

Vision par ordinateur - Reconnaissance Visuelle

Cours 4 : Extraction de primitives - partie 2 : caractéristiques locales et points d'intérêts

Céline Hudelot - Mention IA - CentraleSupélec

2022-2023

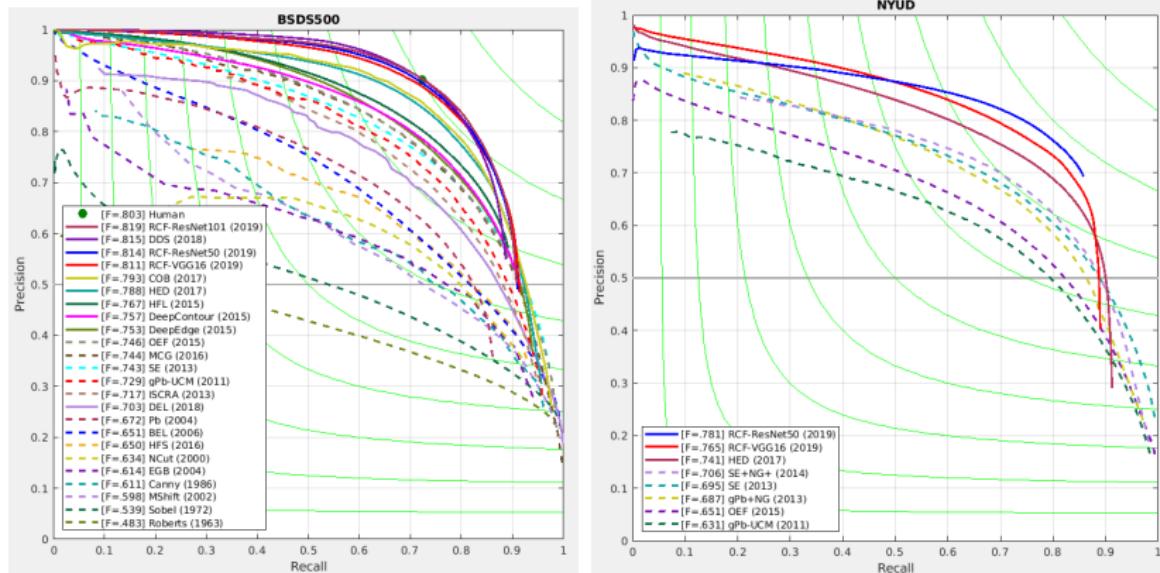
Recap

Contour

- contour : lieu des fortes variations de la fonction d'intensité.
- Plusieurs approches :
 - ▶ Approches discrètes : convolution par un filtre passe-haut + seuillage
 - ★ Gradient : extrema locaux de la norme du gradient.
 - ★ Laplacien : passage à zéros.
 - ▶ Approches optimales : problème d'optimisation multi-critères (détection, localisation, réponse unique)
 - ★ Canny (filtre gaussien)
 - ★ Shen-Castan
 - ★ Deriche

Aujourd'hui : caractéristiques locales et invariantes

Detection de contours aujourd'hui



Source : <https://github.com/MCG-NKU/plot-edge-pr-curves>

Detection de contours aujourd'hui

Plusieurs benchmarks

- Berkeley Segmentation Data Set and Benchmarks 500^a : 500 images naturelles, vérité terrain (annotations humaines) et du code de benchmark.
- NYUD Dataset^b : 1449 paires d'images RVB et de profondeur alignées, étiquetées (images RGBD, capteur : Kinect).
- Multicue Dataset^c : 100 images de scènes naturelles complexes capturées par une caméra stéréo binoculaire.

a. <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html>

b. https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html

c. <https://serre-lab.clps.brown.edu/resource/multicue/>

Une vaste littérature, voir par exemple

<https://github.com/MarkMoHR/Awesome-Edge-Detection-Papers>

Detection de contours aujourd'hui (et un peu hier)

Vue chronologique

- Travaux pionniers : détecteur de Sobel, de Canny (c.f. cours précédent)
- Travaux portant sur les caractéristiques et leur information : hand-design features (Martin et al, PAMI 2004)^a
- Approches basées apprentissage : (Dollár et al, PAMI 2015)^b
- Approches basées apprentissage profond, CNNs, qui mettent en avant l'importance de l'apprentissage de représentation hiérarchique (Xie et al, 2015), (Liu et al, PAMI 2019)^c
- Approches non-supervisées : utilisation du mouvement avec un principe de self-supervision : (Li et al, 2016)^d

a. Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues, <https://people.eecs.berkeley.edu/~efros/courses/AP06/Papers/martin-pami-04.pdf>

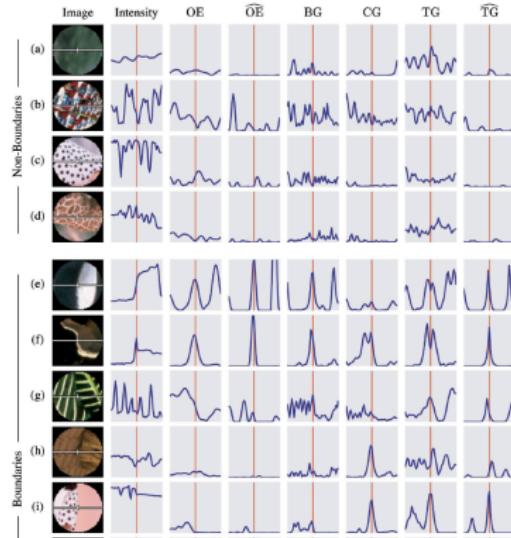
b. Fast Edge Detection Using Structured Forests
<https://arxiv.org/abs/1406.5549>

c. Holistically-Nested Edge Detection <https://arxiv.org/abs/1504.06375> ; Richer Convolutional Features for Edge Detection <https://arxiv.org/abs/1612.02103>

d. Unsupervised Learning of Edges <https://arxiv.org/abs/1511.04166>

Detection de contours aujourd'hui (et un peu hier)

Hand crafted design features - Boundary detection (Martin et al, PAMI 2004)



Etude en chaque pixel des discontinuités locales de plusieurs types de caractéristiques, sur plusieurs échelles et orientations. : OE (oriented energy), BG (brightness gradient), CG (color gradient), TG (raw texture feature) fusionnées à l'aide d'un classifieur supervisé (calcul de $P_b(x, y, \theta)$).

Detection de contours aujourd'hui (et un peu hier)

Détection de contours comme un problème d'apprentissage - Fast Edge Detection Using Structured Forests (Dollár et al, PAMI 2015)

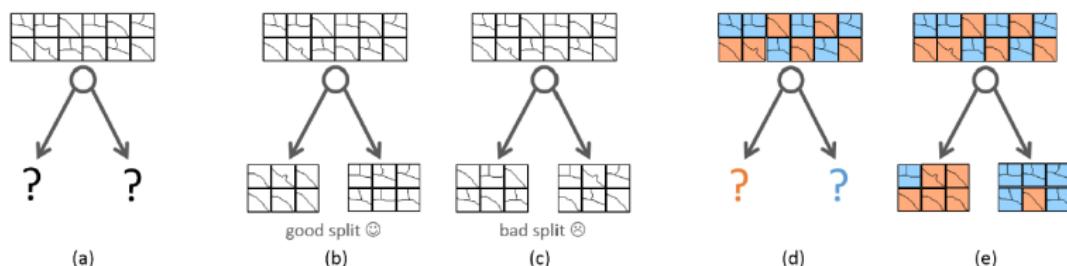
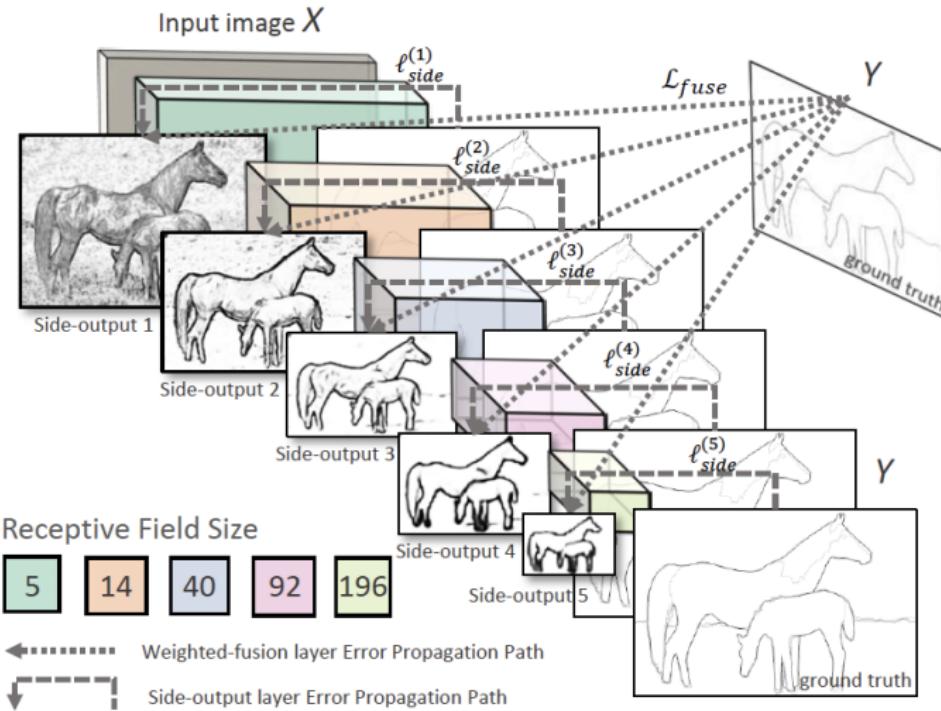


Fig. 2. Illustration of the decision tree node splits: (a) Given a set of structured labels such as segments, a splitting function must be determined. Intuitively a good split (b) groups similar segments, whereas a bad split (c) does not. In practice we cluster the structured labels into two classes (d). Given the class labels, a standard splitting criterion, such as Gini impurity, may be used (e).

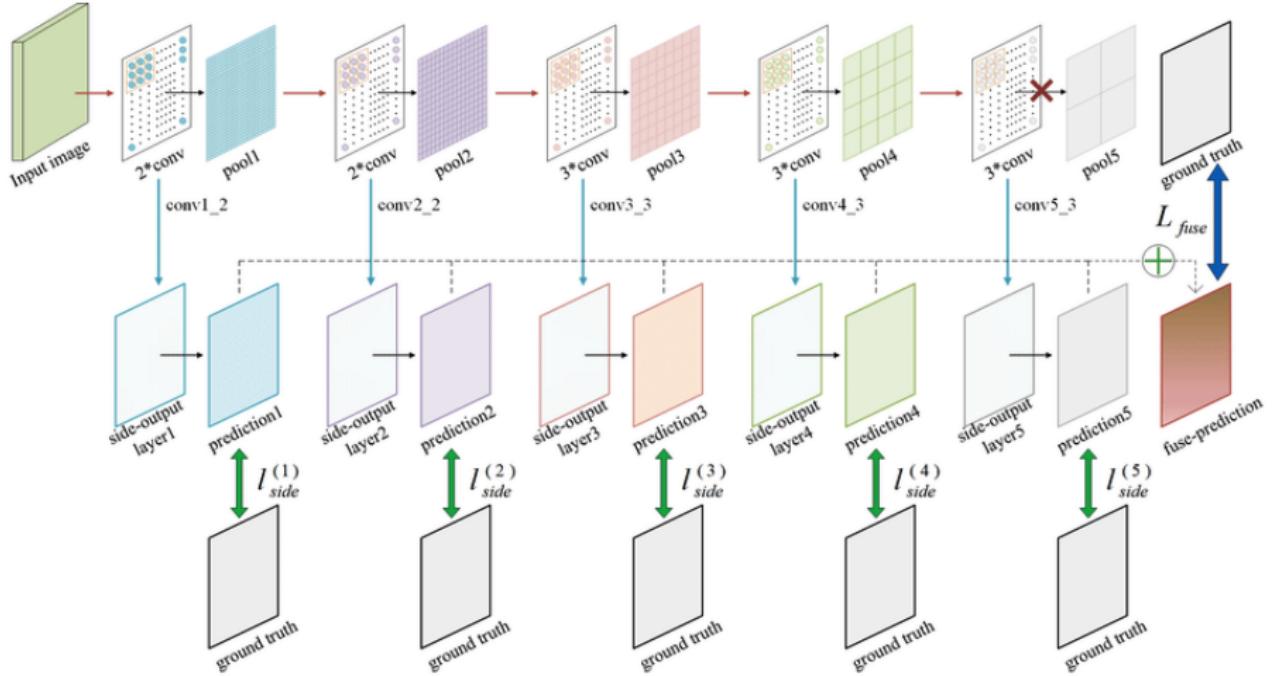
Detection de contours aujourd'hui

Apprentissage profond - Holistically-Nested Edge Detection - (Xie et al, 2015)



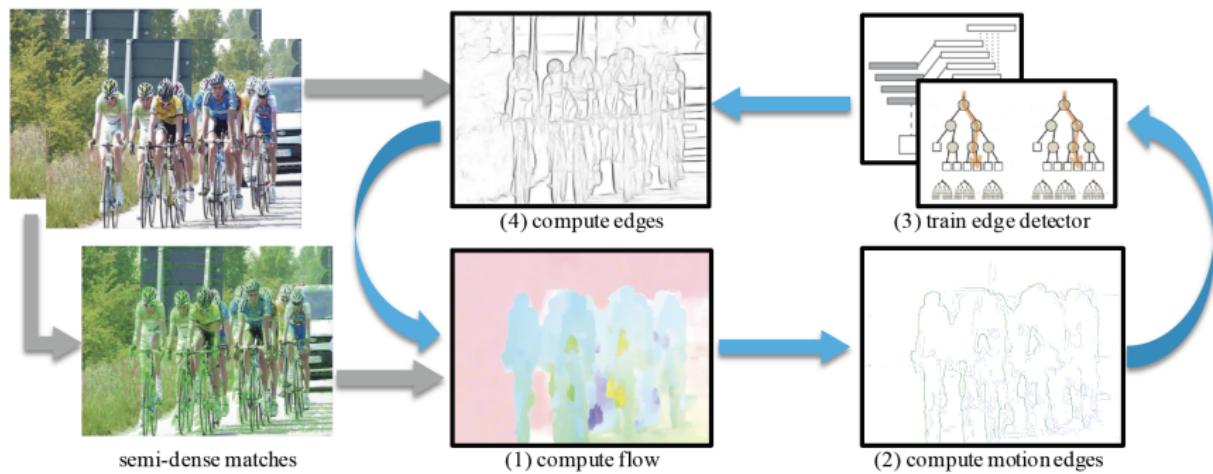
Detection de contours aujourd'hui

Apprentissage profond - Holistically-Nested Edge Detection - (Xie et al, 2015)



Detection de contours aujourd'hui

Apprentissage non supervisé. Unsupervised Learning of Edges, (Li et al, 2016)



Programme du jour

A la recherche de caractéristiques locales robustes et invariantes !

Plan

1 Motivations

2 Points d'intérêt

3 Les coins

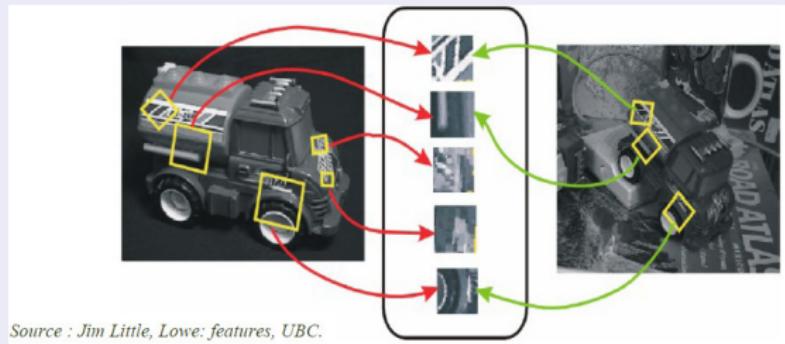
- Moravec
- Harris
- Hessian

4 Invariance à l'échelle

Détection de caractéristiques

Objectifs

- Identifier des caractéristiques robustes dans l'image.
- Motivation principale : mise en correspondance mais aussi reconnaissance, suivi, ...
- Eviter la segmentation qui peut être source d'erreurs.
- Permettre des représentations locales et pas globales.



Détection de caractéristiques

Motivations

- Des caractéristiques locales pour améliorer la robustesse aux :
 - ▶ Occlusions



- ▶ Articulations

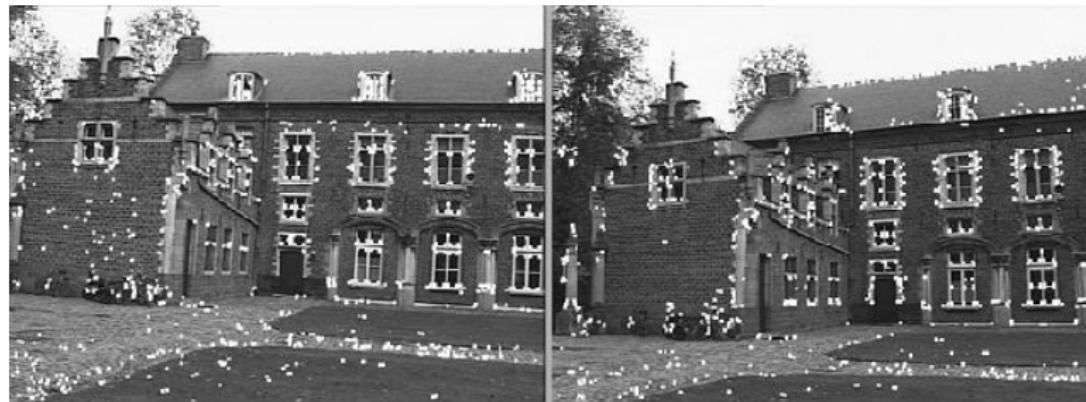


- ▶ Variations au sein d'une même catégorie d'objets.



Détection de caractéristiques

Application : correspondance stéréo



Détection de caractéristiques

Application : recherche d'images contenant un même élément



FIGURE – Source : Ian Reid, Object Detection and Recognition, Oxford

<http://www.robots.ox.ac.uk/~ian/Teaching/ComputerVision/lec8.pdf>

Détection de caractéristiques

Application : recherche d'images contenant un même élément

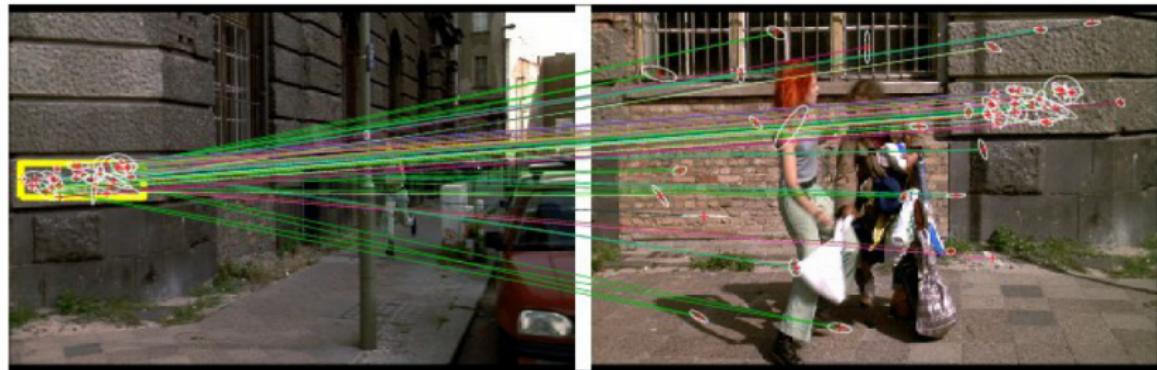


FIGURE – Source : Ian Reid, Object Detection and Recognition, Oxford

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

Détection de caractéristiques

Application : recherche d'images contenant un même élément

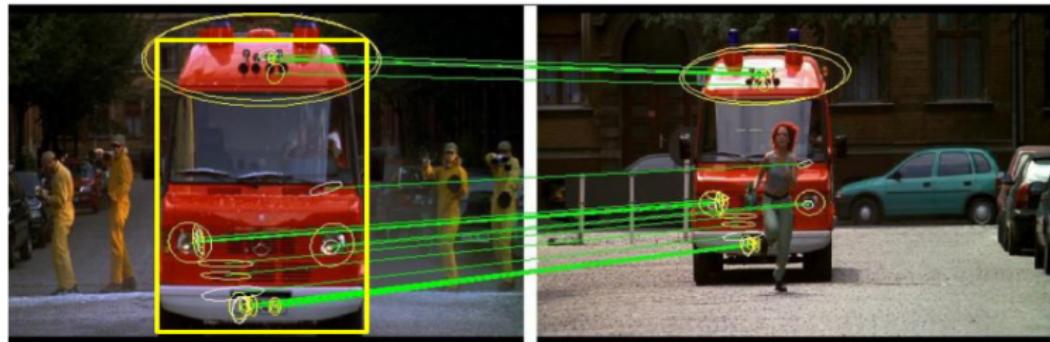
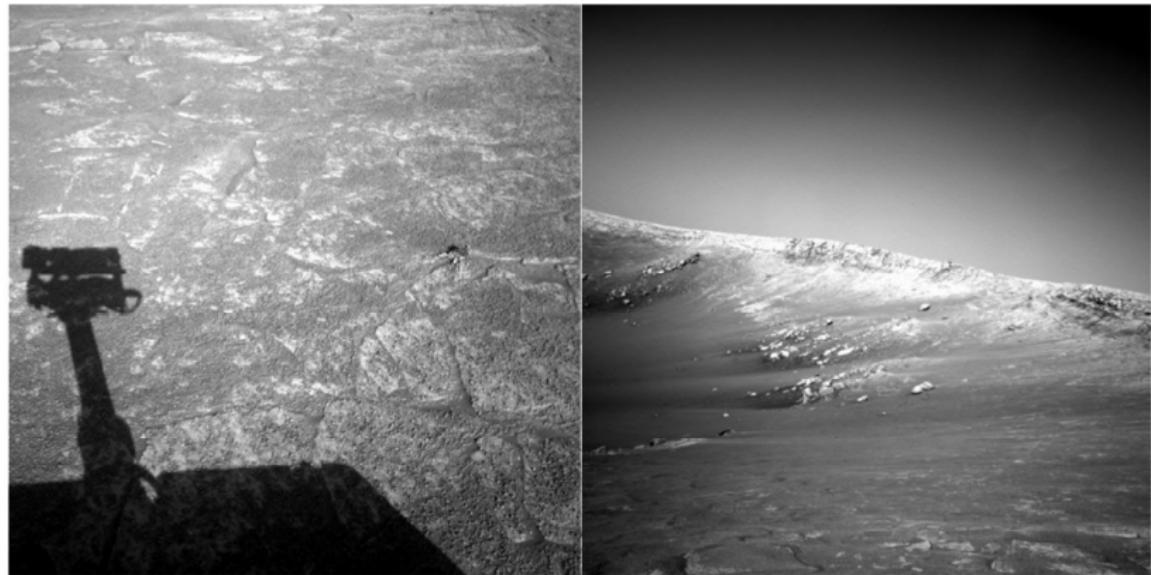


FIGURE – Source : Ian Reid, Object Detection and Recognition, Oxford

Même autobus mais occlusion et taille différente.

Détection de caractéristiques

Mise en correspondance d'images : encore plus difficile

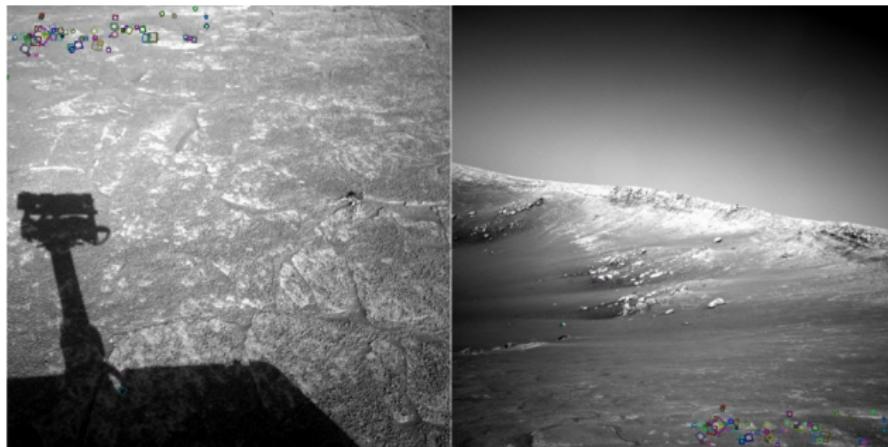


NASA Mars Rover images

FIGURE – Source : Steve Seitz

Détection de caractéristiques

Mise en correspondance d'images : encore plus difficile



NASA Mars Rover images with SIFT feature matches
(Figure by Noah Snavely)

FIGURE – Source : Steve Seitz

La mise en correspondance peut se faire dans ce cas grâce aux détecteurs SIFT (carrés colorés dans les images)

Détection de caractéristiques

Application : création de panorama



FIGURE – Source : Daria Frolova

Détection de caractéristiques

Application : création de panorama

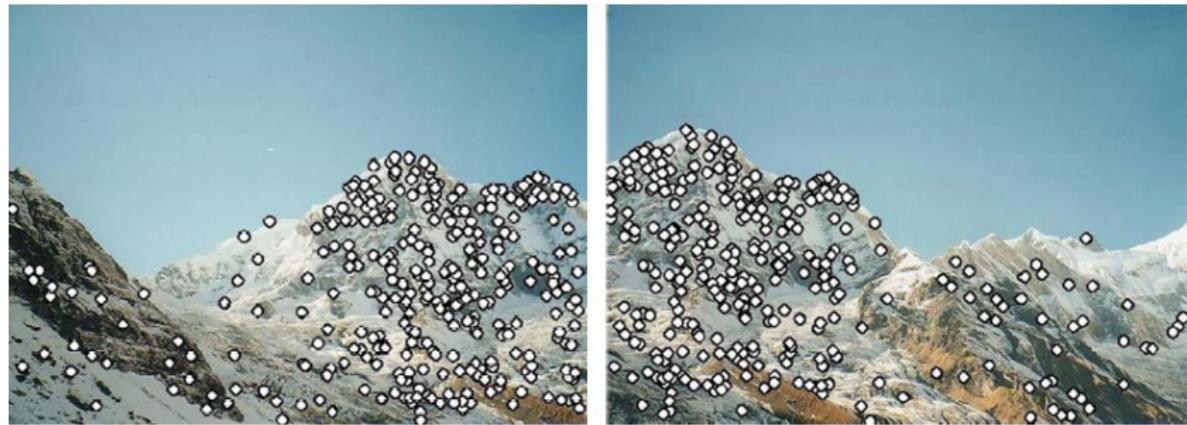


FIGURE – Source : Daria Frolova

- Détection de points caractéristiques dans chaque image.

Détection de caractéristiques

Application : création de panorama

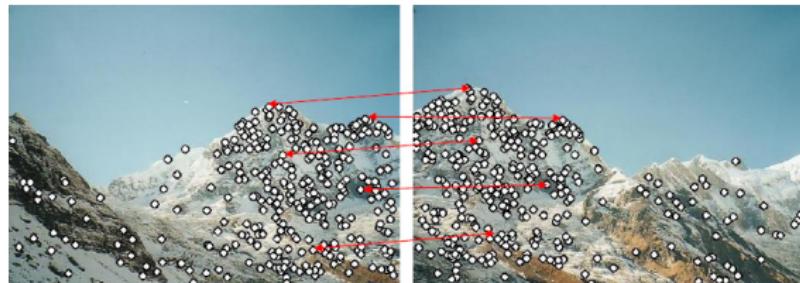


FIGURE – Source : Daria Frolova

- Détection de points caractéristiques dans chaque image.
- Chercher les paires correspondantes.

Détection de caractéristiques

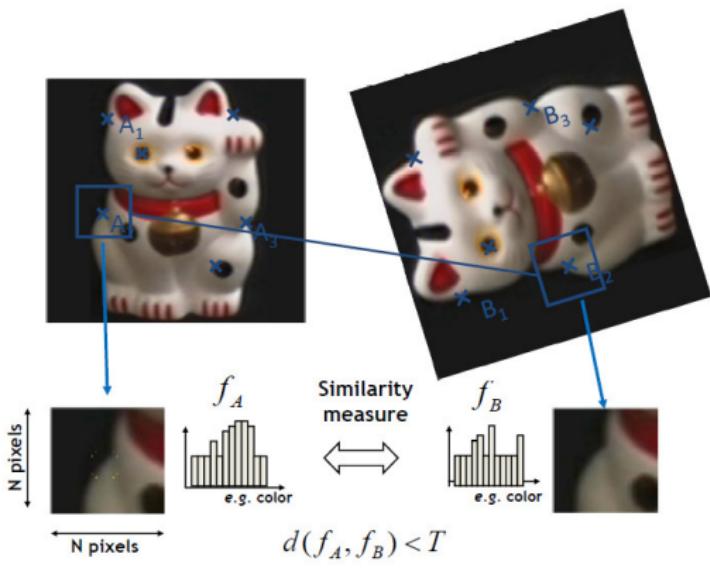
Application : création de panorama



FIGURE – Source : Daria Frolova

- Détection de points caractéristiques dans chaque image.
- Chercher les paires correspondantes.
- Utiliser les paires pour l'alignement.

Détection de caractéristiques



Approche générale

- ➊ Trouver un ensemble de points d'intérêts distinctifs (**feature detection**).
- ➋ Définir une région autour de chaque point.
- ➌ Extraire et normaliser le contenu de la région.
- ➍ Calculer un descripteur local de la région (**feature description**).
- ➎ Mettre en correspondance les descripteurs locaux (**feature matching**).

Détection de caractéristiques

Propriétés idéales des caractéristiques

- **Localité** : une caractéristique locale (vs globale) sera plus robuste aux occlusions.
- **Invariance** : à différentes transformations : translation, rotation, changements d'échelle.
- **Robustesse** : 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 *intéressantes* par objet ou portion d'images.
- **Précision** : position précise dans l'image.
- **Efficacité** : calcul rapide.

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

<http://homes.esat.kuleuven.be/~tuytelaa/tutorial-ECCV06.pdf>

Détection de caractéristiques

Propriétés d'invariance

- Invariance à la géométrie :

- ▶ Rotation



- ▶ Changement d'échelle



- ▶ Transformation affine



- Invariance à la photométrie :

- ▶ Changement affine d'intensité ($I \rightarrow al + b$)



Répétabilité des points

Problème à résoudre 1

Détecter le même point de manière indépendante dans les deux images.



no chance to match!

FIGURE – Source : Daria Frolova

⇒ : besoin d'un **détecteur fiable**.

Descripteur robuste et discriminant

Problème à résoudre 2

Pour chaque point trouver la bonne correspondance.



FIGURE – Source : Daria Frolova

⇒ : besoin d'un **descripteur** fiable et discriminant.

Plan

1 Motivations

2 Points d'intérêt

3 Les coins

- Moravec
- Harris
- Hessian

4 Invariance à l'échelle

Qu'est ce qu'un point d'intérêt ?

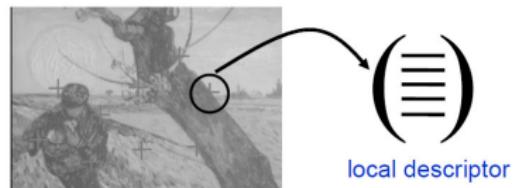
Motivation

Vers des descripteurs photométriques locaux et invariants.

- Local : robuste aux occlusions et sans segmentation.
- Photométrique : distinctif.
- Invariant : aux transformations d'images et changements d'illuminations.

D'après C. Schmid et D. Lowe, CVPR 2003 Tutorial

<http://lear.inrialpes.fr/~schmid/cvpr-tutorial03/>



Qu'est ce qu'un point d'intérêt ?

Motivation

Quels points ? Buts :

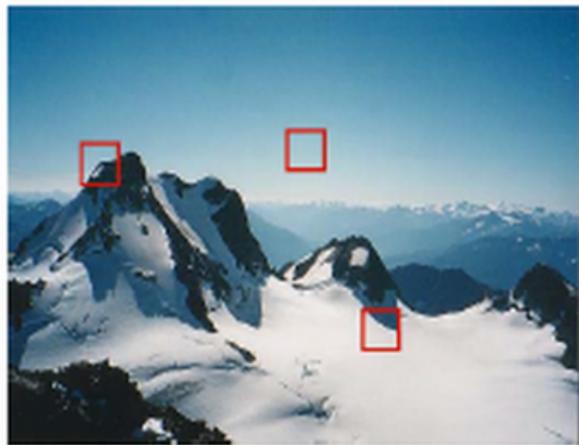
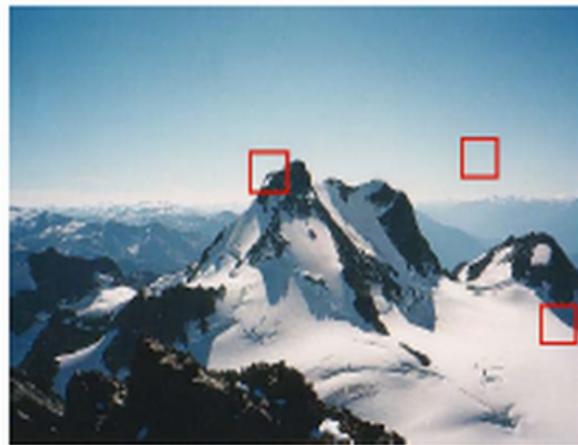
- Répétabilité de la détection.
- Localisation précise.
- Un contenu intéressant.



What are the good features to track?

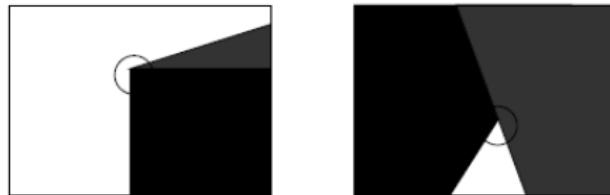
Qu'est ce qu'un point d'intérêt ?

What are the good features to track ?



Source : Szeliski

Qu'est ce qu'un point d'intérêt ?



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

Définition

(Moravec) : Un point d'intérêt est défini lorsque les valeurs de l'intensité varient beaucoup dans toutes les directions.

Pourquoi les points d'intérêt ?

Avantages

- Caractéristiques locales et invariantes
 - ▶ Pas uniquement une méthode pour sélectionner des parties intéressantes de l'image ou pour diminuer le temps de calcul du traitement.
 - ▶ Une nouvelle représentation des images permettant de décrire les objets sans passer par une segmentation.
- Sources d'informations plus fiables que les contours car plus de contraintes sur la fonction d'intensité.
- Plus faciles à extraire que les contours (pas besoin de chaînage).
- Présents dans une grande majorité d'images à la différence des contours.
- Robustesse aux occultations.

Détection des points d'intérêts

Plusieurs approches

- ① **Approche contour** : détection des contours puis extraction des points d'intérêts le long des contours en considérant les points de courbure maximale ainsi que les intersections de contours.
- ② **Approche intensité** : à partir des niveaux de gris de l'image, trouver un opérateur qui est maximal aux points d'interêts.
- ③ **Approche à base de modèles** : identification des points d'intérêts par mise en correspondance de la fonction d'intensité avec un modèle théorique de la fonction des points d'intérêts considérés.

Les approches de la seconde catégorie sont les plus répandues car :

- indépendance vis à vis de la détection de contours.
- indépendance vis à vis du type de points d'intérêt.

Détection des points d'intérêts

Etat de l'art

- En niveaux de gris :

- ▶ Moravec (1977)
- ▶ Beaudet (1978)
- ▶ Kitchen et Rosenfeld (1982)
- ▶ Harris et Stephens (1988)
- ▶ Deriche et Faugeras (1990)
- ▶ Heitger (1992)
- ▶ Förstner (1994)
- ▶ Harris précis (1996)
- ▶ Susan (1997)
- ▶ SIFT (1999,2004)
- ▶ SURF (2006) et ...

- En couleur :

- ▶ Kitchen et Rosenfeld (1998)
- ▶ Harris précis (1998)

Des détecteurs (et descripteurs) qui sont devenus **une brique de base** de nombreuses applications en vision.

Plan

1 Motivations

2 Points d'intérêt

3 Les coins

- Moravec
- Harris
- Hessian

4 Invariance à l'échelle

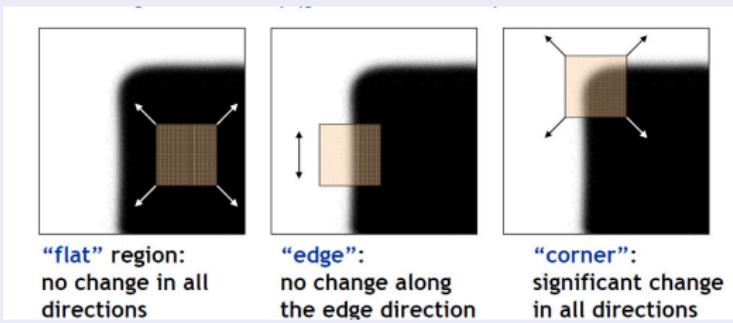
Les coins : des points d'intérêts distinctifs

Propriété clé

Dans la région autour d'un coin, le gradient de l'image a deux ou plusieurs directions dominantes.

Principe de la détection

- Le point devrait pouvoir être reconnu facilement en cherchant dans une petite fenêtre (*localité*).
- Déplacement de la fenêtre dans chaque direction devrait donner de grands changements en intensité (*bonne localisation*)



Plan

1 Motivations

2 Points d'intérêt

3 Les coins

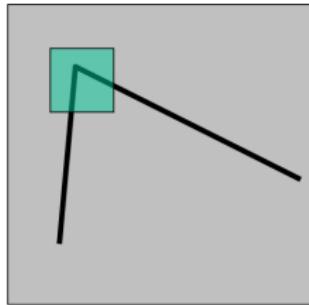
- Moravec
- Harris
- Hessian

4 Invariance à l'échelle

Idée générale

Principe

- On peut facilement reconnaître un coin en regardant dans une petite fenêtre autour de ce coin.
- Déplacer une fenêtre dans toutes les directions peut donner lieu à de grands changements d'intensité.



Detecteur de Moravec

Définition

Variation moyenne de l'intensité pour un petit déplacement (u, v) :

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

- w est une fenêtre définissant le voisinage (valeur 1 à l'intérieur de la fenêtre et 0 à l'extérieur ou fenêtre gaussienne (Harris)).
- processus répété en chaque point (x, y) de l'image, pour divers déplacements (u, v)

Window function $w(x, y) =$



1 in window, 0 outside

or

Gaussian

Source : Szeliski

Detecteur de Moravec

Exemple

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

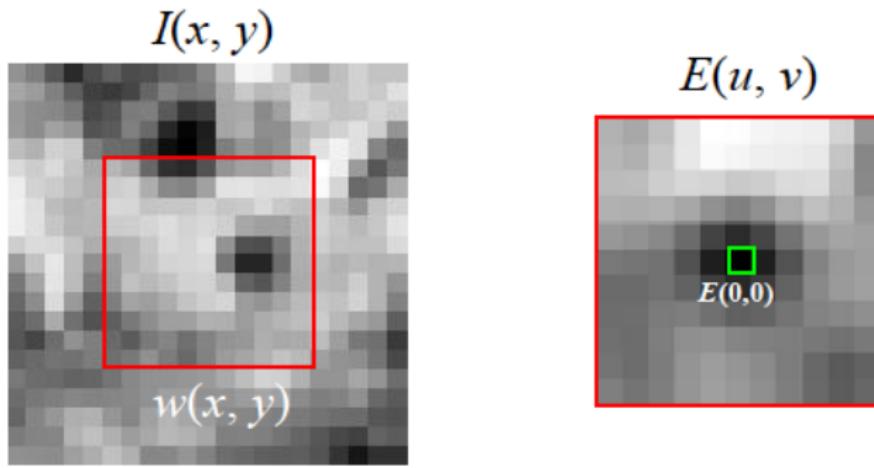
Pour le pixel $p = 6$ de la fenêtre rouge et pour un déplacement $dx = 1, dy = 1$, on calculera la somme des carrés des différences des couples

$$(1-6), (2-7), (3-8), (5-10), (6-11), (7-12), (9-14), (10-15), (11-16)$$

Detecteur de Moravec

Changement d'apparence de la fenêtre $w(x, y)$ pour le déplacement (u, v) .
On calcule la variation moyenne d'intensité entre la région w centrée en (x, y) et celle centrée en $(x + u, y + v)$.

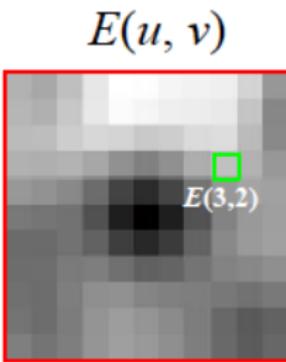
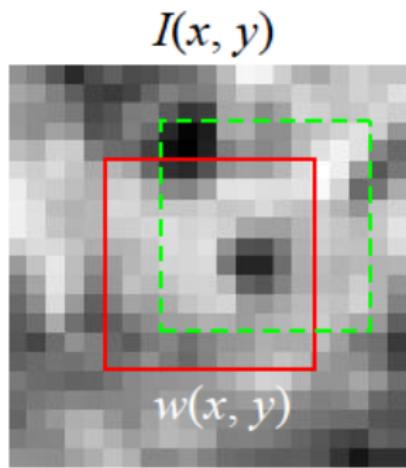
$$E(u, v) = \sum_{x,y} w(x, y) |I(x + u, y + v) - I(x, y)|^2$$



Detecteur de Moravec

Changement d'apparence de la fenêtre $w(x, y)$ pour le déplacement (u, v)

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



Detecteur de Moravec

Visualisation de E

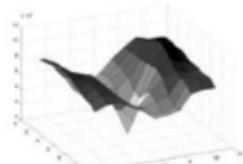
$E(u, v)$



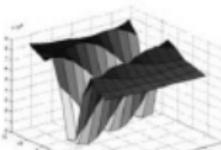
(a)



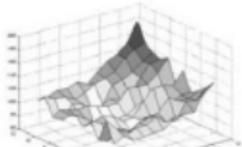
$E(u, v)$



(b)



(c)



(d)

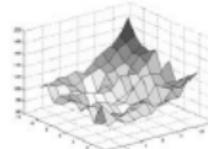
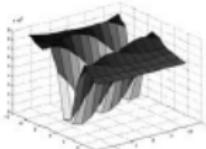
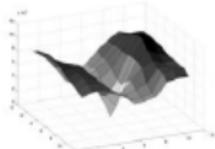
Inverted, and in 3D!

Detecteur de Moravec

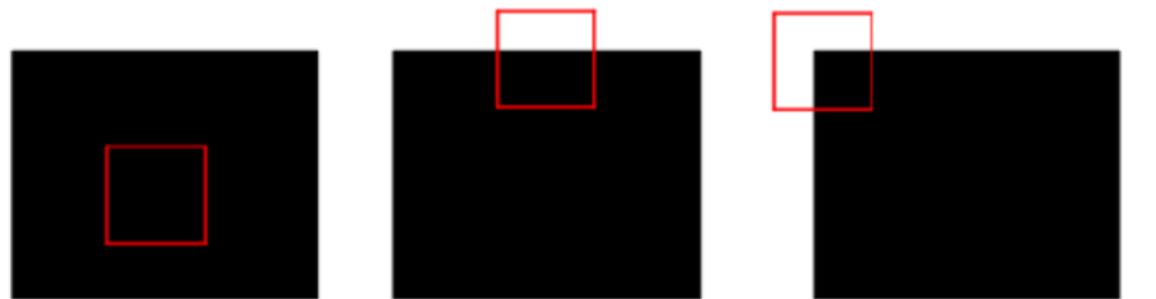
A quelles parties de l'image les surfaces correspondent ?



(a)



Detecteur de Moravec



- Zone d'intensité presque constante : $E(u, v) \approx 0$.
- Contour :
 - ▶ $E(u, v) \approx 0$ pour des déplacements le long du contour ($y \neq 0$).
 - ▶ $E(u, v) > 0$ pour des déplacements perpendiculaires.
- Coin : $E(u, v) > 0$ pour tout $(u, v) \neq (0, 0)$

Detecteur de Moravec

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

Algorithme

- Pour chaque pixel (x, y) , calculer les variations d'intensité $E(u, v)$ pour tous les déplacements : $(u, v) \in \{(1, 0), (1, 1), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1)\}$.
On notera : $E_{u,v}(x, y)$ et on calcule donc pour chaque pixel
 $E_{1,0}, E_{1,1}, E_{0,1}, E_{-1,1}, E_{-1,0}, E_{-1,-1}, E_{0,-1}, E_{1,-1}$
- Construire la carte de *coinité* en calculant la mesure $C(x, y)$ pour chaque pixel (x, y) : $C(x, y) = \min_{u,v} E_{u,v}(x, y)$ (**minimum de E par rapport aux décalages**)
- Trouver les maxima (ou les points supérieurs à un seuil) de cette carte : ils correspondent aux points d'intérêts (**maxima locaux de $C(x, y)$**).

Detecteur de Moravec

Exemple (dans le Lab)

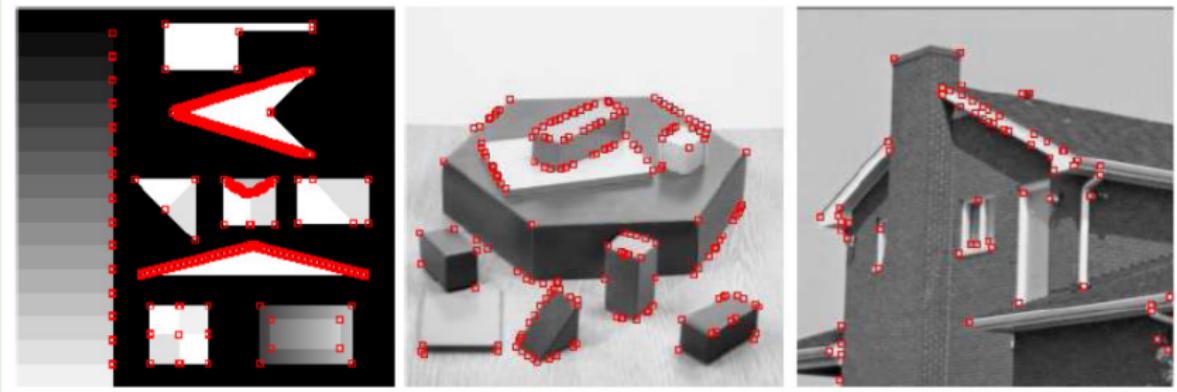
On considère l'image I suivante :

$$I = \begin{pmatrix} 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 0 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 0 & 255 & 0 & 0 & 0 & 0 & 255 & 255 & 255 & 255 & 255 \\ 0 & 0 & 0 & 0 & 0 & 0 & 255 & 255 & 255 & 255 & 255 \\ 0 & 0 & 0 & 0 & 0 & 0 & 255 & 255 & 255 & 255 & 255 \\ 0 & 0 & 0 & 0 & 0 & 0 & 255 & 255 & 255 & 255 & 255 \end{pmatrix}$$

- ➊ Calculer $E(u,v)$ et $C(x,y)$ aux pixels en gras de l'image ci-dessus.
- ➋ Que représentent ces points ?
- ➌ Que pensez-vous des performances du détecteur de Moravec ?

Detecteur de Moravec

Exemple



Detecteur de Moravec

Quelques limitations

- Problème 1 : Beaucoup de bruit lié à l'utilisation d'une fenêtre binaire.
- Problème 2 : Seuls des déplacements à 45 degrés sont considérés.
- Problème 3 : Seul le minimum de E est pris en compte (et donc grande sensibilité aux contours).

Plan

1 Motivations

2 Points d'intérêt

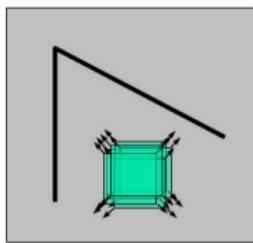
3 Les coins

- Moravec
- Harris
- Hessian

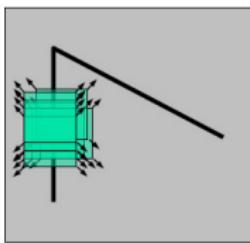
4 Invariance à l'échelle

Detecteur de Harris

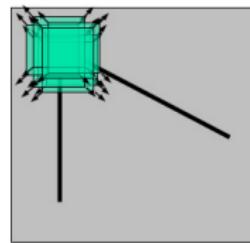
- Amélioration du détecteur de Moravec :
 - ▶ Introduction d'une fenêtre gaussienne ([réponse au problème 1](#)).
 - ▶ Se base sur l'idée de l'auto-corrélation
- Un coin = changements dans plusieurs directions ([n'importe quel angle : réponse au problème 2](#)).



Zone “homogène” : aucun changement dans aucune direction



“contour” : changement dans une seule direction



“coin” : changements dans plusieurs directions

C.Harris and M. Stephens. *A Combined Corner and Edge Detector*. 1988.
<http://www.bmva.org/bmvc/1988/avc-88-023.pdf>

Du détecteur de Moravec au détecteur de Harris

- Déplacement d'une fenêtre dans toutes les directions et recherche des changements d'intensité (*n'importe quel angle : réponse au problème 2*).
- On considère le développement de Taylor de la fonction d'intensité I au voisinage du pixel (x, y) :

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

- D'où :

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

- On néglige le terme $o(u^2, v^2)$ (valide pour les petits déplacements), et on obtient :

$$E(u, v) = Au^2 + 2Cuv + Bv^2$$

- Avec $A = \sum_{x,y} w(x, y) I_x^2(x, y)$; $B = \sum_{x,y} w(x, y) I_y^2(x, y)$ et $C = \sum_{x,y} I_x I_y(x, y)$
- w : fenêtre gaussienne

Du détecteur de Moravec au détecteur de Harris

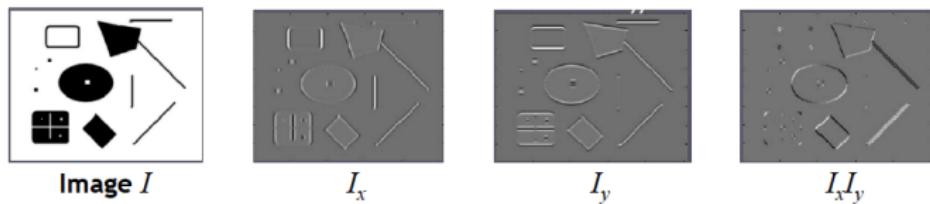
- $E(u, v) = (u, v).M.(u, v)^T$

- Avec :

$$M = \begin{pmatrix} A & C \\ C & B \end{pmatrix} = \sum_{x,y} w(x, y) \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

La matrice M , appelée matrice d'auto-corrélation, représente la structure locale de la fonction $E(x, y)$ au voisinage du pixel (x, y) .

I_x : dérivée en X de l'image, I_y : dérivée en Y de l'image.



Detecteur de Harris

Principe

- Calcul basé sur l'auto-corrélation de l'image multipliée par une fonction de lissage (gaussienne).
- Calcul sur une fenêtre de taille définie :

$$M = G(\sigma) * (\nabla I)(\nabla I)^T = G(\sigma) * \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

- I_x : dérivée en X de l'image, I_y : dérivée en Y de l'image.
- Lissage des dérivées (gradients) sur une fenêtre de taille définie.
- Capture la structure du voisinage local.
- Les **valeurs propres** (λ_1, λ_2) de la matrice renseignent sur les caractéristiques de l'image :
 - ▶ Une valeur propre élevée : contour.
 - ▶ Deux valeurs propres élevées : coin.

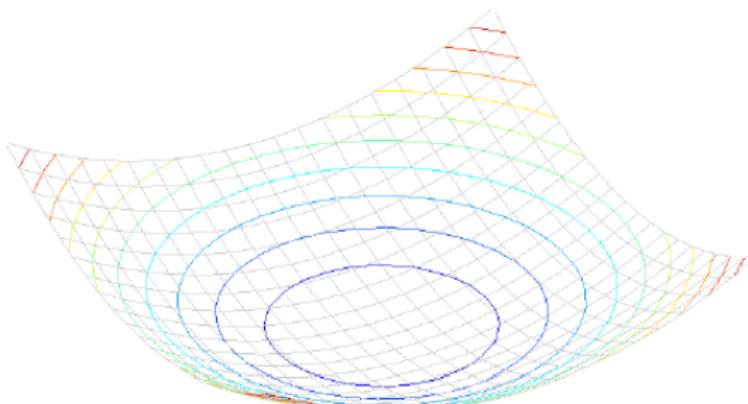
Réponse au problème 3 : on considère une nouvelle mesure de la présence d'un point : analyse des valeurs propres de la matrice M

Detecteur de Harris

Interprétation de M

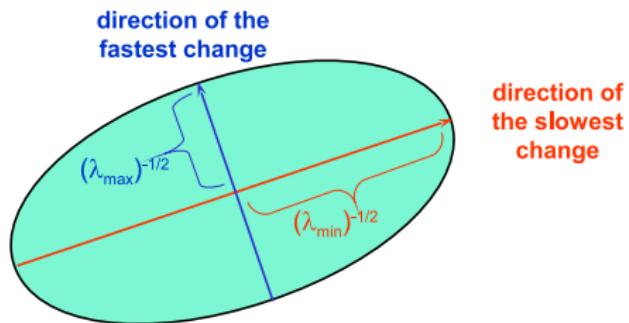
Si on considère une coupe horizontale de la forme quadratique approximant $E : E(u, v) = \text{cste}$.

On a l'équation d'une ellipse.



Detecteur de Harris

- Comme M est symétrique, on a $M = R^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} R$
Décomposition en valeurs propres.
- On peut visualiser M comme une ellipse ($E(u, v) = \text{cste}$) dont les axes sont déterminés par les valeurs propres et l'orientation par R .



Source : Daria Frolova, Invariant feature detectors and descriptors.

[http://www.wisdom.weizmann.ac.il/~vision/courses/2004_2/files/
InvariantFeatures.ppt](http://www.wisdom.weizmann.ac.il/~vision/courses/2004_2/files/InvariantFeatures.ppt)

Detecteur de Harris

Interprétation des valeurs propres

Les valeurs propres de M correspondent aux courbures principales associées à E .

- **Intensité presque constante** : les deux courbures sont de faible valeur.
- **Contour** : une des courbures est de forte valeur, l'autre de faible valeur.
- **Point** : les deux courbures sont de fortes valeurs.

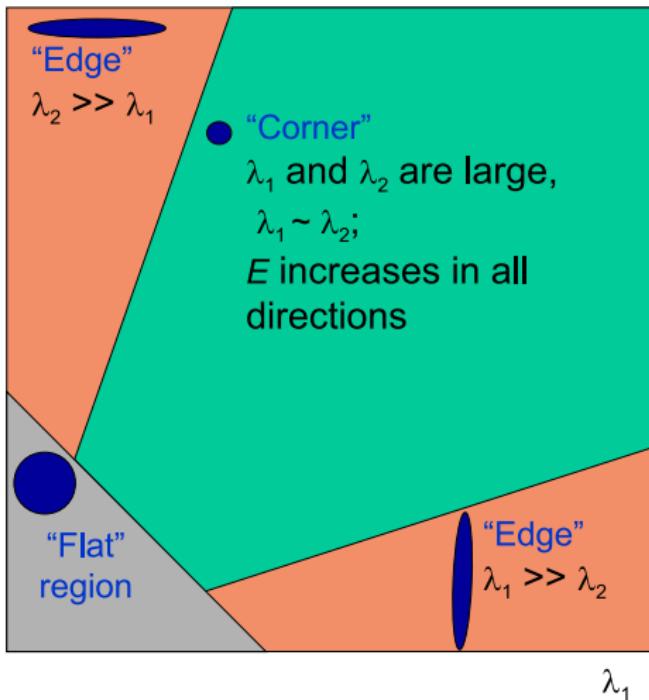
Detecteur de Harris

Interprétation des valeurs propres.

Classification of image points using eigenvalues of M :

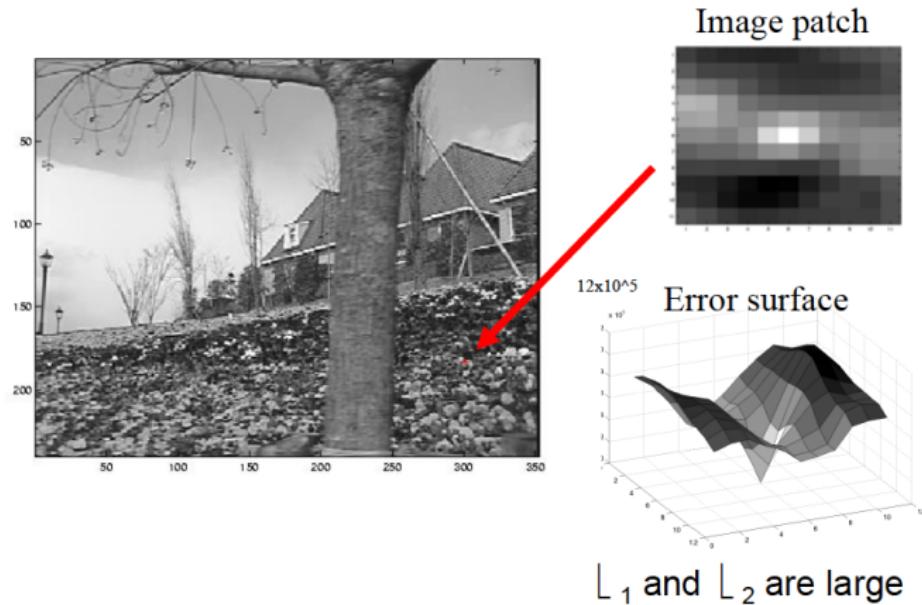
 λ_2

λ_1 and λ_2 are small;
E is almost constant in all directions



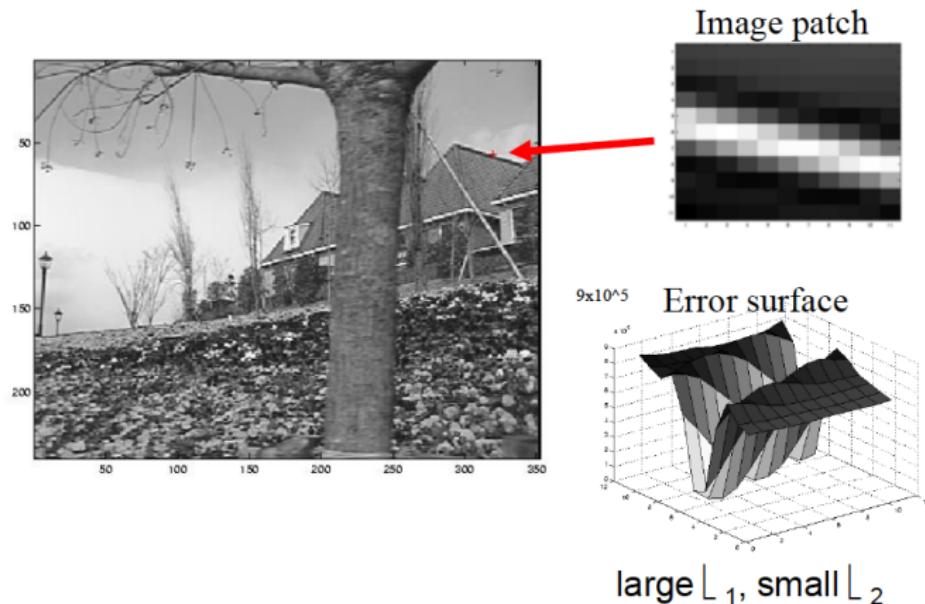
Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris

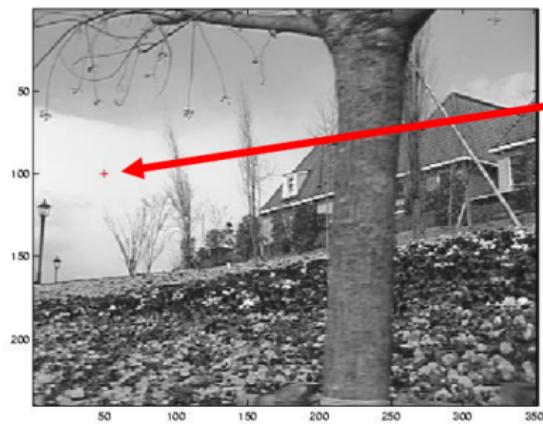
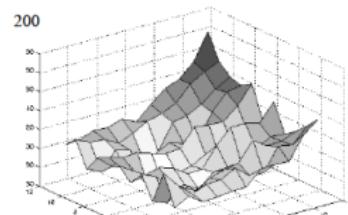


Image patch



Error surface
(vertical scale exaggerated relative to previous plots)
small L_1 , small L_2

Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris

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

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

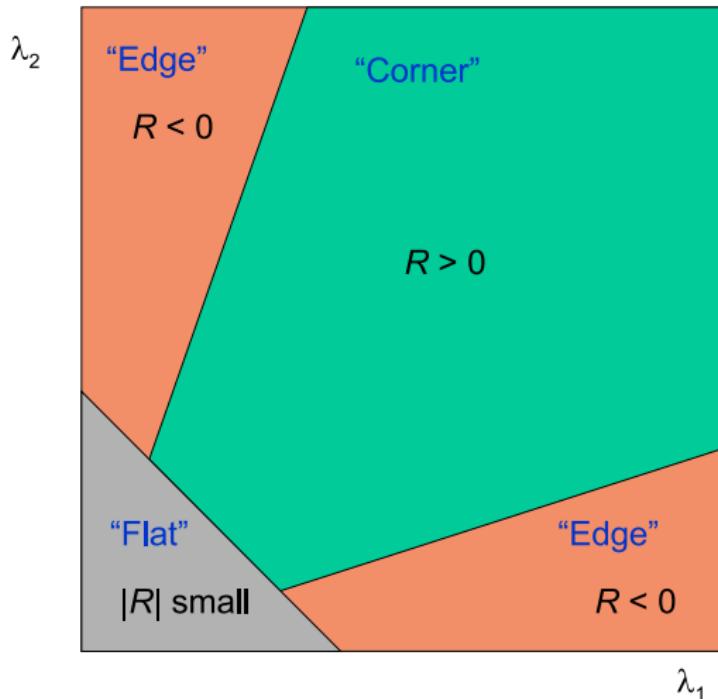
$$\text{trace}(M) = A + B = \lambda_1 + \lambda_2$$

- On calcule alors 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.
- Constante $k = 0.04$ (typiquement) (k permet de réduire l'effet des contours *forts*).
- Points d'intérêts = maxima locaux de R .

Detecteur de Harris



Source : Daria Frolova, Invariant feature detectors and descriptors.

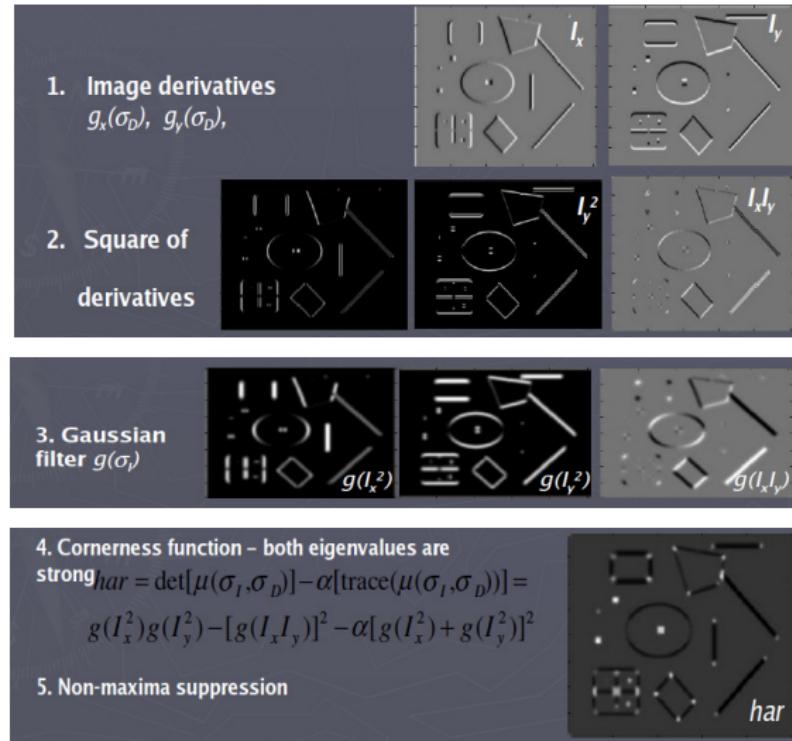
<http://www.wisdom.weizmann.ac.il/~daryaf/InvariantFeatures.ppt>

Detecteur de Harris : schéma de fonctionnement

Algorithme

- ① Calcul du gradient horizontal et vertical de l'image (I_x, I_y) : cela peut se faire par convolution de l'image originale avec un filtre de dérivée de gaussiennes (c.f. cours précédent).
- ② Calcul des trois images correspondant au produit des gradients.
- ③ Convolution de chaque image avec un noyau gaussien (plus large).
- ④ Calcul de R .
- ⑤ Trouver les maxima locaux au dessus d'un certain seuil : ce sont les localisations des points d'intérêts.

Detecteur de Harris : schéma de fonctionnement



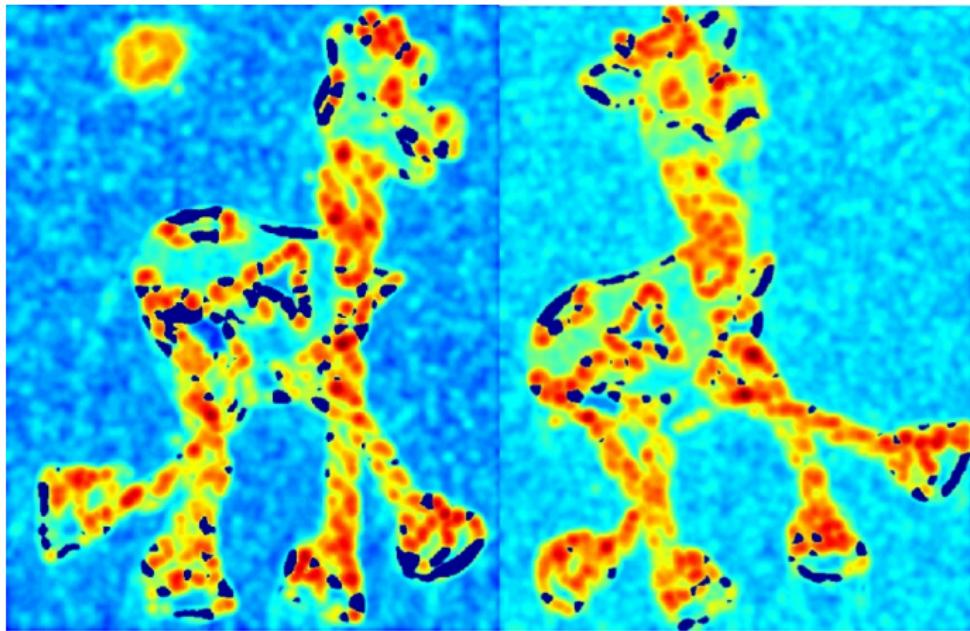
Detecteur de Harris : Exemple de fonctionnement



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris : Exemple de fonctionnement

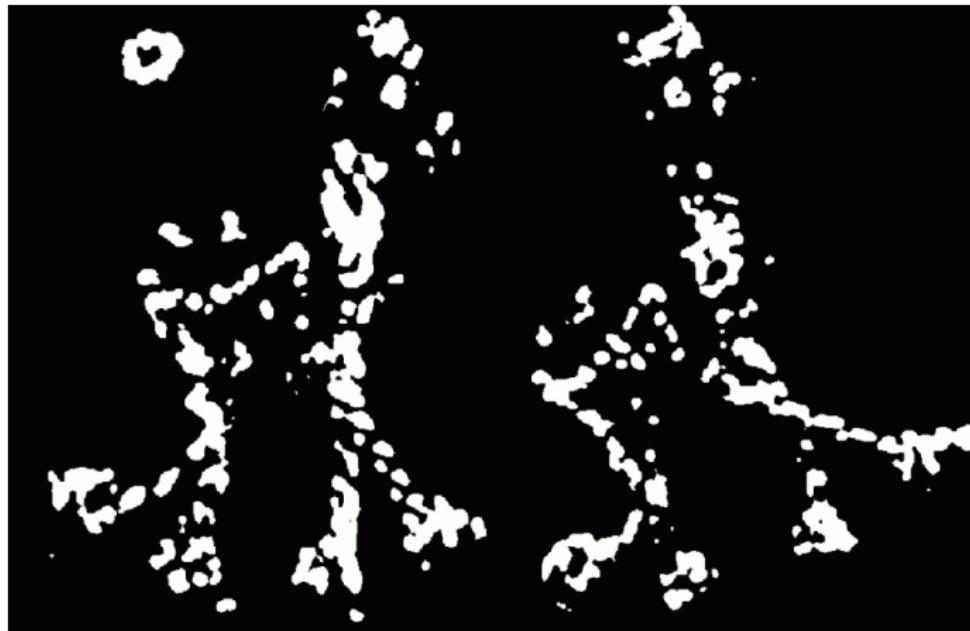
Calcul de R



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris : Exemple de fonctionnement

Seuillage des pixels tel que $R > \text{seuil}$



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris : Exemple de fonctionnement

Recherche des maxima locaux de R



Source : Daria Frolova, Invariant feature detectors and descriptors.

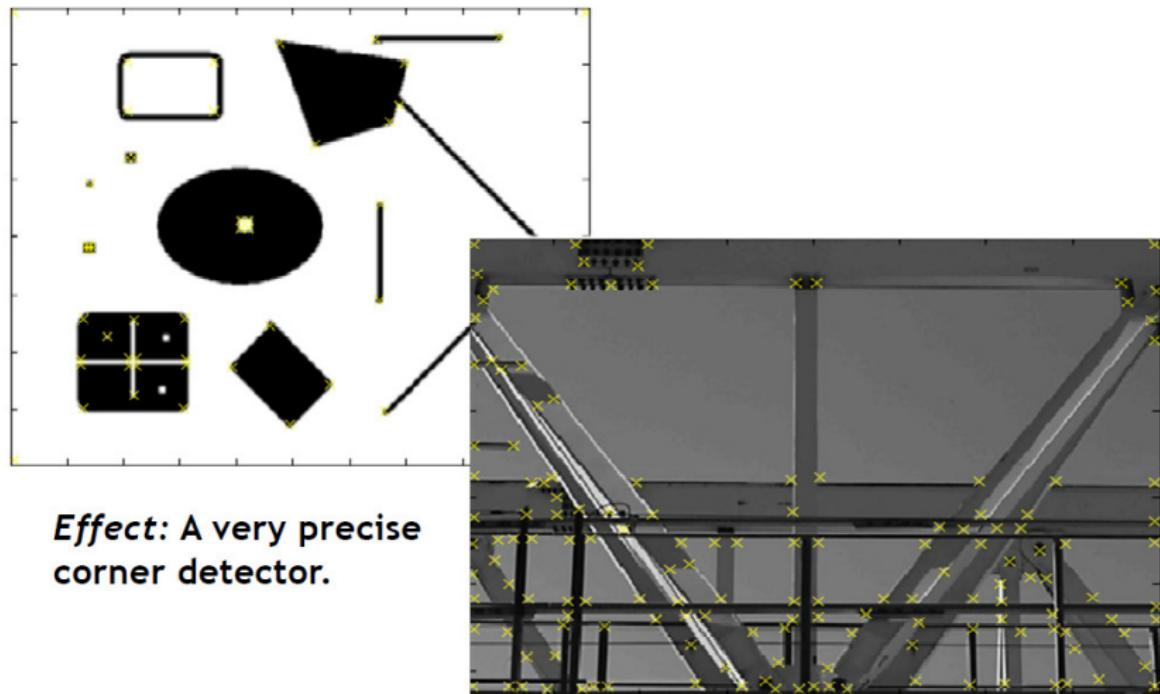
Detecteur de Harris : Exemple de fonctionnement

Superposition aux images originales



Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris : exemple



Effect: A very precise corner detector.

FIGURE – Source : Krystian Mikolajczyk

Detecteur de Harris : exemple



Detecteur de Harris : exemple

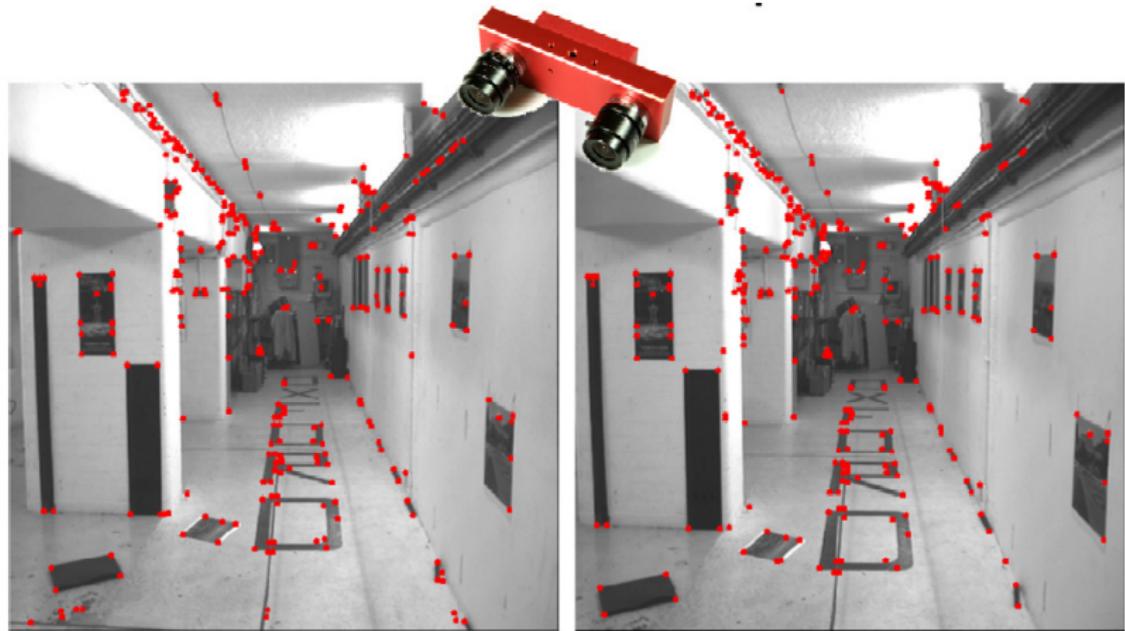


FIGURE – Source : Kristen Grauman

Les résultats sont bons pour la stéréo-correspondance.

Detecteur de Harris : Propriétés

Invariance et covariance

On veut que le détecteur de points d'intérêts soit :

- **Invariant** aux transformations photométriques : l'image est transformée mais la localisation des points ne changent pas.
- **Covariant** aux transformations géométriques : l'image est transformée et les points doivent être détectés à des positions correspondantes.

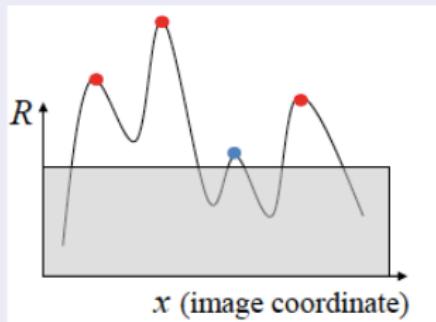
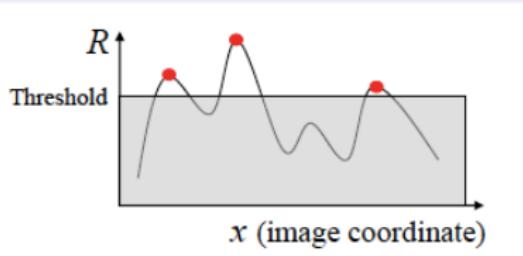


Detecteur de Harris : Propriétés

Changement affine d'intensité

$$\square \rightarrow \blacksquare$$

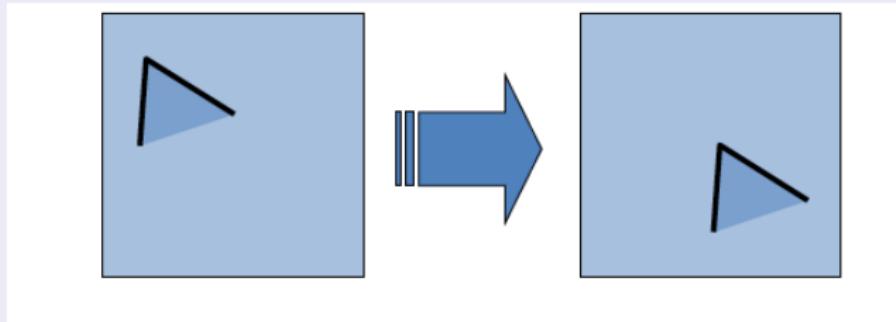
$$I \rightarrow a I + b$$



Invariance partielle (translation d'intensité)

Detecteur de Harris : Propriétés

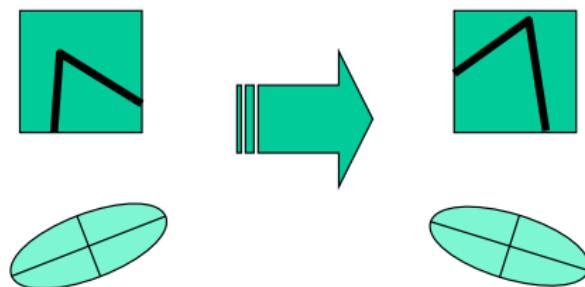
Translation



Dérivée et fenêtre donc invariance à la translation.

Detecteur de Harris : Propriétés

Invariant en rotation (filtre gaussien).

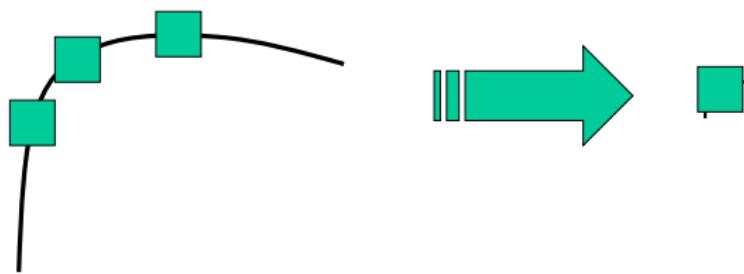


Même après rotation, la forme de l'ellipse reste la même et les valeurs propres de la matrice aussi.

Source : Daria Frolova, Invariant feature detectors and descriptors.

Detecteur de Harris : Propriétés

Non invariance aux changements d'échelle



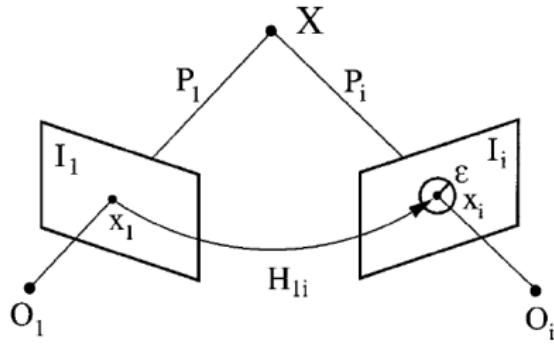
All points will be
classified as **edges**

Corner !

Qualité de détection

Repétabilité

- C. Schmid et al : *Evaluation of Interest Point Detectors*, IJVC 2000
- Un point répétable est détecté dans toutes les vues d'une scène (rotation, zoom...)
- Point répétable = sa projection 2D via une homographie est dans un ϵ voisinage d'un point d'intérêt de l'image cible
- Taux de répétabilité = nb points répétés / (min) nb point détectés

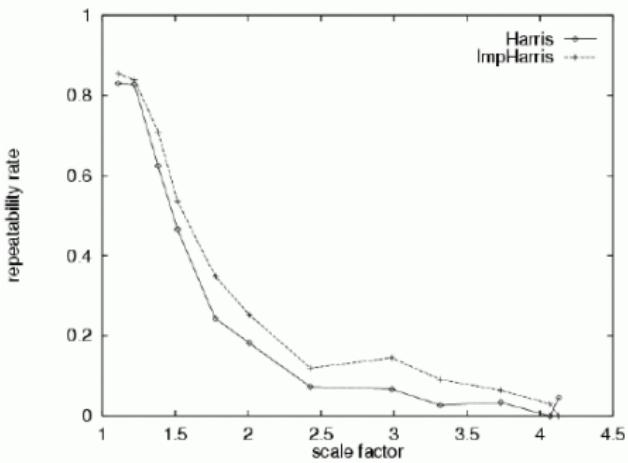
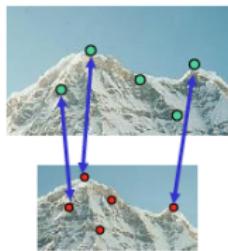


Qualité de détection

Qualité du détecteur de Harris pour différents changements d'échelle .

Repeatability rate:

$$\frac{\# \text{ correspondences}}{\# \text{ possible correspondences}}$$



C. Schmid et al : *Evaluation of Interest Point Detectors*, IJVC 2000

(Parenthèse) Detecteur de Harris Couleur

Principe

- Généralisation de la fonction d'auto-corrélation à la couleur.
- Utilisation du tenseur multispectral :

$$M = G(\sigma) \begin{pmatrix} R_x^2 + V_x^2 + B_x^2 & R_x R_y + V_x V_y + B_x B_y \\ R_x R_y + V_x V_y + B_x B_y & R_y^2 + V_y^2 + B_y^2 \end{pmatrix}$$

Detecteur de Harris Couleur

Exemple



Plan

1 Motivations

2 Points d'intérêt

3 Les coins

- Moravec
- Harris
- Hessian

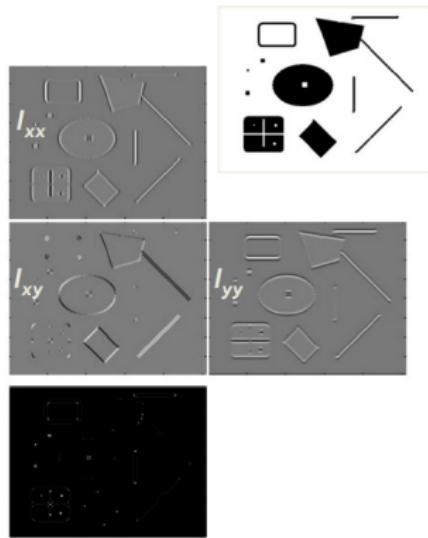
4 Invariance à l'échelle

Détecteur Hessian

- Déterminant hessien :

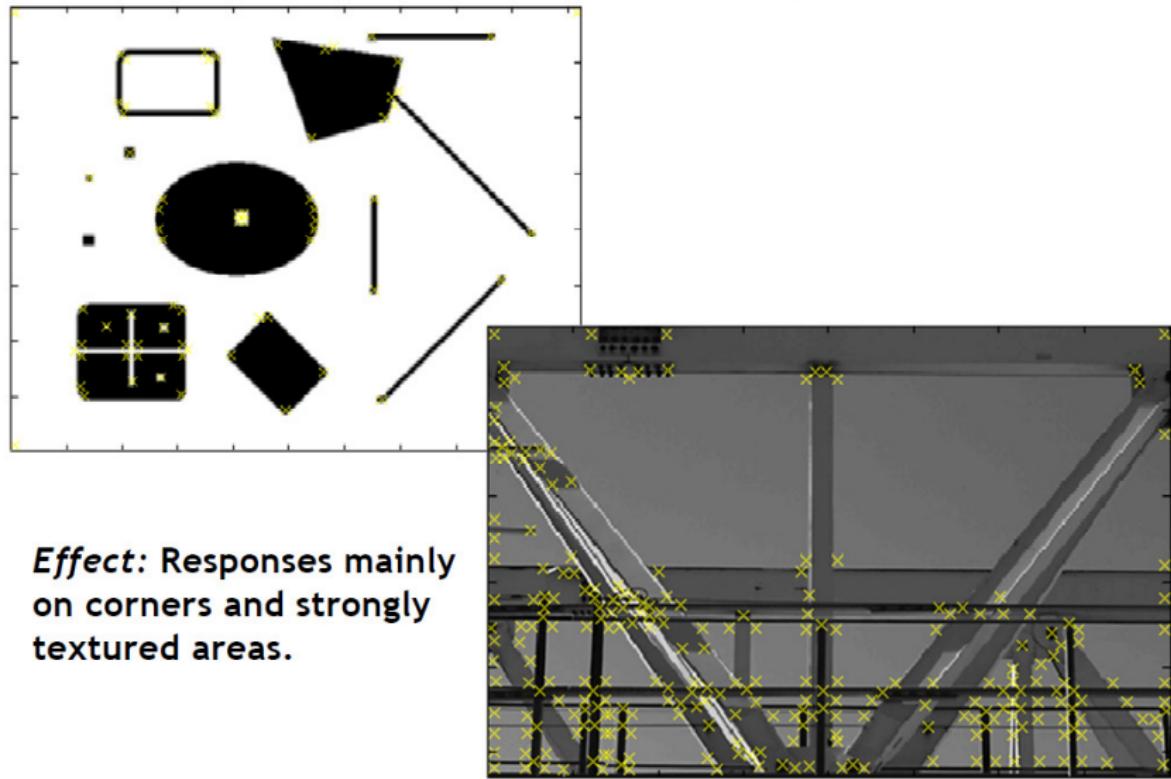
$$\text{Hessian}(I) = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}$$

$$\bullet \det(\text{Hessian}(I)) = I_{xx}I_{yy} - I_{xy}^2$$



Recherche des dérivées fortes dans deux directions perpendiculaires.

Detecteur Hessian : exemple



Effect: Responses mainly on corners and strongly textured areas.

Detecteur Hessian : exemple



Plan

1 Motivations

2 Points d'intérêt

3 Les coins

- Moravec
- Harris
- Hessian

4 Invariance à l'échelle

Des points aux régions

- Les détecteurs de coins permettent de détecter des points d'intérêts avec :
 - ▶ une localisation précise,
 - ▶ une forte répétabilité.
- Pour comparer ces points, il est nécessaire de définir un descripteur sur une région.
- Question : Comment détecter et définir des régions d'intérêts avec invariance à l'échelle ?
- autrement dit : quelle taille de région ?

Introduction à la problématique de l'échelle de détection

- On considère des régions de différentes tailles autour d'un point, par exemple des cercles.
- Les régions de taille correspondante vont apparaître similaires dans les deux images.

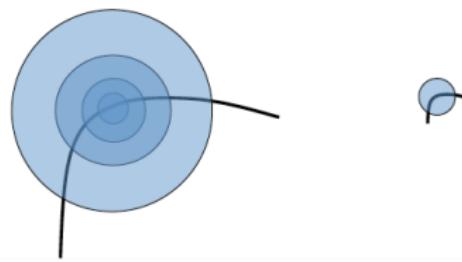


FIGURE – Source : Fei Fei Li

Introduction à la problématique de l'échelle de détection

Problème

Comment choisir les régions correspondantes de manière indépendante dans chaque image ?

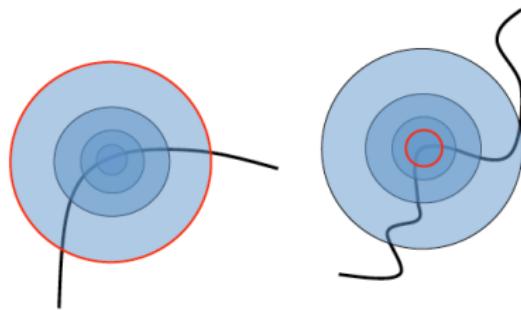
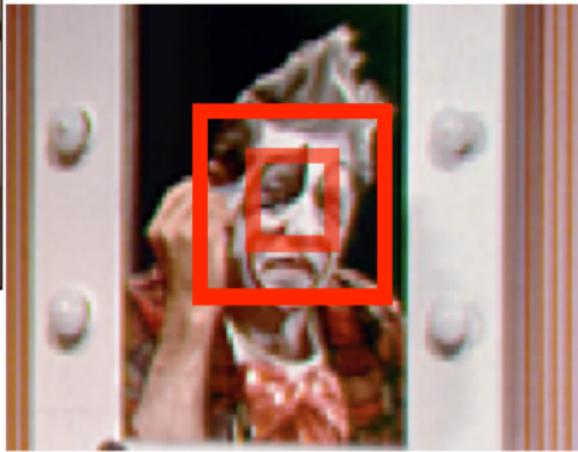


FIGURE – Source : Fei Fei Li

Introduction à la problématique de l'échelle de détection



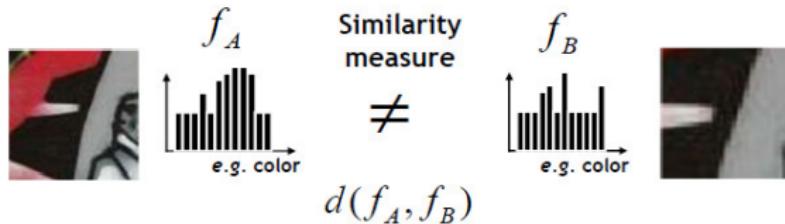
Introduction à la problématique de l'échelle de détection



Approche naïve : recherche exhaustive

Principe

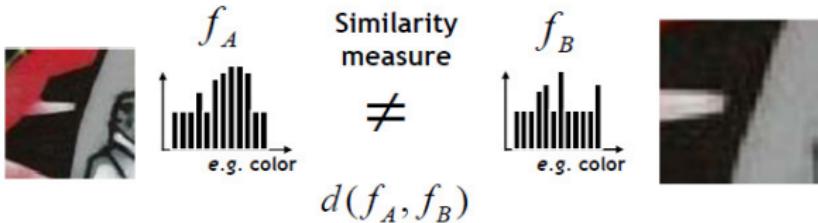
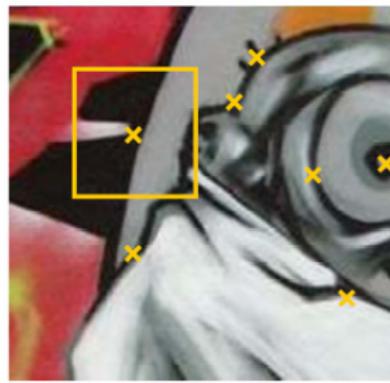
Comparaison des descripteurs en faisant varier la taille du patch.



Approche naïve : recherche exhaustive

Principe

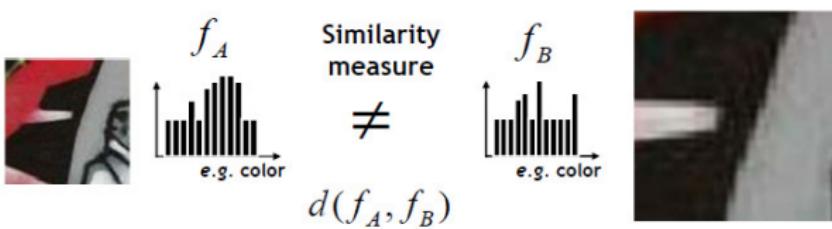
Comparaison des descripteurs en faisant varier la taille du patch.



Approche naïve : recherche exhaustive

Principe

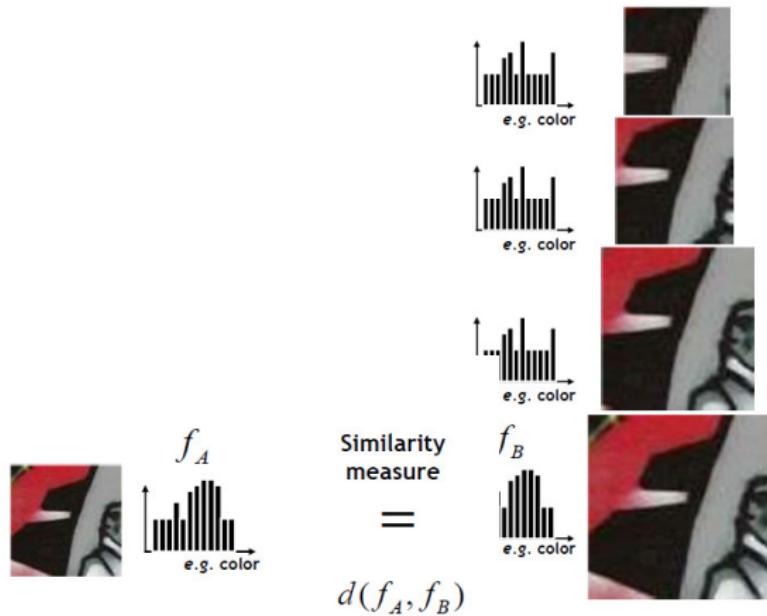
Comparaison des descripteurs en faisant varier la taille du patch.



Approche naïve : recherche exhaustive

Principe

Comparaison des descripteurs en faisant varier la taille du patch.



Approche naïve : recherche exhaustive

Bilan

- Inefficace en terme de calcul.
- Possible pour la mise en correspondance.
- Ne permet pas de traiter de grandes bases d'images.
- Pas possible pour la reconnaissance.

Sélection automatique de l'échelle

Idée

- Concevoir une fonction sur les régions qui est invariante à l'échelle (*la même pour les régions correspondantes, même si elles sont à différentes échelles*).
- Exemple : intensité moyenne. Pour des régions correspondantes, même de différentes tailles, elle sera la même.
- Pour un point dans l'image, le considérer comme une fonction de la taille de la région.

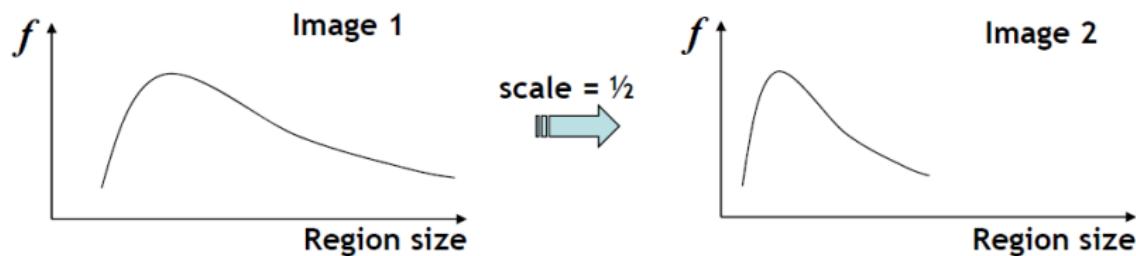


FIGURE – Source : Kristen Grauman

Sélection automatique de l'échelle

Approche classique

- Prendre un maximum local de cette fonction.
- Observation : la taille de la région pour laquelle le maximum est trouvé devrait être invariante à l'échelle.
- Important : cette taille doit être trouvée dans chaque image de manière indépendante.

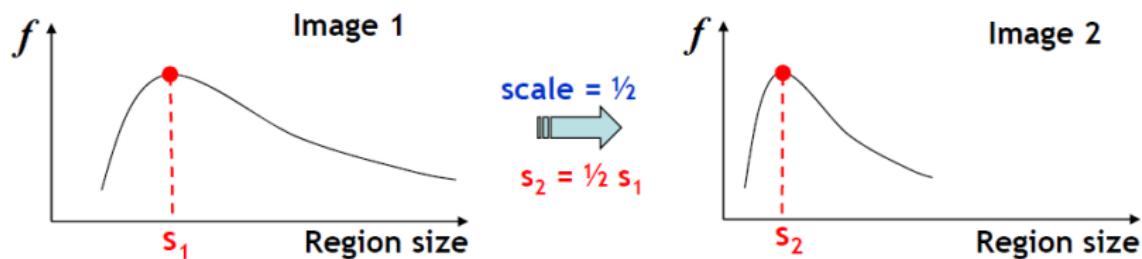
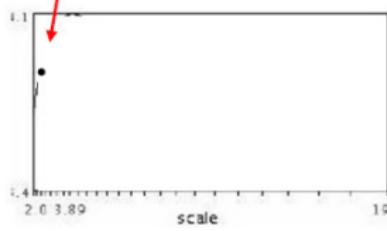


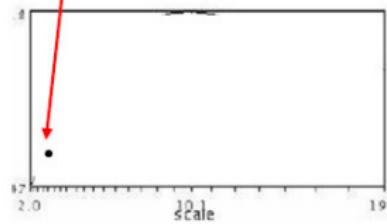
FIGURE – Source : Kristen Grauman

Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.



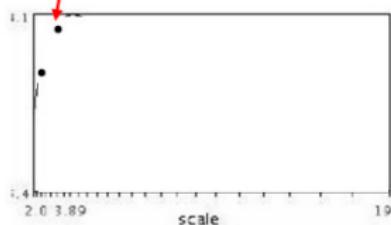
$$f(I_{i_1 \dots i_m}(x, \sigma))$$



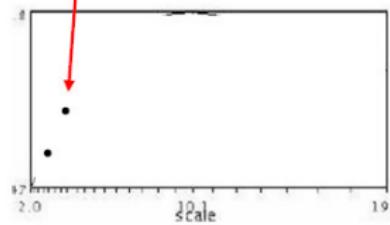
$$f(I_{i_1 \dots i_m}(x', \sigma))$$

Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.



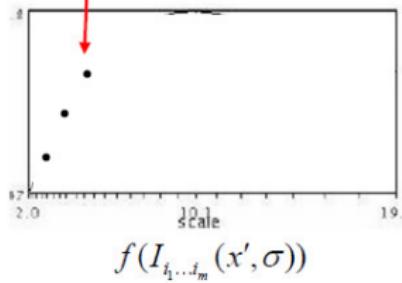
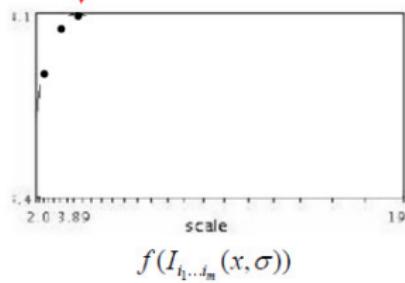
$$f(I_{i_1\dots i_m}(x, \sigma))$$



$$f(I_{i_1\dots i_m}(x', \sigma))$$

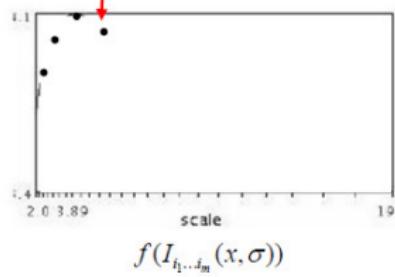
Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.

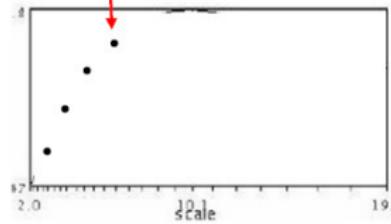


Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.



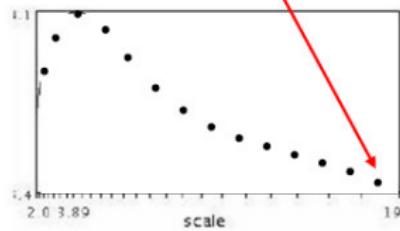
$$f(I_{i_1\dots i_m}(x, \sigma))$$



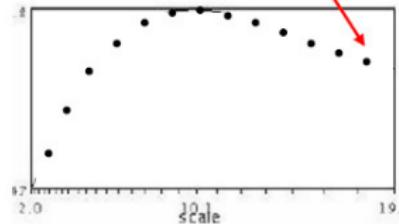
$$f(I_{i_1\dots i_m}(x', \sigma))$$

Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.



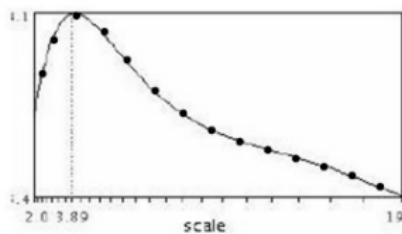
$$f(I_{h...i_m}(x, \sigma))$$



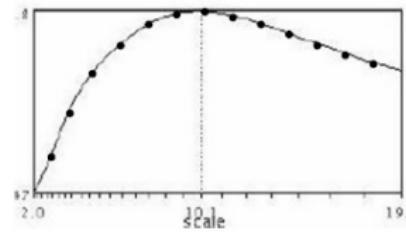
$$f(I_{i...i_m}(x', \sigma))$$

Sélection automatique de l'échelle

Valeurs de la fonction pour des échelles croissantes.



$$f(I_{i_1...i_m}(x, \sigma))$$



$$f(I_{i_1...i_m}(x', \sigma'))$$

Sélection automatique de l'échelle

Normalisation : rééchelonnage à une taille fixée.

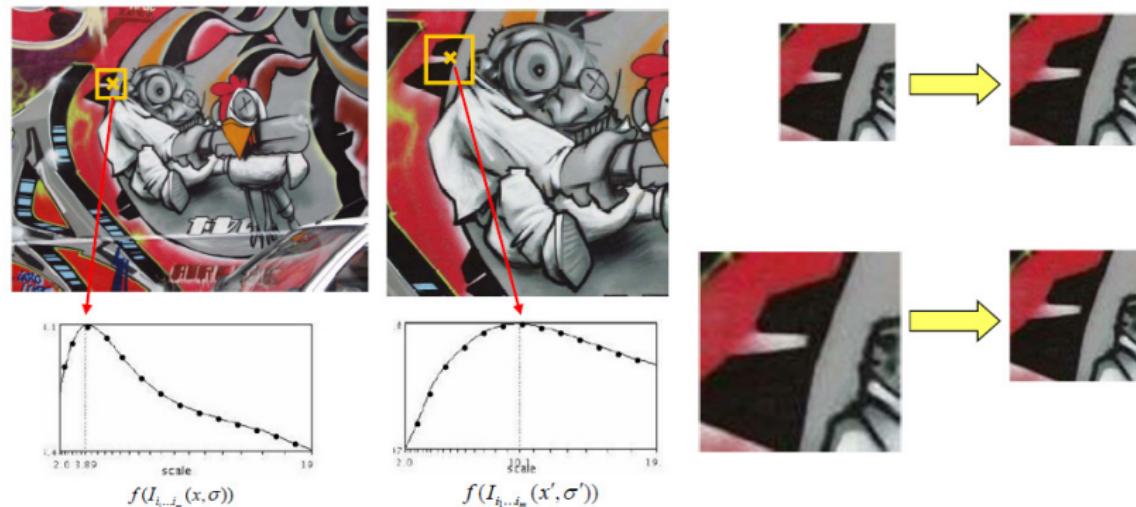


FIGURE – Source : Tinne Tuyttellaars

Quelle fonction choisir ?

Une bonne fonction pour la détection de l'échelle caractéristique doit avoir un pic fin et stable

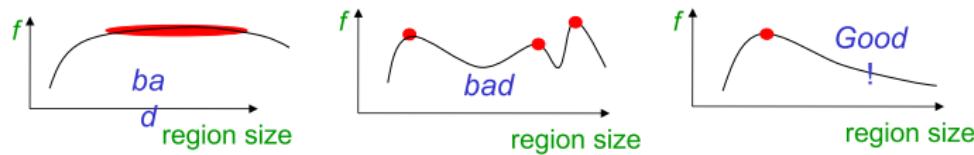


FIGURE – Source : Darya Frolova

Pour des images classiques, une *bonne* fonction est une fonction qui a une bonne réponse au contraste.

Quelle fonction choisir ?

Il faut une méthode pour sélectionner des points de manière stable en position et en échelle

- Le seul noyau raisonnable est le noyau gaussien [Koenderink,84], [Lindeberg, 94] .
- Détection de pics dans la pyramide des différences de gaussiennes [Burt, 83], [Crowley,84] .

Quelle fonction choisir ?

Laplacien de gaussienne (LoG) : détecteur de blobs.

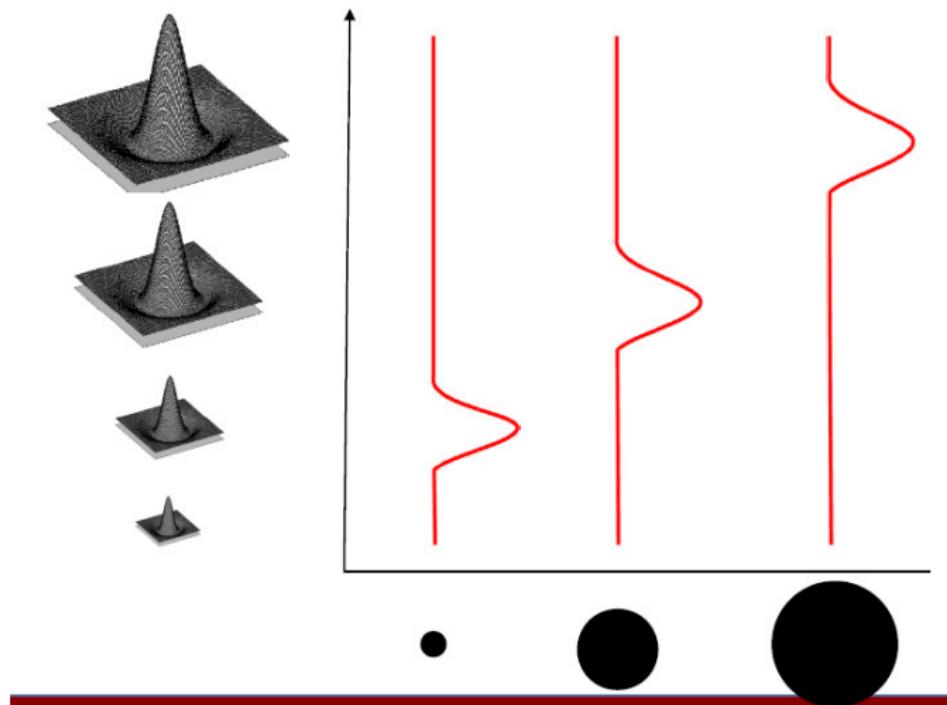


FIGURE – Source : Fei Fei Li

Echelle caractéristique

L'échelle caractéristique est définie comme l'échelle qui produit un pic dans la réponse du Laplacien.

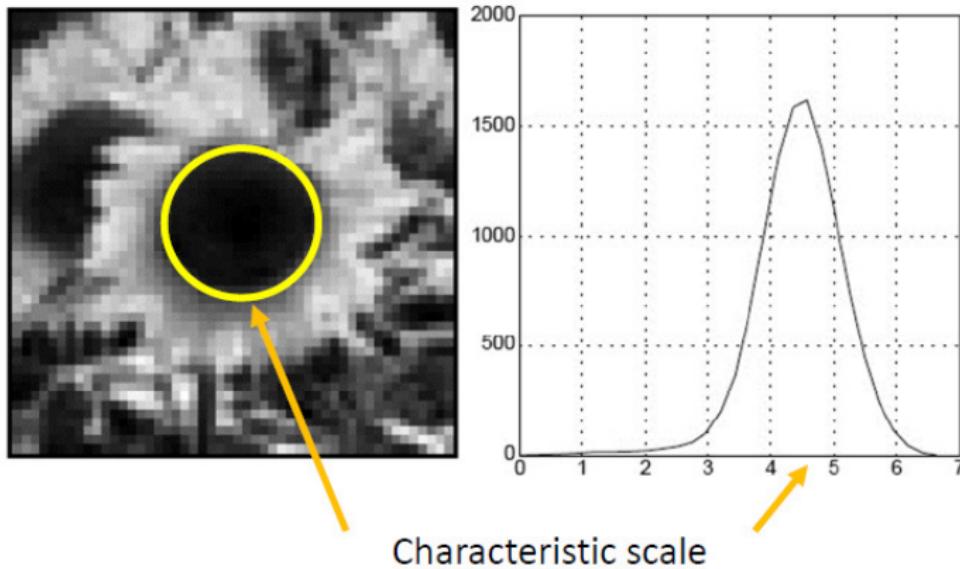


FIGURE – Source : Fei Fei Li

Détection invariante à l'échelle

Fonction pour déterminer l'échelle : $f = \text{kernel} * \text{Image}$

- Noyaux :

- ▶ Laplacien :

$$L = \sigma^2(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$

- ▶ Différence de gaussiennes :

$$L = G(x, y, k\sigma) - G(x, y, \sigma)$$

Avec $G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$

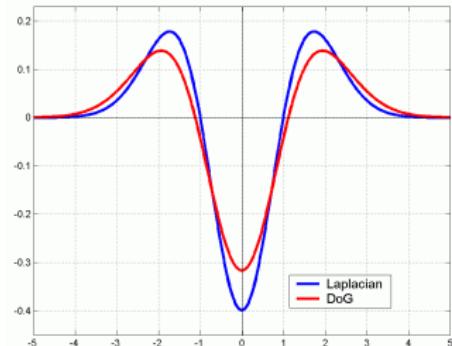
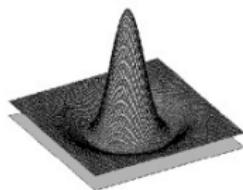


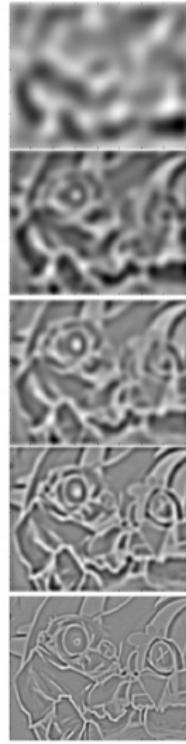
FIGURE – Source : Darya Frolova

Laplacien de Gaussiennes

- Interest points:
 - Local maxima in scale space of Laplacian-of-Gaussian



$$L_{xx}(\sigma) + L_{yy}(\sigma)$$

 σ
 σ^2
 σ^3
 σ^4
 σ^5


Laplacien de Gaussiennes

- Interest points:
 - Local maxima in scale space of Laplacian-of-Gaussian

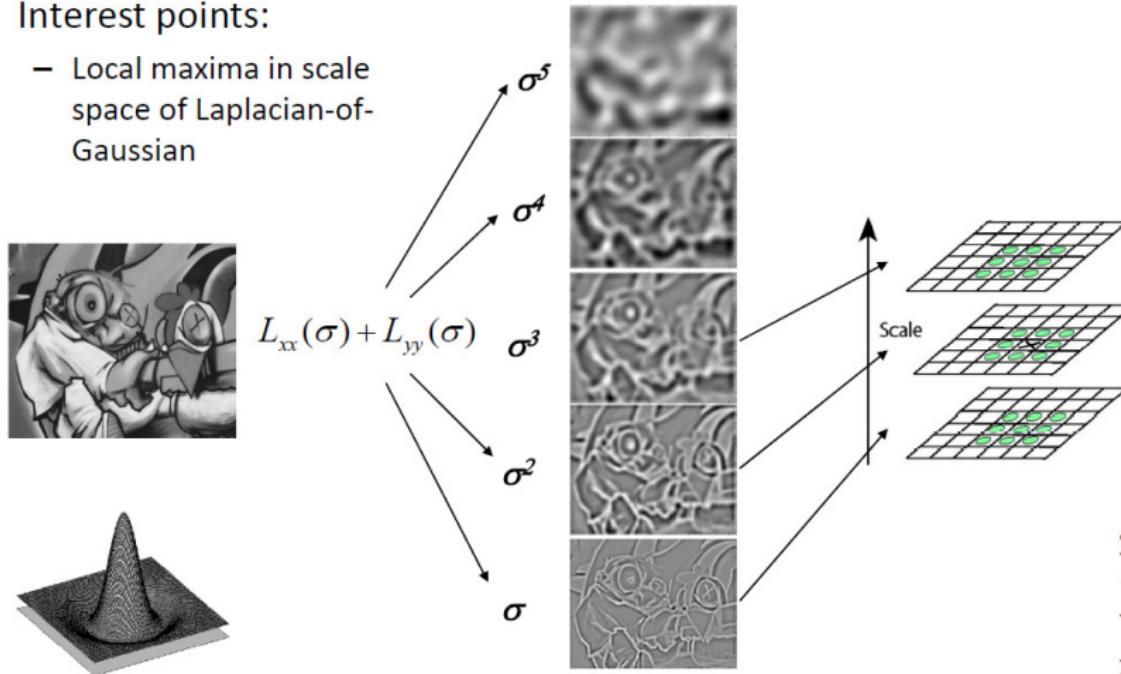


FIGURE – Source : Krystian Mikolajczyk

Laplacien de Gaussiennes

- Interest points:
 - Local maxima in scale space of Laplacian-of-Gaussian

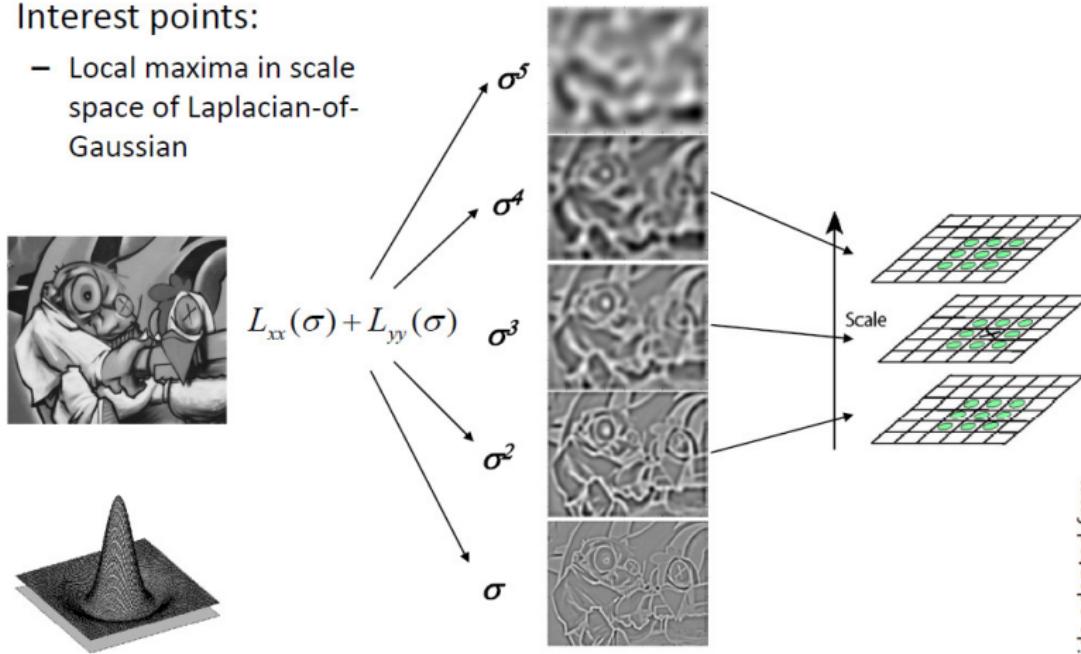


FIGURE – Source : Krystian Mikolajczyk

Laplacien de Gaussiennes

- Interest points:
 - Local maxima in scale space of Laplacian-of-Gaussian

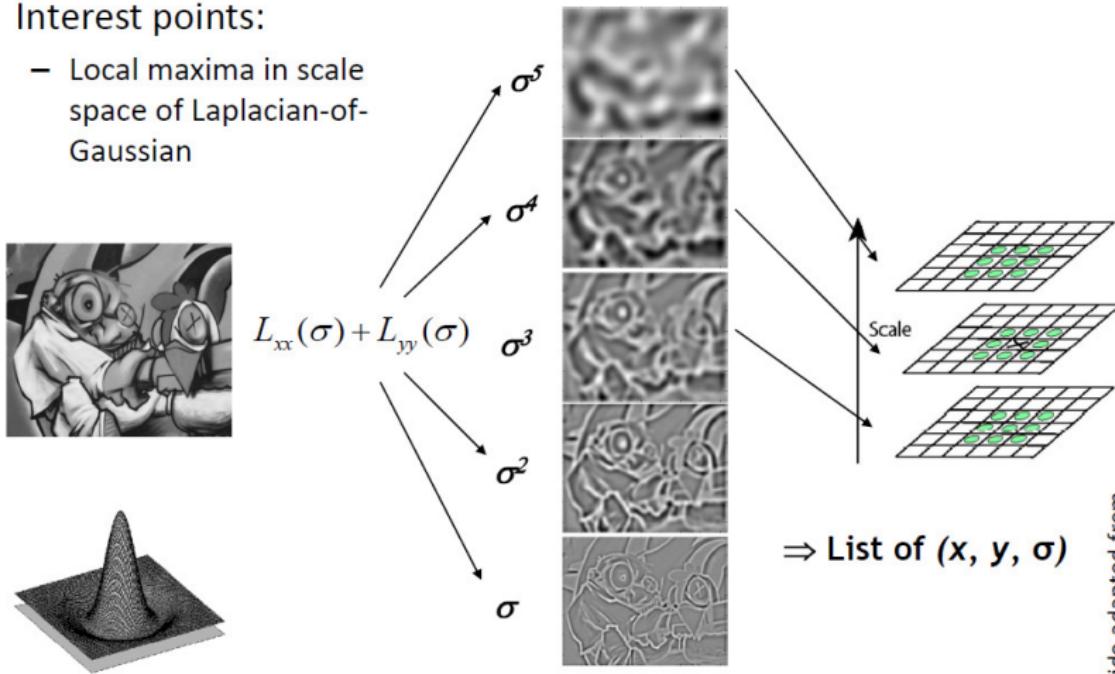


FIGURE – Source : Krystian Mikolajczyk

Détecteur LoG : fonctionnement



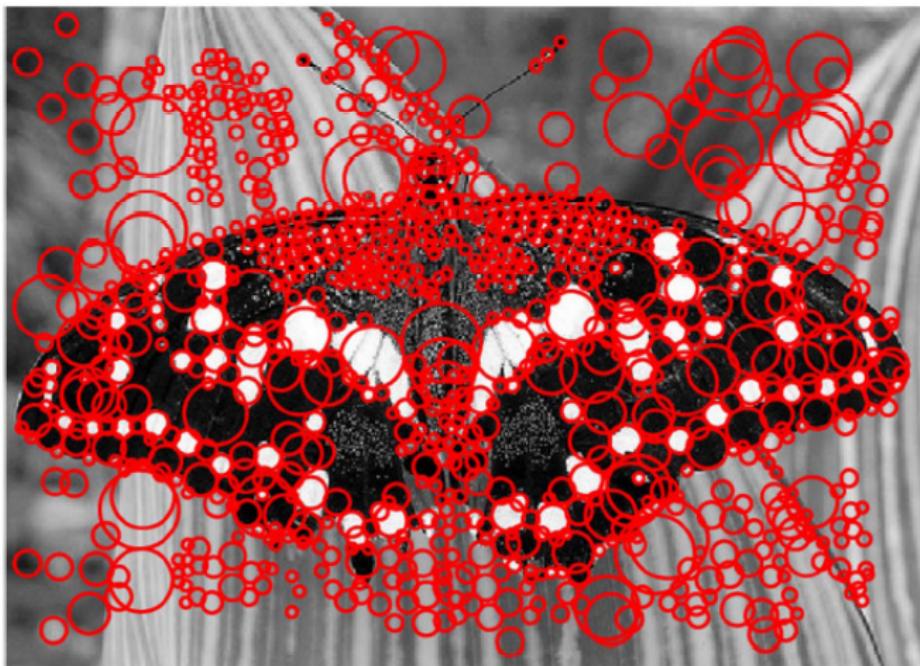
FIGURE – Source : Svetlana Lazebnik

Détecteur LoG : fonctionnement



$\sigma = 11.9912$

Détecteur LoG : fonctionnement



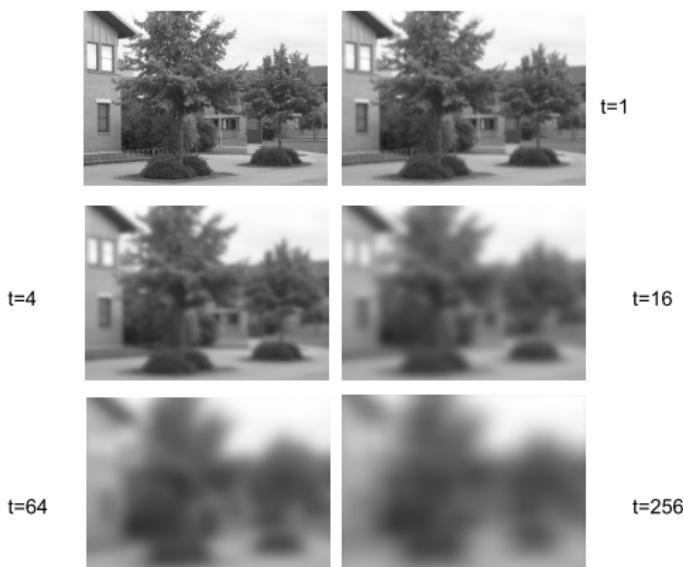
Détecteur LoG : Représentation multi-échelle

- L'image du Laplacien de Gaussienne permet de définir une nouvelle **représentation multi-échelle** de l'image.
- Chaque niveau permet d'observer les objets à un niveau différent de détails.
- Cette représentation permet de détecter :
 - ▶ Les **contours** : passage par zero
 - ▶ Les **blobs** : régions délimitées par les passages par zéros.
 - ▶ Les **crêtes** (ridges and valley) : lignes d'extrema locaux.
 - ★ Séquence de points plus hauts (bas) que ceux des deux côtés.
Intuitivement une ligne qui se trouve au milieu d'une tâche allongée.
 - ▶ Les **pics** (peak) : points d'extrema locaux. Sommet (ou creux) dans toutes les directions.

Représentation multi-échelle

Exemple de représentation d'une image dans l'espace d'echelle gaussien.

$$g(x, y, t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t}$$
$$L(., ., t) = g(., ., t) * f(., .)$$



http://en.wikipedia.org/wiki/Scale_space

Détecteur LoG : Représentation multi-échelle

Exemple

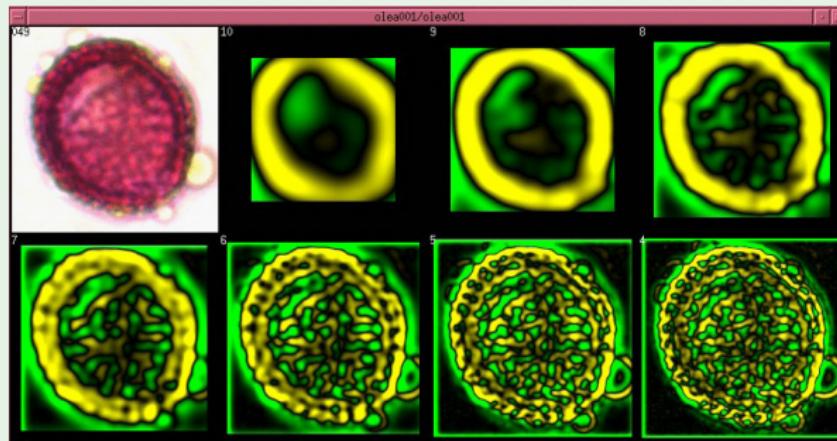


FIGURE – Source : Alain Boucher

Détecteur LoG : Représentation multi-échelle

Exemple de détection de crêtes

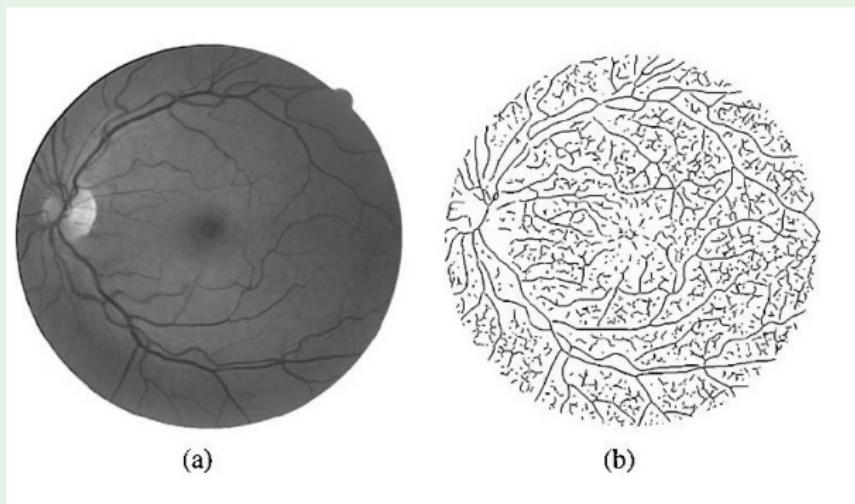


FIGURE – Source : Alain Boucher

Détecteur LoG : Représentation multi-échelle

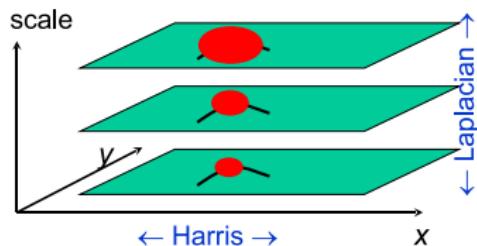
Exemple de détection de crêtes

De nombreuses applications pour la détection de routes dans des images aériennes ou la détection de vaisseaux sanguins dans des images de la rétine.



Detecteur de Harris Laplacien

Objectif : invariance à l'échelle.



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

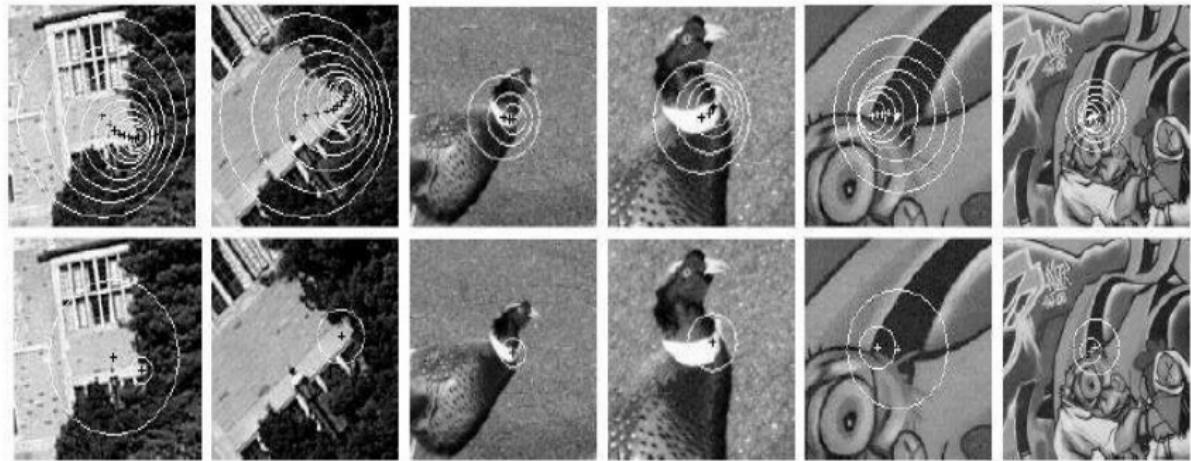
Principe

Trouver le maximum local pour :

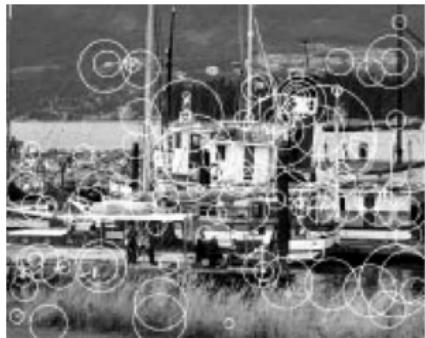
- Les coins de Harris dans l'espace.
- Le laplacien en échelle.

Detecteur de Harris Laplacien : Exemple

1. Initialization: Multiscale Harris corner detection
2. Scale selection based on Laplacian



Detecteur de Harris Laplacien : Exemple



(a) Scale change of 3.9 and rotation of 17° .



(b) Scale change of 1.8 and viewpoint change of 30°

SIFT : Scale Invariant Feature Transform

Caractéristiques

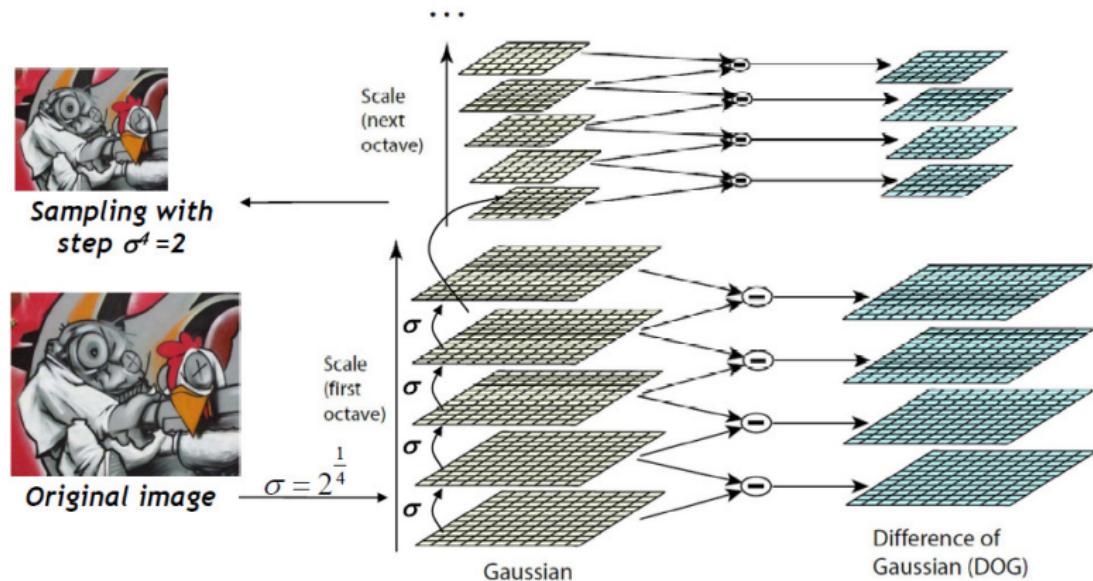
- Nouvelle approche de points d'intérêts très populaire.
 - ▶ Invariance en échelle.
 - ▶ Invariance par rotation.
 - ▶ Invariance au point de vue.
 - ▶ Invariance aux conditions d'éclairage.
- Proposé par D. Lowe en 1999 et 2004.
<http://www.cs.ubc.ca/~lowe/keypoints/>
<http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>
- Pas juste un détecteur mais aussi un descripteur.

SIFT : Scale Invariant Feature Transform

Etape 1 : Traitement multi-échelle de l'image :

- Pyramide de l'image à différentes tailles.
- Différences de gaussienne :

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$



SIFT : Scale Invariant Feature Transform

Etape 1 : Traitement multi-échelle de l'image :

- L'espace d'échelles est séparé en octaves :
 - ▶ Octave 1 : échelle σ
 - ▶ Octave 2 : échelle 2σ
 - ▶ ...
- Dans chaque octave, chaque image initiale est convoluée avec des noyaux gaussiens pour avoir une représentation multi-échelle.
- Différence de gaussiennes
- Après chaque octave, l'image est réduite d'une facteur 2 : image dont la taille est $\frac{1}{4}$ de l'image précédente.

SIFT : Scale Invariant Feature Transform

Etape 1 : Traitement multi-échelle de l'image : construction de la pyramide

$$k^4 \sigma_0$$



$$k^3 \sigma_0$$



$$k^2 \sigma_0$$



$$k \sigma_0$$



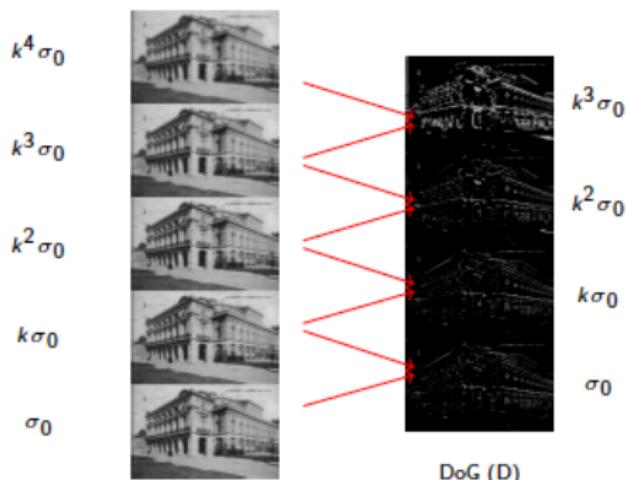
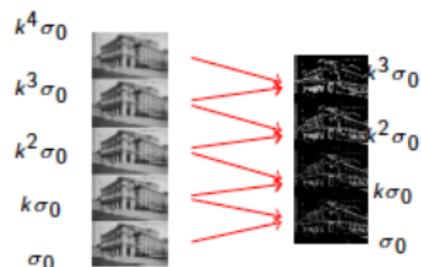
$$\sigma_0$$



$$\sigma_0 = 1.6, k = \sqrt{2}, 3 \text{ octaves}, 5 \text{ échelles}$$

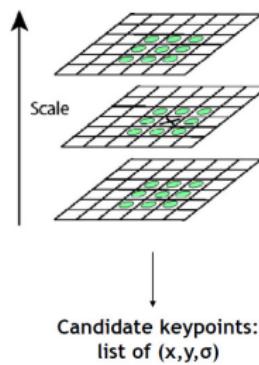
SIFT : Scale Invariant Feature Transform

Etape 1 : Traitement multi-échelle de l'image : construction de la pyramide



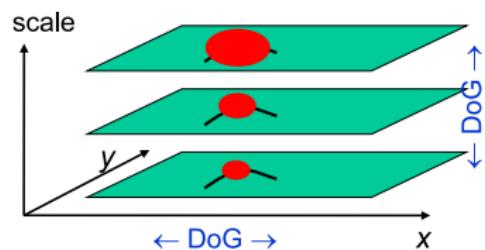
SIFT : Scale Invariant Feature Transform

- Etape 2 : Extraction des extrema locaux des différences de gaussiennes.
 - ▶ Un point = (x, y, σ) .



- Etape 3 : Filtrage des points obtenus
 - ▶ Rejet des points à faible contraste.
 - ▶ Rejet des points sur les bords (courbures faibles).

SIFT : Scale Invariant Feature Transform



Détection de points invariante à l'échelle

- **Données** : Deux images de la même scène avec des grandes différences d'échelles.
- **But** : Trouver les mêmes points d'intérêts indépendamment de l'échelle.
- **Solution** : Trouver les maxima de fonctions adaptées en échelle et en image.

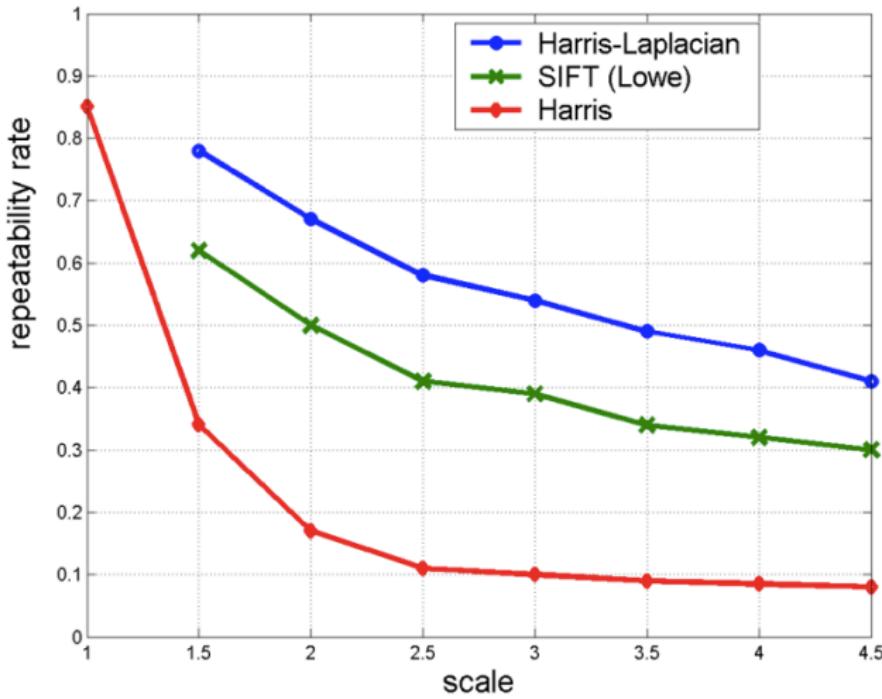
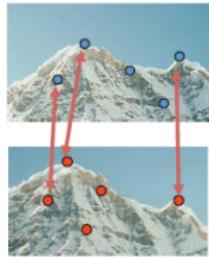
Méthodes

- **Harris Laplacian** : maximiser le laplacien sur l'échelle et la mesure des coins de Harris sur l'image [Mikolajczyk, Schmid]
- **SIFT** : maximiser les différences de gaussienne en échelle et en espace [Lowe]

Détection de points invariante à l'échelle

Repeatability rate:

$$\frac{\# \text{ correspondences}}{\# \text{ possible correspondences}}$$



Résumé

- Importance d'avoir des caractéristiques invariantes en vision par ordinateur.
- Localisation des points d'intérêts :
 - ▶ Détecteur de Moravec
 - ▶ Détecteur de Harris
 - ▶ Détecteur Hessien
- Sélection d'une région invariante à l'échelle :
 - ▶ Sélection automatique de l'échelle
 - ▶ Laplacien de gaussienne
 - ▶ Différence de gaussienne
 - ▶ Combinaisons : Harris-Laplace, SIFT.

Quels descripteurs ? : SIFT, SURF, HOG, GLOH, BOW, ...