvement global de l'image

$\mathbf{d} \leftarrow$  Initialisation par bloc fixe ou H-S ou ... ;

$\frac{\partial T}{\partial \mathbf{p}} \leftarrow$  Calculer la matrice des dérivées partielles du modèle de transformation;

$$\mathbf{p} \leftarrow \frac{\begin{bmatrix} \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix}^T \mathbf{d}}{\begin{bmatrix} \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix} \begin{bmatrix} \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix}^T};$$

**retourner**  $\mathbf{p}$

# Amélioration par estimation robuste

On peut améliorer nos résultats en **retirant les pixels** qui sont trop différents du mouvement global trouvé ou en **séparant les images en régions** :

- ➊ Prendre tous les pixels des images comme région ;
- ➋ Appliquer la méthode directe ou indirecte pour toute la région ;
- ➌ Évaluer l'erreur pour tous les pixels de la région ;
- ➍ Éliminer les pixels ayant une erreur trop élevée ;
- ➎ Répéter les étapes 2-4 pour les pixels exclus.