

Informatique visuelle - Vision par ordinateur

Extractions de caractéristiques - les points d'intérêt

Elise Arnaud
elise.arnaud@imag.fr

cours inspiré par X. Descombes, J. Ros, A. Boucher, A. Manzanera, E. Boyer, M Black, J.H. Thomas

Elise Arnaud elise.arnaud@imag.fr

M2P UFR IMA

Vision par ordinateur

Points d'intérêts

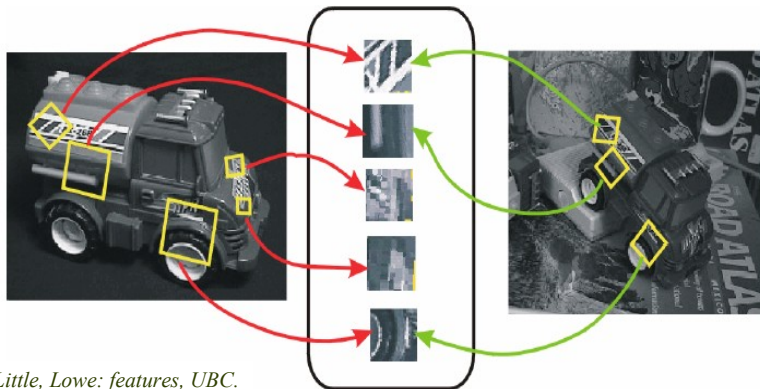
Alain Boucher - IFI

Elise Arnaud elise.arnaud@imag.fr

M2P UFR IMA

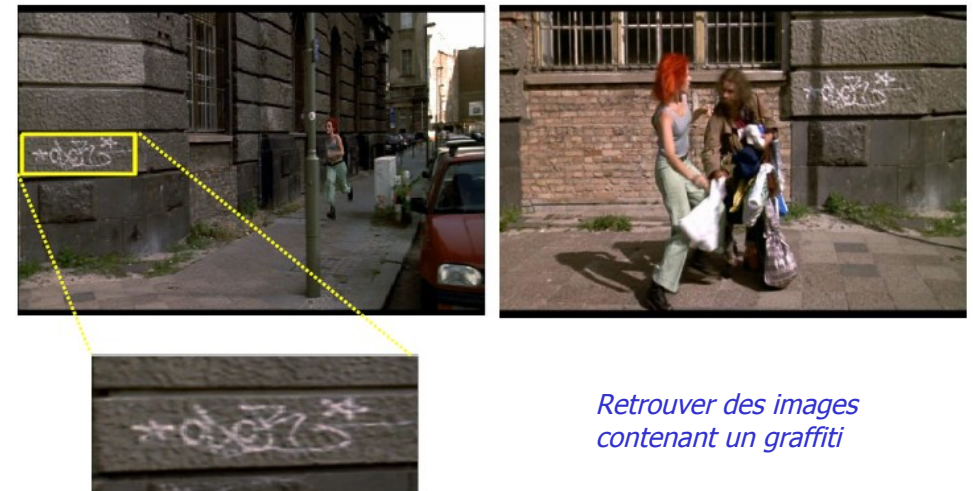
Détection de caractéristiques

- **But** : identifier des **caractéristiques robustes** dans l'image pour mise en correspondance (3D), reconnaissance, suivi (mouvement), ...
- On cherche souvent à **éviter la segmentation** (source d'erreurs) en identifiant d'autres caractéristiques



Source : Jim Little, Lowe: features, UBC.

Recherche d'images similaires



*Retrouver des images
contenant un graffiti*

Source : Ian Reid, Object Detection and Recognition, Oxford (UK).

Mise en correspondance d'images

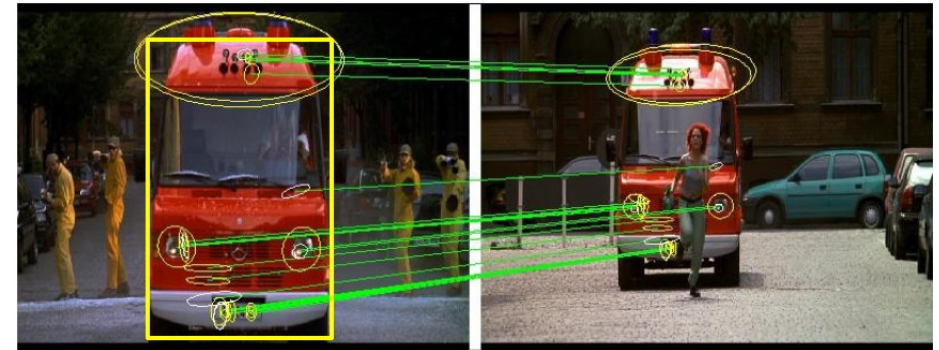


Des scènes sont différentes (angle de vue, zoom, lumière, ...), mais on veut identifier des éléments semblables...

Source : Ian Reid, *Object Detection and Recognition*, Oxford (UK).

4

Recherche d'objets dans les images



Ici, c'est le même autobus malgré la taille différente et le fait qu'il soit en partie caché sur une image...

Source : Ian Reid, *Object Detection and Recognition*, Oxford (UK).

5

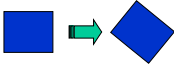
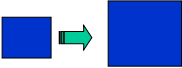


Propriétés idéales des caractéristiques

- **Localité** : une caractéristique locale (*vs globale*) sera plus robuste aux occlusions
- **Invariance** : à la translation, la rotation, le changement d'échelle, ...
- **Robuste** : au bruit, aux conditions d'acquisition de l'image, à la compression, ...
- **Discriminante** : permet d'identifier quelques objets spécifiques parmi beaucoup d'autres
- **Quantité** : pouvoir générer beaucoup de caractéristiques par objet / portion d'images
- **Précision** : position précise dans l'image / l'objet
- **Efficacité** : calcul rapide (temps réel)

Source : Tinne Tuytelaars, *Local Invariant Features: What? Why? When? How?*, ECCV 2006.

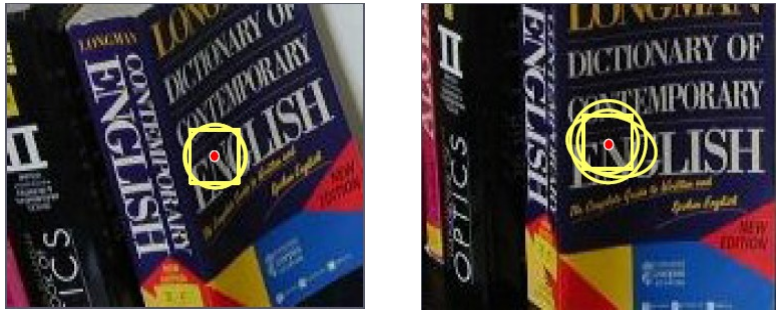
6

Propriété d'invariance

- Invariance à la géométrie
 - **Rotation** 
 - **Changement d'échelle** 
 - **Transformation affine** 
- Invariance à la photométrie
 - **Changement affine d'intensité** ($I \rightarrow aI + b$) 

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Transformations géométriques



Source : Tinne Tuytelaars, Local Invariant Features: What? Why? When? How?, ECCV 2006.

Transformations photométriques

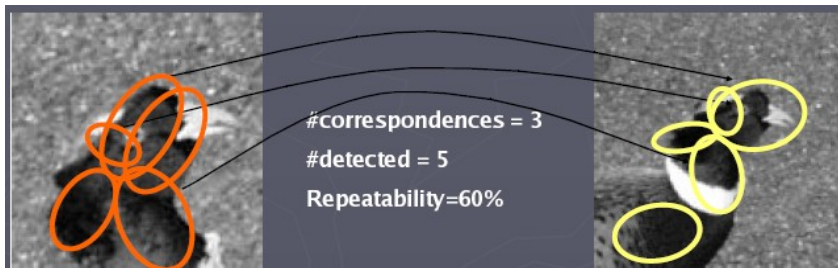


Source : Tinne Tuytelaars, Local Invariant Features: What? Why? When? How?, ECCV 2006.

Répétabilité des correspondances

Critère d'évaluation : correspondances entre images

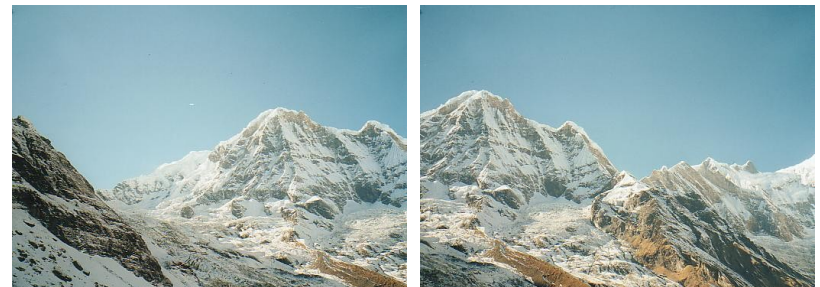
$$\text{Répétabilité} = \frac{\# \text{correspondances}}{\# \text{détectés}} \cdot 100\%$$



Source : Tinne Tuytelaars, Local Invariant Features: What? Why? When? How?, ECCV 2006.

Exemple : panorama

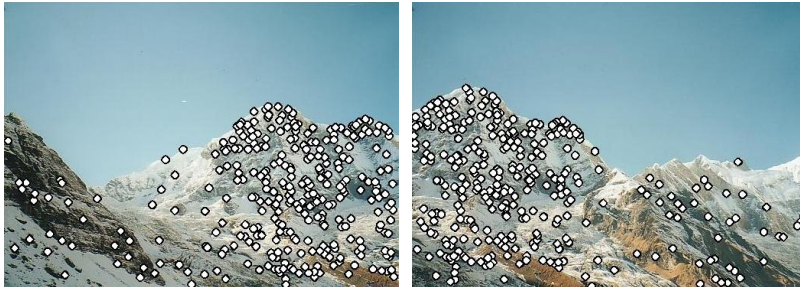
(1) Aligner les deux images



Source : Frolova & Simakov, Invariant Feature Detectors and Descriptors, Weizmann Ins. (Israel).

Exemple : panorama

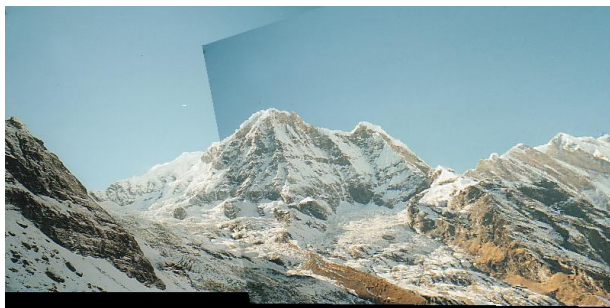
(2) détection des points caractéristiques dans les 2 images



Source : Frolova & Simakov, Invariant Feature Detectors and Descriptors, Weizmann Ins. (Israel).

Exemple : panorama

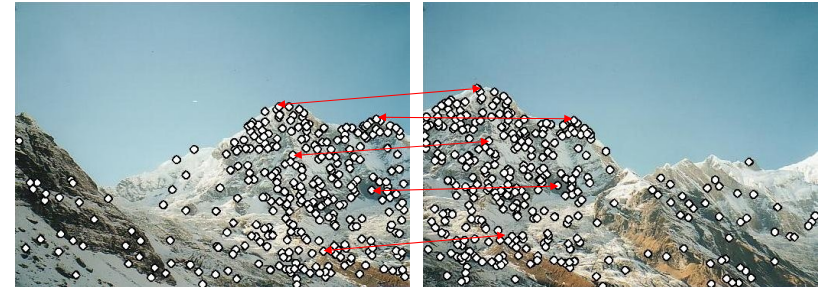
(4) Assembler les deux images pour en former une seule



Source : Frolova & Simakov, Invariant Feature Detectors and Descriptors, Weizmann Ins. (Israel).

Exemple : panorama

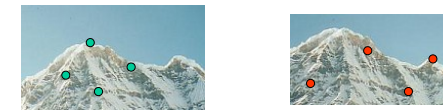
(3) trouver des paires de points correspondants



Source : Frolova & Simakov, Invariant Feature Detectors and Descriptors, Weizmann Ins. (Israel).

Exemple : panorama

- Problème à résoudre no 1 :
 - Détecter le même point dans les deux images



cas impossible

Répétabilité des points

Source : Frolova & Simakov, Invariant Feature Detectors and Descriptors, Weizmann Ins. (Israel).



Exemple : panorama

- Problème à résoudre no 2 :
 - Pour chaque point, trouver la bonne correspondance



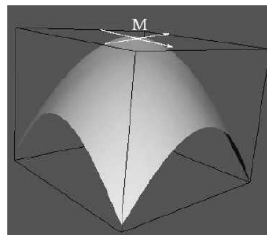
Descripteur robuste et discriminant

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Qu'est ce qu'un point d'intérêt ?



- **contour** : discontinuité dans une direction de la fonction d'intensité ou de ses dérivées
- **point d'intérêt** : dans deux directions



Plan

- ▶ détection de points d'intérêt
- ▶ mise en correspondance de points

Qu'est ce qu'un point d'intérêt ?

Avantages des points d'intérêt :

- ▶ Sources d'informations plus fiable que les contours car plus de contraintes sur la fonction d'intensité.
- ▶ Robuste aux occultations (soit occulté complètement, soit visible).
- ▶ Plus facile à extraire que les contours
- ▶ Présents dans une grande majorité d'images (\neq contours !).

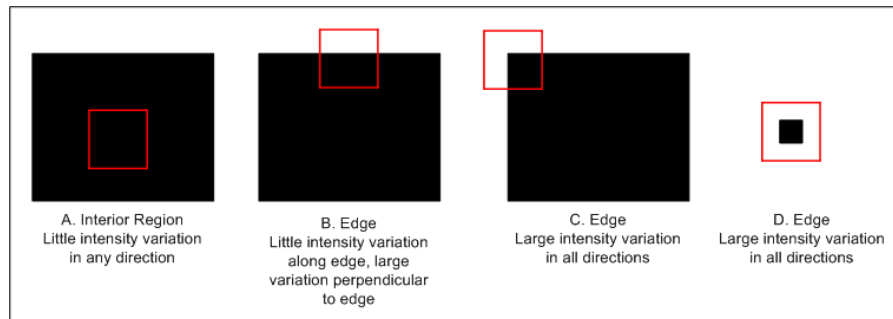
Détection de points d'intérêt

Différentes approches

1. **Approches contours** : détecter les contours puis extraction des points d'intérêts le long des contours en considérant les points de courbures maximales ainsi que les intersections de contours.
2. **Approches intensité** : à partir des niveaux de gris de l'image, trouver un opérateur qui est maximal aux points d'intérêt
3. **Approches à base de modèles** : identification des points d'intérêts par mise en correspondance de la fonction d'intensité avec un modèle théorique de cette fonction des points d'intérêts considérés.

→ Les approches de la deuxième catégorie sont celles utilisées généralement car (a) indépendance vis à vis de la détection de contours (b) indépendance vis à vis du type de points d'intérêts

Détecteur de Moravec



- A. **intensité presque constante** : $E(x, y) \approx 0$
- B. **contour** : $E(x, y) \approx 0$ pour des déplacements le long du contour ($y \neq 0$) ; $E(x, y) > 0$ pour des déplacements perpendiculaires
- C. **coin** : $E(x, y) > 0$ pour tout $(x, y) \neq (0, 0)$
- D. **pixel seul** : idem coin

Détecteur de Moravec (1980)

Variation moyenne de l'intensité pour un petit déplacement (x, y)

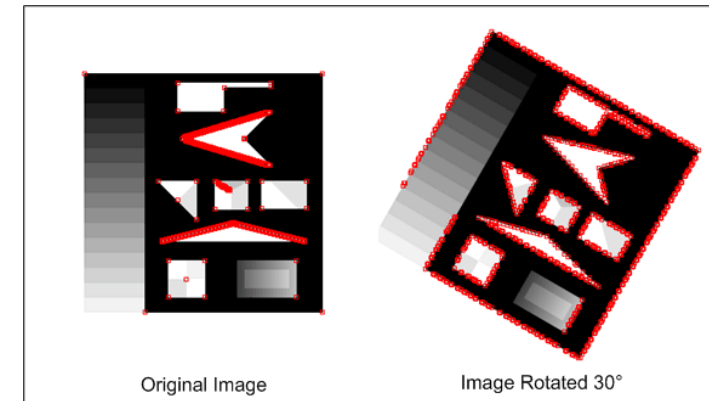
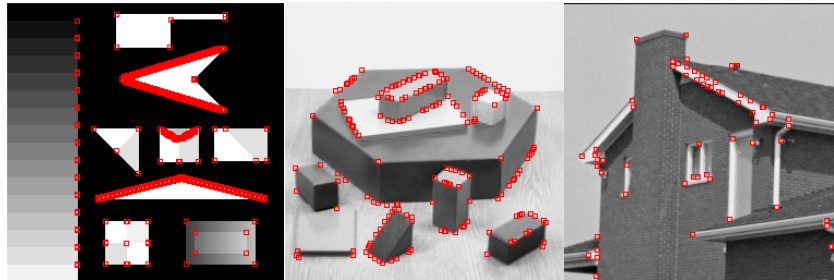
$$E(x, y) = \sum_{u, v} w(u, v) |I(x + u, y + u) - I(u, v)|^2$$

- ▶ w spécifie le voisinage considéré (valeur 1 à l'intérieur de la fenêtre et 0 à l'extérieur);
- ▶ $I(u, v)$ est l'intensité au pixel (u, v)

Détecteur de Moravec

- ⇒ un coin est un maximum local de E
- ⇒ pbl : la valeur de E est la même pour un coin que pour un pixel isolé

1. pour chaque pixel (u, v) , calculer les variations d'intensité $E(x, y)$ pour $(x, y) = \{(1, 0), (1, 1), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1)\}$
2. Construire la carte de "coinité" en calculant la mesure $C(u, v)$ pour chaque pixel (u, v) : $C(u, v) = \min E(x, y)$
3. Trouver les maxima de cette carte (correspondent aux points d'intérêt)



Du détecteur de Moravec au détecteur de Harris (1988)

On considère le développement de Taylor de la fonction d'intensité I au voisinage du pixel (u, v) :

$$I(x + u, y + v) = I(u, v) + x \frac{\delta I}{\delta x} + y \frac{\delta I}{\delta y} + o(x^2, y^2)$$

D'o :

$$E(x, y) = \sum_{u, v} w(u, v) \left[x \frac{\delta I}{\delta x} + y \frac{\delta I}{\delta y} + o(x^2, y^2) \right]^2$$

En négligeant le terme $o(x^2, y^2)$ (valide pour les petits déplacements) :

$$E(x, y) = Ax^2 + 2Cxy + By^2,$$

avec:

- ▶ $A = \frac{\delta I^2}{\delta x} \otimes w$; $B = \frac{\delta I^2}{\delta y} \otimes w$; $C = \left(\frac{\delta I}{\delta x} \frac{\delta I}{\delta y} \right) \otimes w$
- ▶ w : fenêtre gaussienne (+ isotrope)

Détecteur de Harris

$$E(x, y) = (x, y) \cdot M \cdot (x, y)^t,$$

avec :

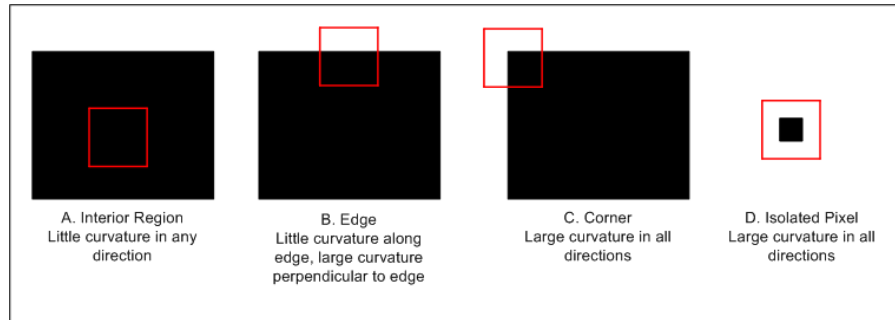
$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$

avec:

$$A = \frac{\delta I^2}{\delta x} \otimes w \quad ; \quad B = \frac{\delta I^2}{\delta y} \otimes w \quad ; \quad C = \left(\frac{\delta I}{\delta x} \frac{\delta I}{\delta y} \right) \otimes w$$

M : symétrique, définie positive \Rightarrow décomposition en valeurs propres

Détecteur de Harris



les valeurs propres de M correspondent aux courbures principales associées à E :

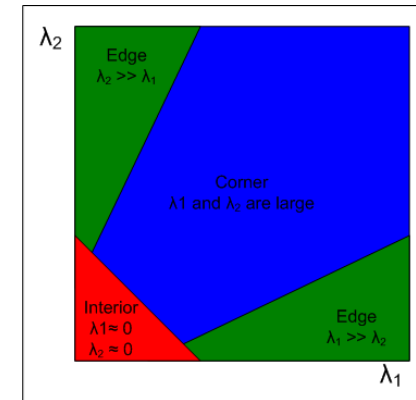
- A. **intensité presque constante** : les deux courbures sont de faibles valeurs
- B. **contour** : une des courbures est de forte valeur, l'autre est de faible valeur
- C. **point** : les deux courbures sont de fortes valeurs

Elise Arnaud elise.arnaud@imag.fr

M2P UFR IMA

Détecteur de Harris

valeurs propres de M : λ_1 et λ_2



Elise Arnaud elise.arnaud@imag.fr

M2P UFR IMA

Détecteur de Harris

Plutôt que de calculer les valeurs propres, il est possible de calculer

$$\det(M) = AB - C^2 = \lambda_1 \cdot \lambda_2$$

$$\text{trace}(M) = A + B = \lambda_1 + \lambda_2$$

et on calcule la réponse :

$$R = \det(M) - k \text{trace}^2(M)$$

Les valeurs de R sont positives au voisinage d'un coin, négatives au voisinage d'un contour et faibles dans une région d'intensité constante ($k = 0.04$)

⇒ coins/point d'intérêt = max locaux de R

Elise Arnaud elise.arnaud@imag.fr

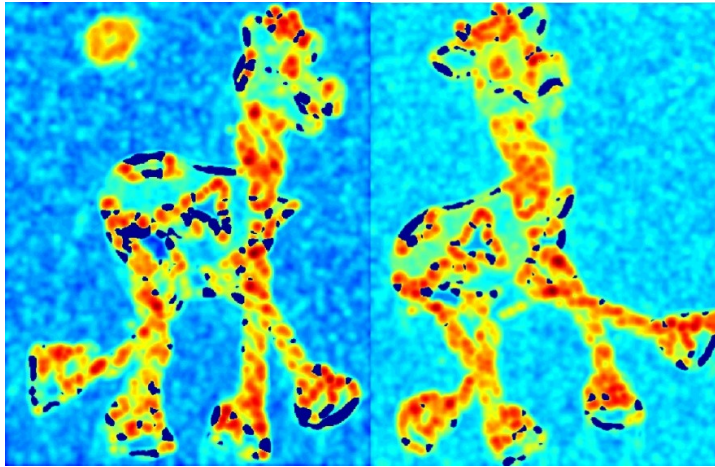
M2P UFR IMA

Détecteur de Harris : fonctionnement



Images originales

Détecteur de Harris : fonctionnement

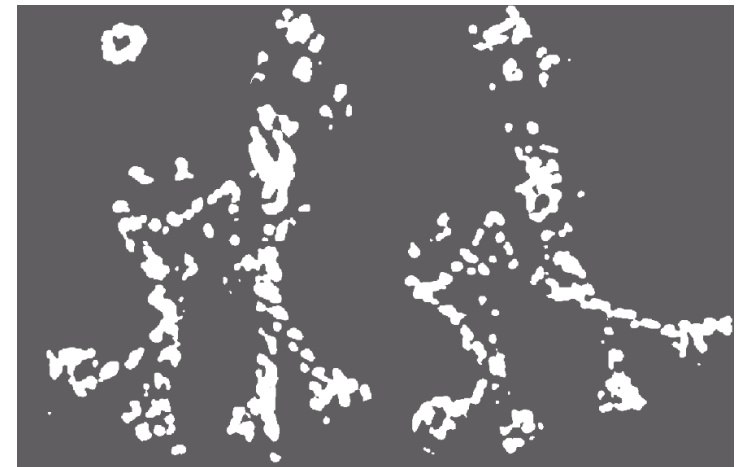


Calcul du détecteur de Harris

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

25

Détecteur de Harris : fonctionnement

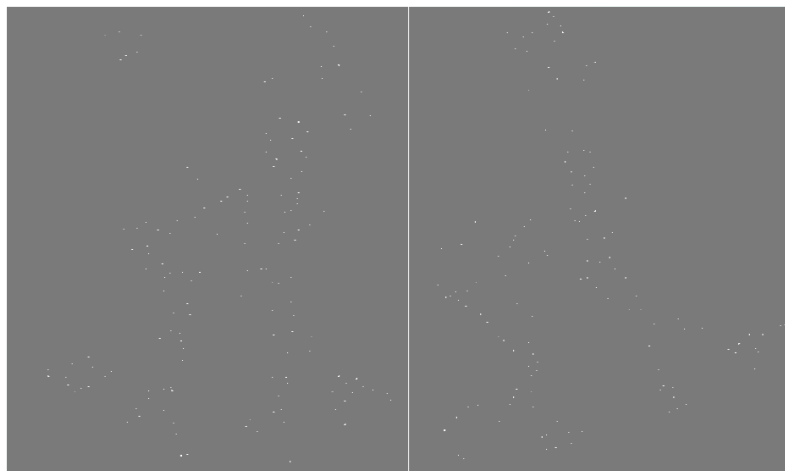


Seuillage des pixels (coins)

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

26

Détecteur de Harris : fonctionnement



Maxima locaux

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

27

Détecteur de Harris : fonctionnement

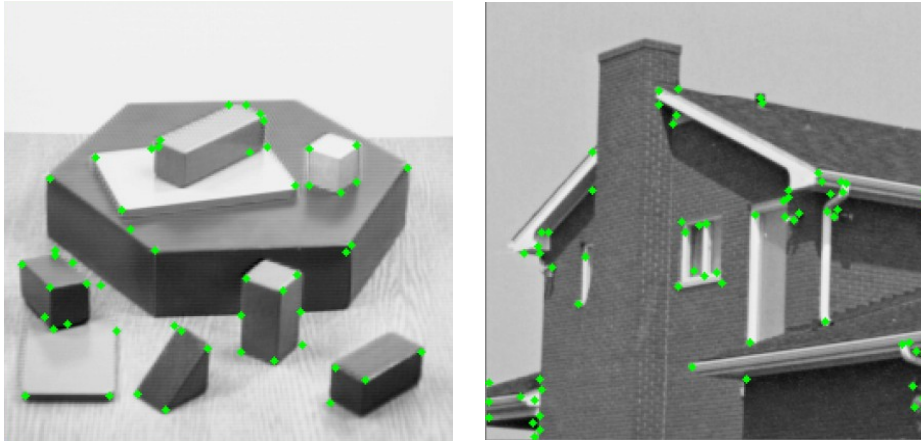


Superposition sur les images originales

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

28

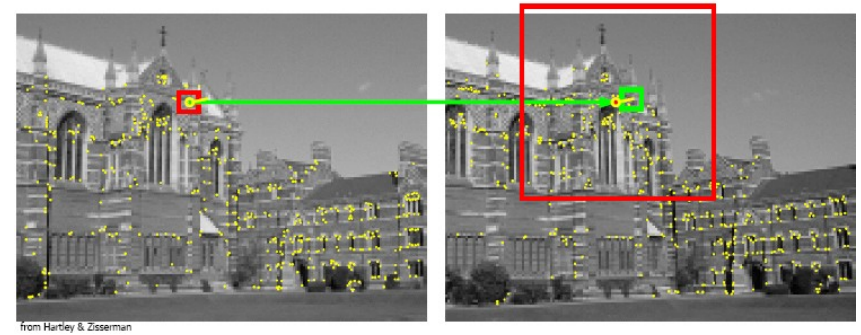
Exemple de détection de coins



Source : Dmitrij Csetverikov, *Finding Patterns in Images* (Hungary).

29

Détecteur de Harris : utilité



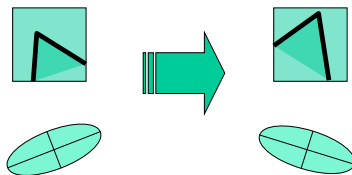
*Correspondance entre images (plus loin dans le cours)...
pour le 3D, le mouvement, la recherche d'images, ...*

Source : Jim Little, *CPSC 505 Image Understanding I: Image Analysis*, UBC (Canada)

30

Détecteur de Harris : propriétés

- Invariant en rotation

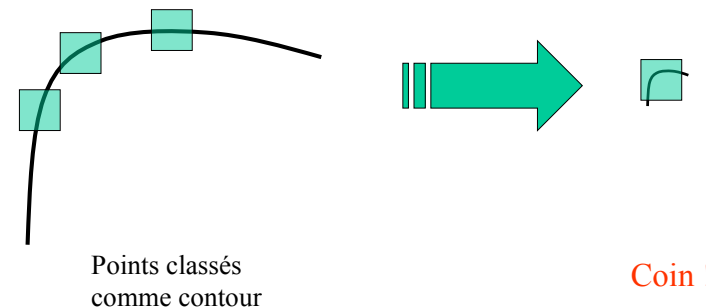


Même après rotation, la forme reste la même, et les valeurs propres de la matrice aussi

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Détecteur de Harris : propriétés

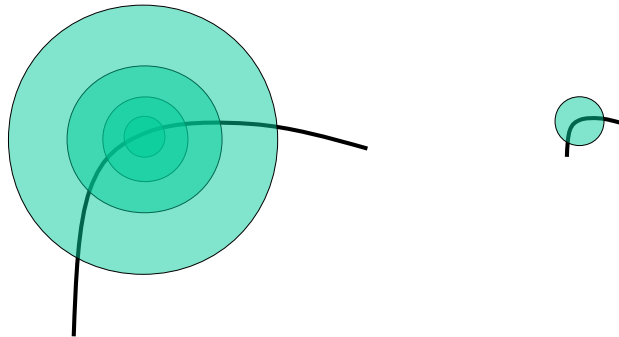
- Le détecteur de Harris n'est pas invariant à l'échelle
 - Ne fonctionne pas si changement d'échelle



Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Invariance à l'échelle

- Il existe des variantes permettant d'être invariant à l'échelle
- Exemple : Harris-Laplacien



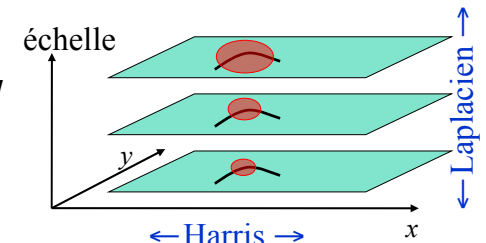
Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Mise en correspondance

- ▶ Les méthodes de corrélation sont utilisées depuis longtemps pour mettre en correspondance des pixels sur la base d'informations d'intensités.
- ▶ L'idée est de définir une mesure de similarité entre les pixels de deux images.
- ▶ Les pixels sont les primitives les mieux adaptés pour la mise en correspondance.
- ▶ Les régions sont en effet mal adaptées la mise en correspondance (la taille d'une région est différente d'une image à une autre), tout comme les contours

Harris-Laplacien

- **Harris-Laplacien¹**
Trouver le maximum local pour :
 - Les coins de Harris dans l'espace (x, y)
 - Le Laplacien en échelle



¹ K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

Source : Frolova & Simakov, *Invariant Feature Detectors and Descriptors*, Weizmann Ins. (Israel).

Mise en correspondance - principe

Le principe est de considérer, pour un pixel p_1 de l'image 1, une fenêtre rectangulaire centrée en p_1 et de calculer sa corrélation/distance avec une fenêtre dans la deuxième image. La fonction de corrélation est alors maximum en p_2 correspondant de p_1 dans la deuxième image (distance minimum)



Mise en correspondance - fonctions de dissimilarité

Pour une fenêtre de taille $2N + 1 \times 2P + 1$

- Sum of absolute differences (SAD)

$$SAD(p_1(u_1, v_1), p_2(u_2, v_2)) = \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} |I_1(u_1+i, v_1+j) - I_2(u_2+i, v_2+j)|$$

- Sum of squared differences (SSD)

$$SSD(p_1(u_1, v_1), p_2(u_2, v_2)) = \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} (I_1(u_1+i, v_1+j) - I_2(u_2+i, v_2+j))^2$$

Mise en correspondance - fonctions de similarité

Pour une fenêtre de taille $2N + 1 \times 2P + 1$

- Corrélation Croisée (CC) (peut se déduire de la SSD)

$$CC(p_1(u_1, v_1), p_2(u_2, v_2)) = \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} I_1(u_1+i, v_1+j) \cdot I_2(u_2+i, v_2+j)$$

Mise en correspondance - fonctions de dissimilarité

Pour une fenêtre de taille $2N + 1 \times 2P + 1$

- Zero Mean Sum of squared differences (ZSSD)

$$\begin{aligned} ZSSD(p_1(u_1, v_1), p_2(u_2, v_2)) \\ = \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} [(I_1(u_1+i, v_1+j) - \overline{I_1(u_1, v_1)}) \\ - (I_2(u_2+i, v_2+j) - \overline{I_1(u_2, v_2)})]^2 \end{aligned}$$

Mise en correspondance - fonctions de similarité

Pour une fenêtre de taille $2N + 1 \times 2P + 1$

- Zero-Mean normalized cross-correlation (ZNCC)

$$\begin{aligned} ZNCC(p_1(u_1, v_1), p_2(u_2, v_2)) \\ = \frac{1}{\sigma_1 \cdot \sigma_2} \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} (I_1(u_1+i, v_1+j) - \overline{I_1(u_1, v_1)}) \\ \cdot (I_2(u_2+i, v_2+j) - \overline{I_1(u_2, v_2)}) \end{aligned}$$

avec

$$\sigma_1 = \sqrt{\frac{1}{(2N+1)(2P+1)} \sum_{i=-N}^{i=N} \sum_{j=-P}^{j=P} (I_1(u_1+i, v_1+j) - \overline{I_1(u_1, v_1)})^2}$$

repose sur des hypothèses fortes :

- ▶ Les changements de points de vue n'altèrent pas l'aspect des surfaces
- ▶ Pas d'occultations lors de la recherche d'un correspondant.
- ▶ Une région rectangulaire dans l'image 1 correspond à une région rectangulaire dans l'image 2.
- ▶ Deux régions de couleurs constantes présentent une distance normalisée (ZSAD, ZSSD) nulle. Une solution consiste à normaliser non pas la région mais l'ensemble de l'image.

Points d'intérêts : SIFT

- SIFT (**Scale Invariant Feature Transform**) [Lowe, 2004]
 - <http://www.cs.ubc.ca/~lowe/keypoints/>
- Nouvelle approche de points d'intérêts très populaire
 - invariance par échelle
 - invariance par rotation
 - invariance par point de vue,
 - invariance aux conditions d'éclairage
- Points orientés : valeurs + vecteur

Source : David Lowe. Distinctive Image Features from Scale Invariant Keypoints », IJCV, 60(2), 2004. 36

SIFT

Pour une image, on obtient un ensemble de vecteurs avec attributs

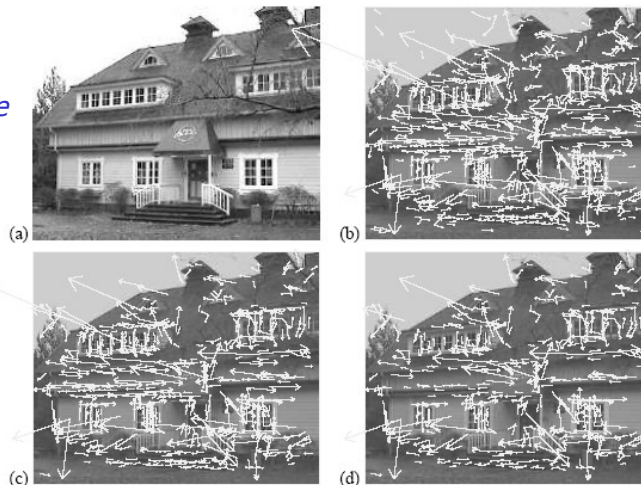
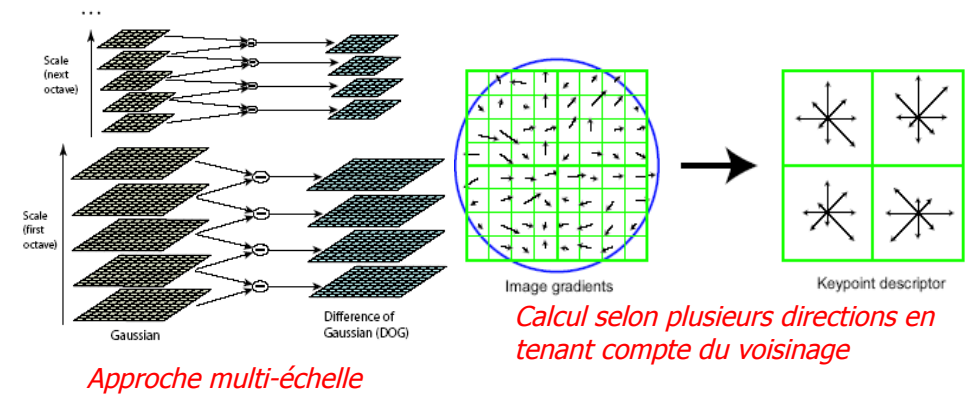


Figure 5: This figure shows the stages of keypoint selection. (a) The 233x189 pixel original image (b) The initial 832 keypoints locations at maxima and minima of the difference-of-Gaussian function. Keypoints are displayed as vectors indicating scale, orientation, and location. (c) After applying a threshold on minimum contrast, 729 keypoints remain. (d) The final 536 keypoints that remain following an additional threshold on ratio of principal curvatures.

Caractéristiques du calcul SIFT



Pour l'algorithme complet, lire les articles disponibles sur <http://www.cs.ubc.ca/~lowe/keypoints/>

Recherche d'objets dans une image



Figure 12: The training images for two objects are shown on the left. These can be recognized in a cluttered image with extensive occlusion, shown in the middle. The results of recognition are shown on the right. A parallelogram is drawn around each recognized object showing the boundaries of the original training image under the affine transformation solved for during recognition. Smaller squares indicate the keypoints that were used for recognition.