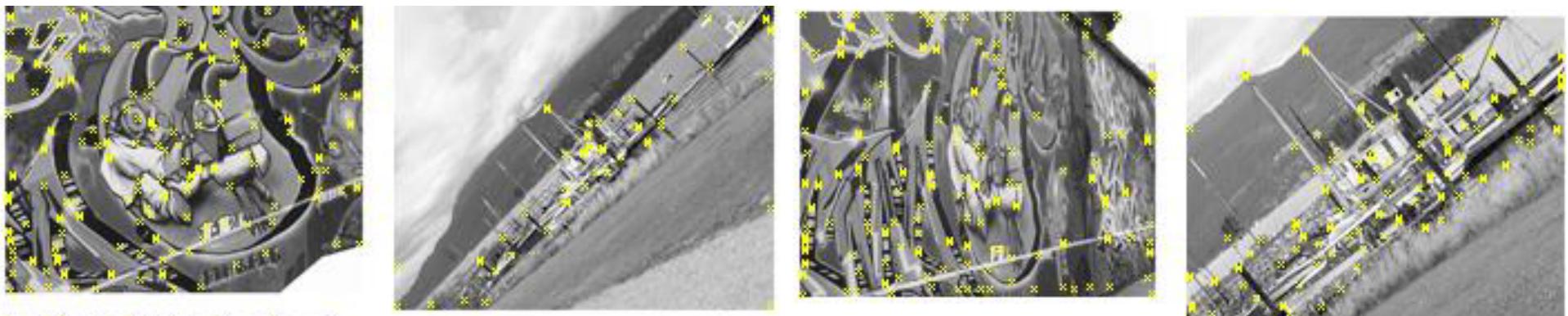


Slides from A. Manzanera

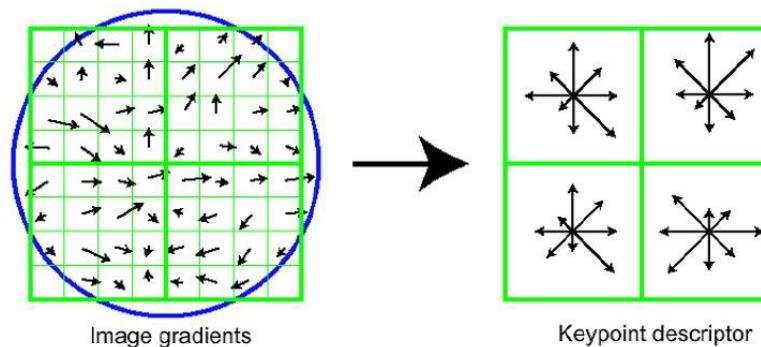
Resp.: Matthieu Cord —LIP6 / UPMC Paris 6

Course objective: local features

1. DETECTION: Points/Regions of Interest (always looking for repeatability)



2. DESCRIPTION: Local features (always looking for invariance)



QU'EST-CE QU'UNE BONNE PRIMITIVE VISUELLE ?

Objectif : mettre en correspondance des points / ensembles / images avec d'autres points / ensembles / classes / catégories visuelles.

Les propriétés souhaitées d'une bonne primitive sont :

- **Robustesse** : L'information visuelle doit être fidèlement représentées indépendamment des modifications qu'elle peut subir d'une instance à l'autre : distorsions géométriques, changements d'illumination, occultations, variations intra-classes...
- **Discrimination** : L'objet représenté doit se distinguer facilement des autres objets, en particulier de ceux qui l'environnent.
- **Efficacité** : le calcul de la primitive doit être rapide, et les descripteurs économies en taille mémoire...

Une primitive visuelle peut caractériser l'information image à plusieurs niveaux :

- **Local** : une primitive par point / région / courbe...
- **Global** : une primitive qui concerne l'ensemble de l'information...

BASES DE GEOMETRIE DIFFERENTIELLE POUR LES IMAGES

La « géométrie locale » dans une image se décrit naturellement par des concepts de la géométrie différentielle : direction, courbure,...

Dans le modèle différentiel, l'image est assimilée à une fonction $I: \mathbb{R}^2 \rightarrow \mathbb{R}$ continue et différentiable.

Le comportement local de l'image autour de chaque point peut être prédit par ses dérivées locales (Formule de Taylor) :

$$I(x_0 + \varepsilon, y_0 + \eta) = \sum_{k=0}^r \sum_{i=0}^k C_i^k \varepsilon^{k-i} \eta^i \frac{\partial^k I}{\partial x^{k-i} \partial y^i}(x_0, y_0) + o((\varepsilon^2 + \eta^2)^{r/2})$$

Dans les images discrètes, la notion de dérivable sera remplacée par la notion de régularité locale.

Cette régularité pouvant être imposée explicitement par filtrage (convolution), l'estimation de la dérivée correspondra à une convolution, et sera toujours relative à une échelle (espaces d'échelles).

ORDRE 1 : GRADIENT ET ISOPHOTE

A l'ordre 1, la grandeur de base est le vecteur gradient :

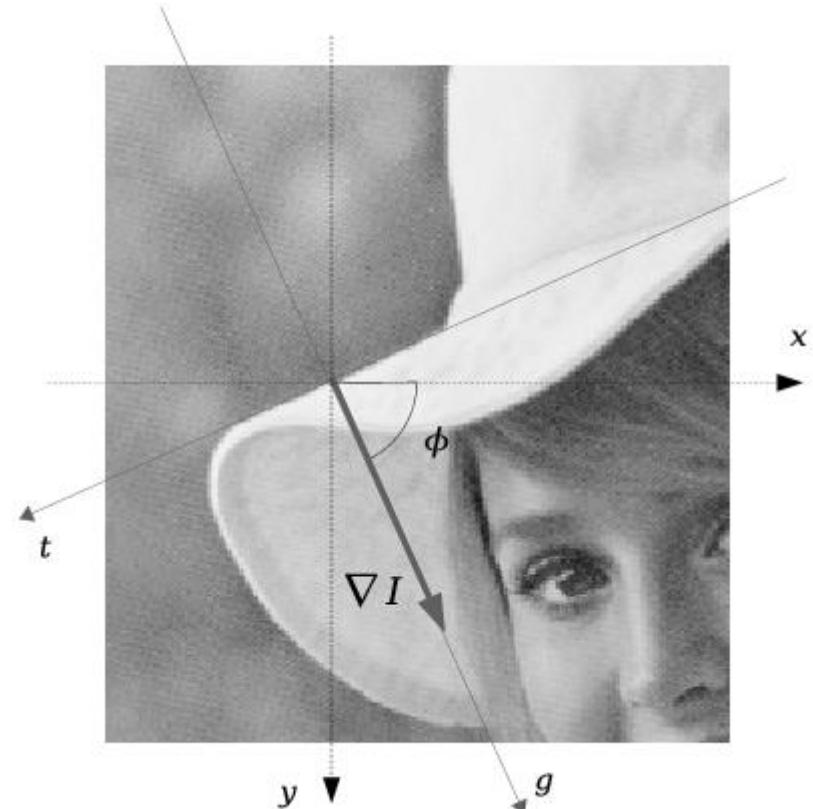
$$\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)^T$$

- Son argument, $\arg \nabla I$, correspond à la direction de plus grande pente.
- Sa norme, $\|\nabla I\|$, mesure le contraste local.
- Il permet de calculer la dérivée partielle dans toute direction du plan. Soit v un vecteur unitaire :

$$\frac{\partial I}{\partial v} = \nabla I \cdot v^T$$

- Dans le repère local (g, t) avec $g = \frac{\nabla I}{\|\nabla I\|}$ et $t = g^\perp$:

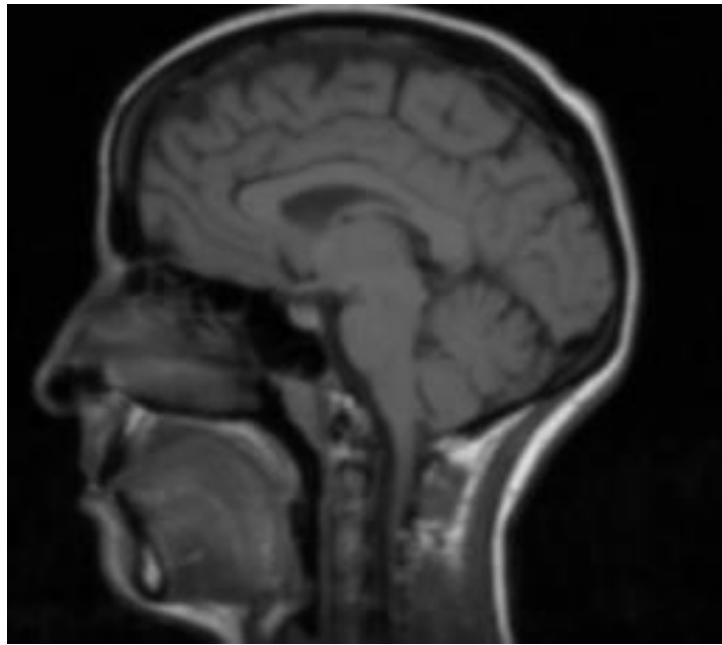
$$\frac{\nabla I}{\nabla g} = \|\nabla I\| \text{ (direction principale)} ; \frac{\nabla I}{\nabla t} = 0 \text{ (isophote)}$$



GRANDEURS DIFFÉRENTIELLES D'ORDRE 1

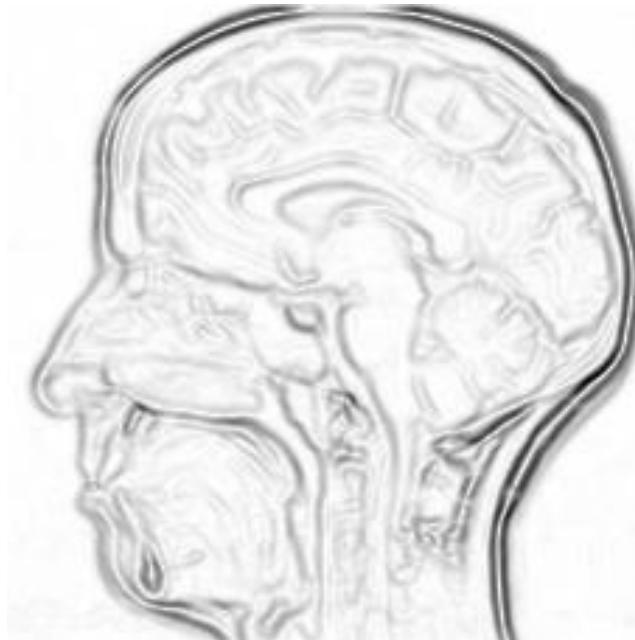
original

I

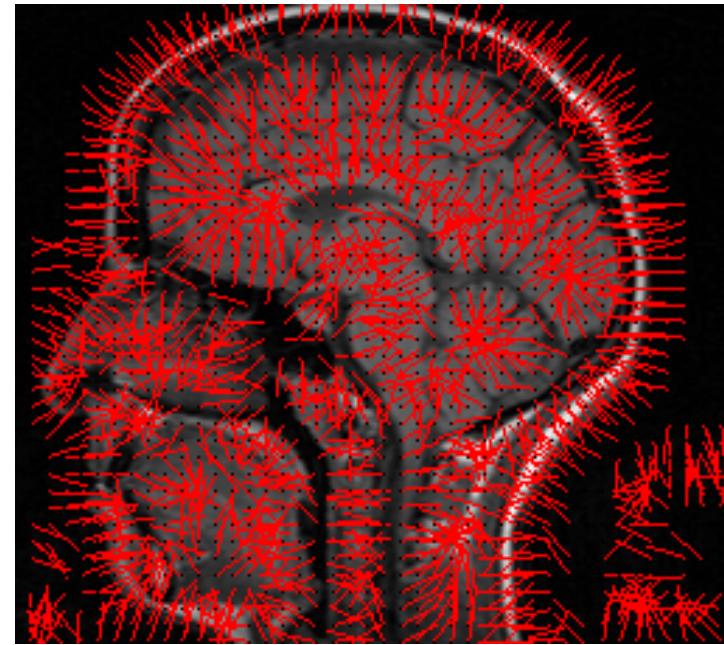


norme du
gradient

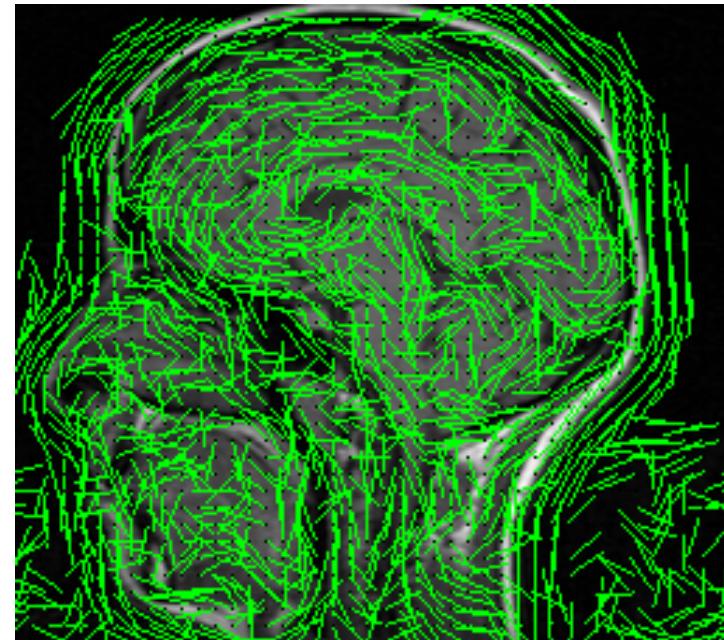
$\|\nabla I\|$



direction
du gradient
 $\arg \nabla I$



direction de
l'isophote
 $\arg \nabla I^\perp$

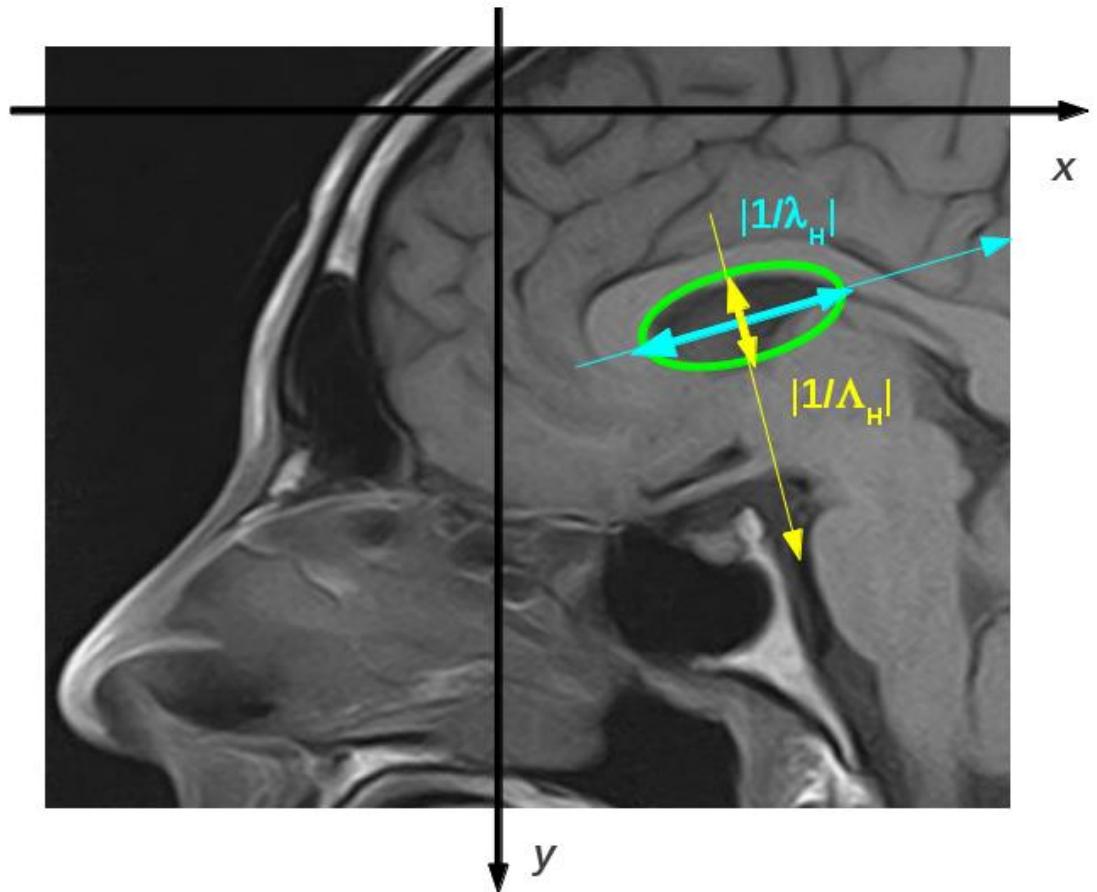


ORDRE 2 : HESSIENNE ET COURBURE

A l'ordre 2, la grandeur de base est la matrice hessienne :

- Ses vecteurs propres (resp. ses valeurs propres Λ_H et λ_H) correspondent aux directions (resp. intensités) de courbure principale.
- Sa norme de Frobénius, $\|H_I\|_F$, mesure l'intensité de la courbure globale.

$$H_I = \begin{pmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{pmatrix}$$



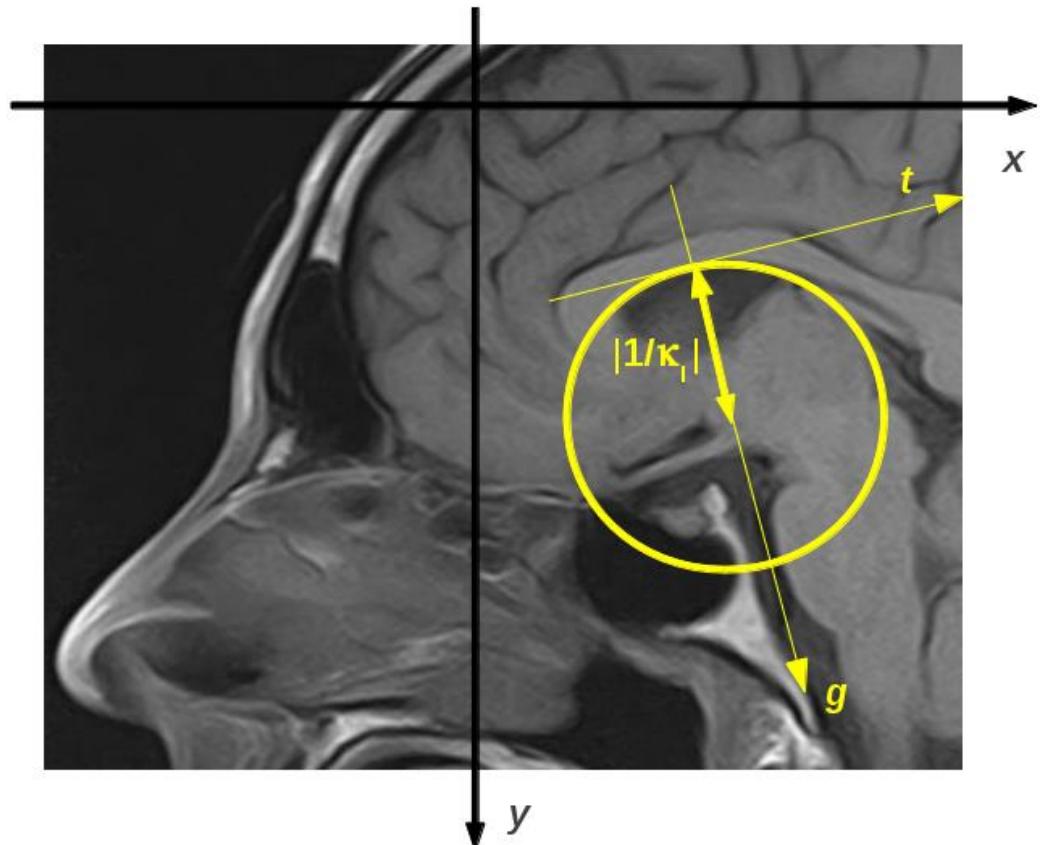
ORDRE 2 : HESSIENNE ET COURBURE

- Soit u et v deux vecteurs unitaires. La dérivée seconde selon u et v se calcule par :
- En particulier la courbure de l'isophote est égal à l'inverse du rayon du cercle osculateur au contour :

$$\kappa_I = I_{tt} = \frac{I_{xx}I_y^2 - 2I_xI_yI_{xy} + I_{yy}I_x^2}{\|\nabla I\|^3}$$

(Notations :

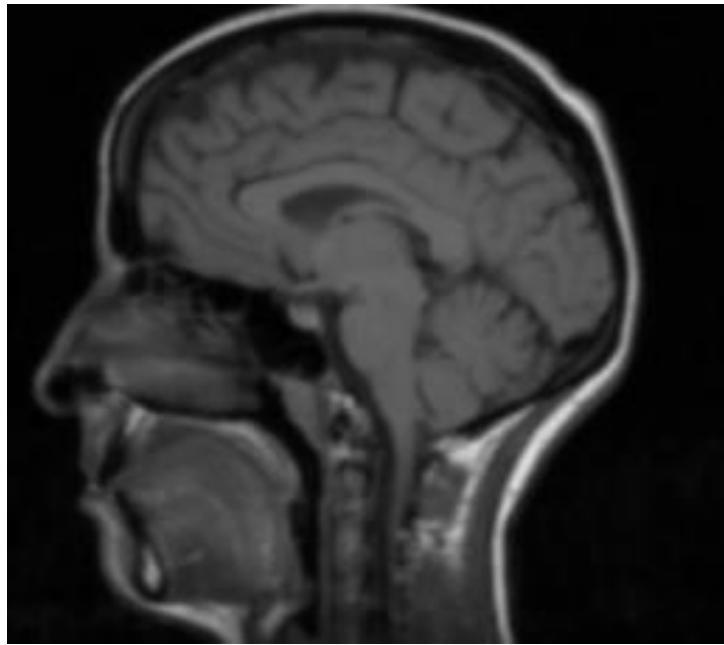
$$I_u = \frac{\partial I}{\partial u}; I_{uv} = \frac{\partial^2 I}{\partial u \partial v}, \text{ etc.})$$



GRANDEURS DIFFÉRENTIELLES D'ORDRE 2

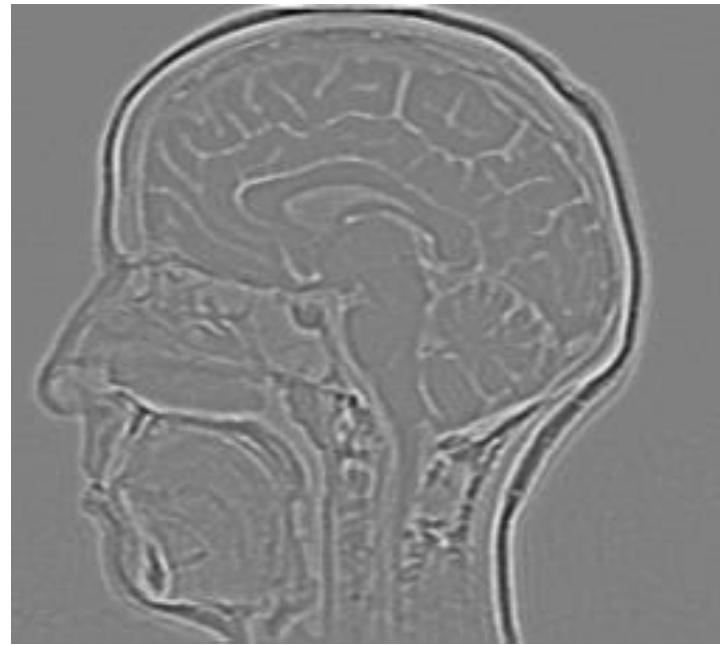
original

I



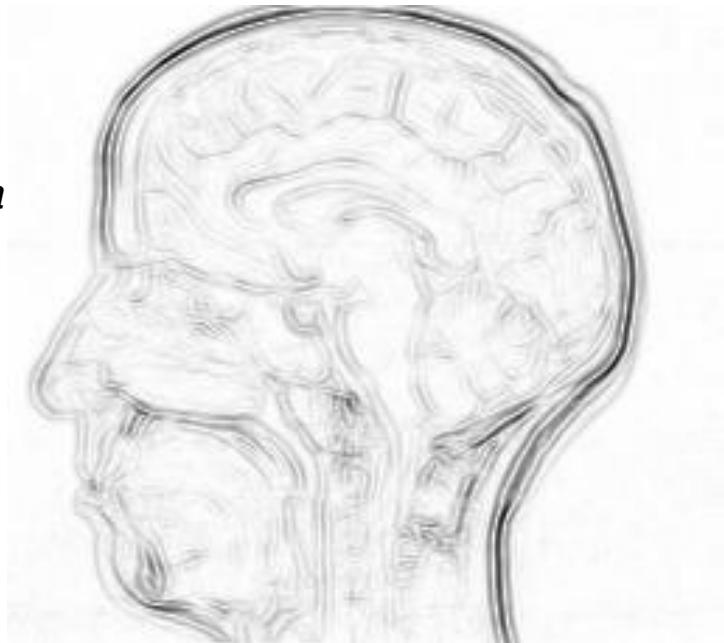
*trace de la
hessienne, ou
courbure totale
= laplacien*

ΔI



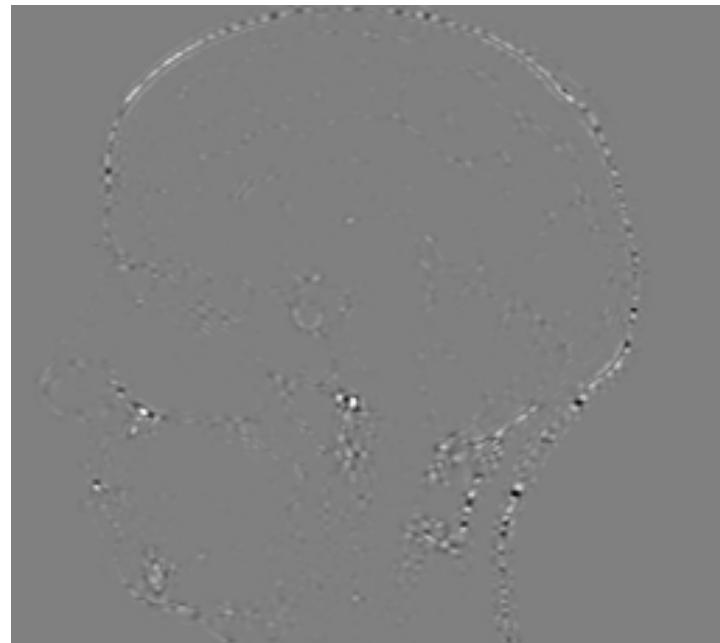
*norme de la
hessienne*

$\|H_I\|_F$



*déterminant
de la
hessienne*

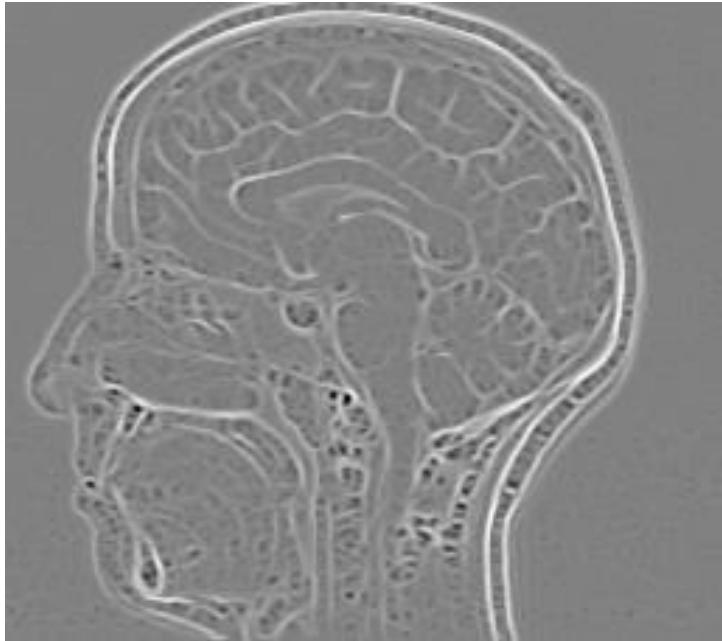
$\det\|H_I\|_F$



GRANDEURS DIFFÉRENTIELLES D'ORDRE 2

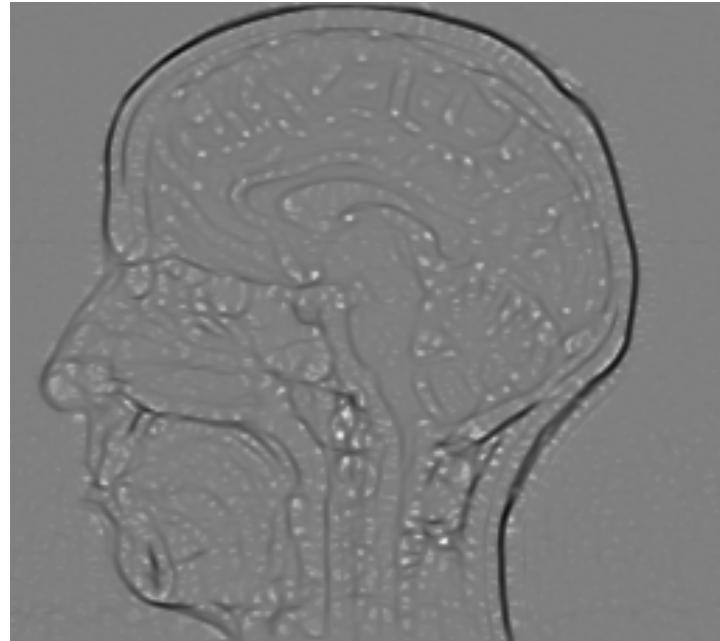
« grande »
valeur
propre

$$\Lambda_I$$

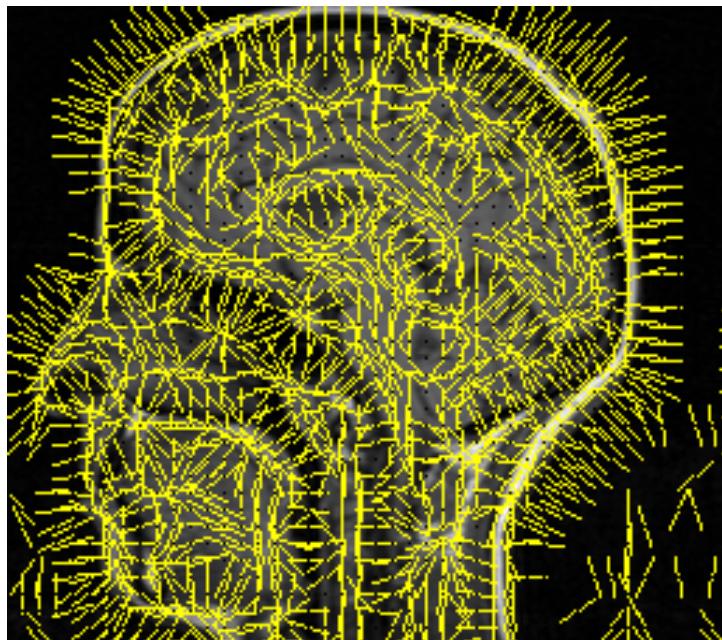


« petite »
valeur
propre

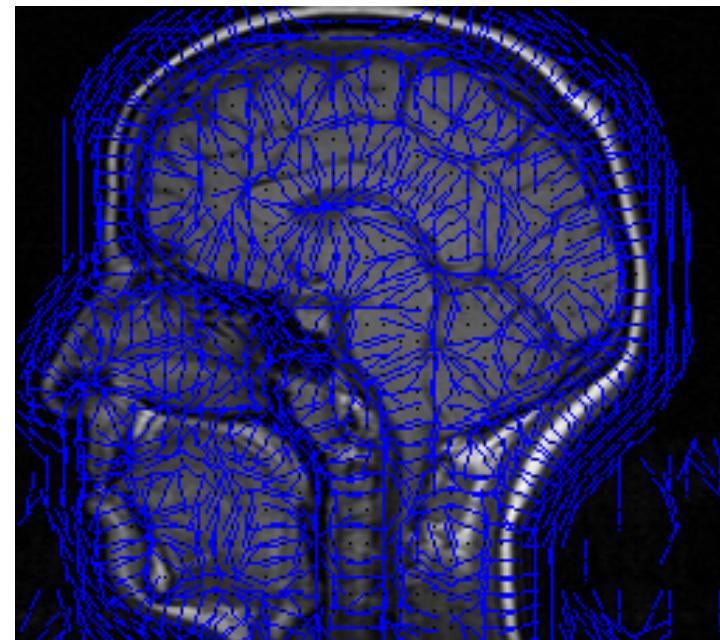
$$\lambda_I$$



direction
du
« grand »
vecteur
propre



direction
du
« petit »
vecteur
propre

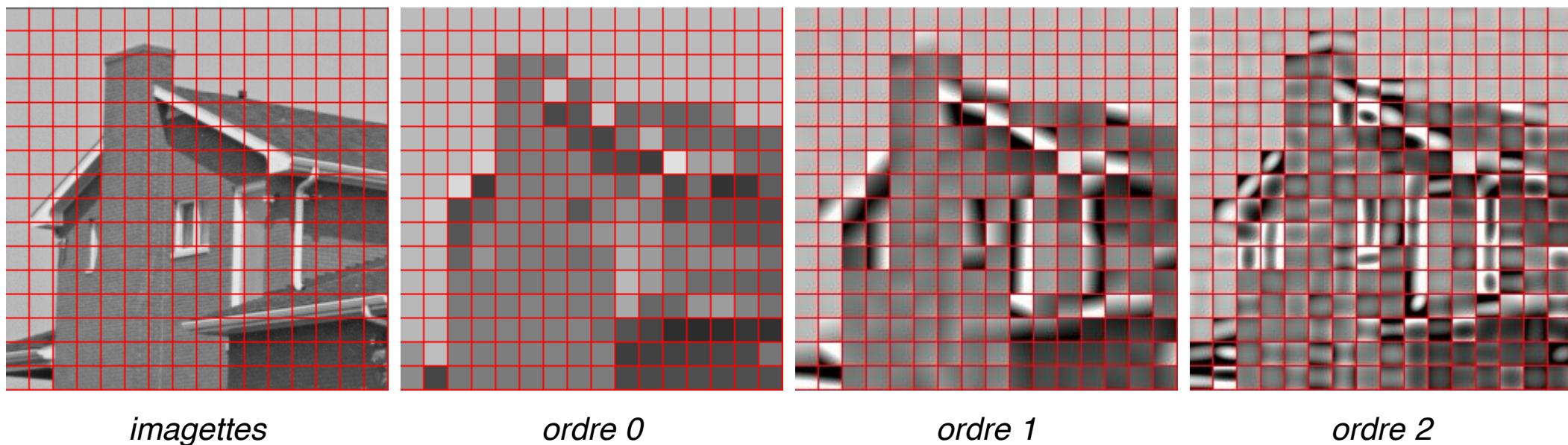


REPRESENTATION PAR LES DERIVEES LOCALES

Expression de la formule de Taylor à l'ordre 2, à partir du vecteur gradient et de la matrice hessienne :

$$I(x_0 + \varepsilon, y_0 + \eta) = I(x_0, y_0) + (\varepsilon, \eta)^T \cdot \nabla I + (\varepsilon, \eta)^T \cdot H_I \cdot (\varepsilon, \eta) + o(\varepsilon^2 + \eta^2)$$

Reconstruction d'imagettes à partir des dérivées partielles calculées au centre de l'imagette, à l'ordre 0, 1 et 2 :



imagettes

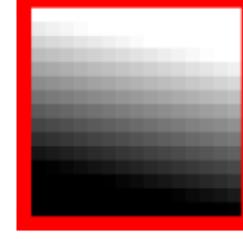
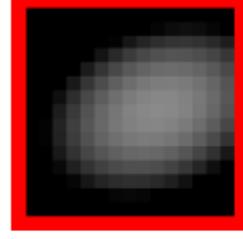
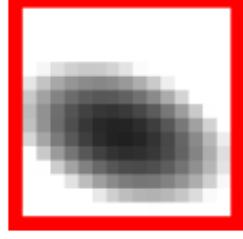
ordre 0

ordre 1

ordre 2

CATEGORISATION DIFFERENTIELLE LOCALE

La valeur des dérivées jusqu'à l'ordre 2 permettent de catégoriser, selon l'ordre dominant, la géométrie locale des pixels en 4 catégories (6 en tenant compte de la polarité) :

0	1	2	
$\ \nabla_I\ \simeq 0$ $\ H_I\ _F \simeq 0$ Plateau	$\ \nabla_I\ \gg 0$ $\ H_I\ _F \simeq 0$ Contour	$\Lambda_I \lambda_I > 0$ Courbure elliptique	$\ H_I\ _F \gg 0$ $\Lambda_I \lambda_I < 0$ Courbure tubulaire
			
		$\Lambda_I < 0$ $\lambda_I < 0$	$\Lambda_I > 0$ $\lambda_I > 0$
		$\Lambda_I < 0$ $\lambda_I > 0$	$\Lambda_I > 0$ $\lambda_I < 0$

ESTIMATION DES DERIVEES ET ESPACE D'ECHELLES

L'idée clef des espaces d'échelles pour le traitement d'images est que toute mesure est relative à une échelle d'estimation.

En particulier une dérivée n'a de sens qu'estimée à une échelle donnée, correspondant à une hypothèse de régularité qui est explicitement réalisée par lissage de l'image. Cette estimation repose sur la commutativité entre dérivation et convolution :

$$\partial^n(I \star g) = I \star (\partial^n g)$$

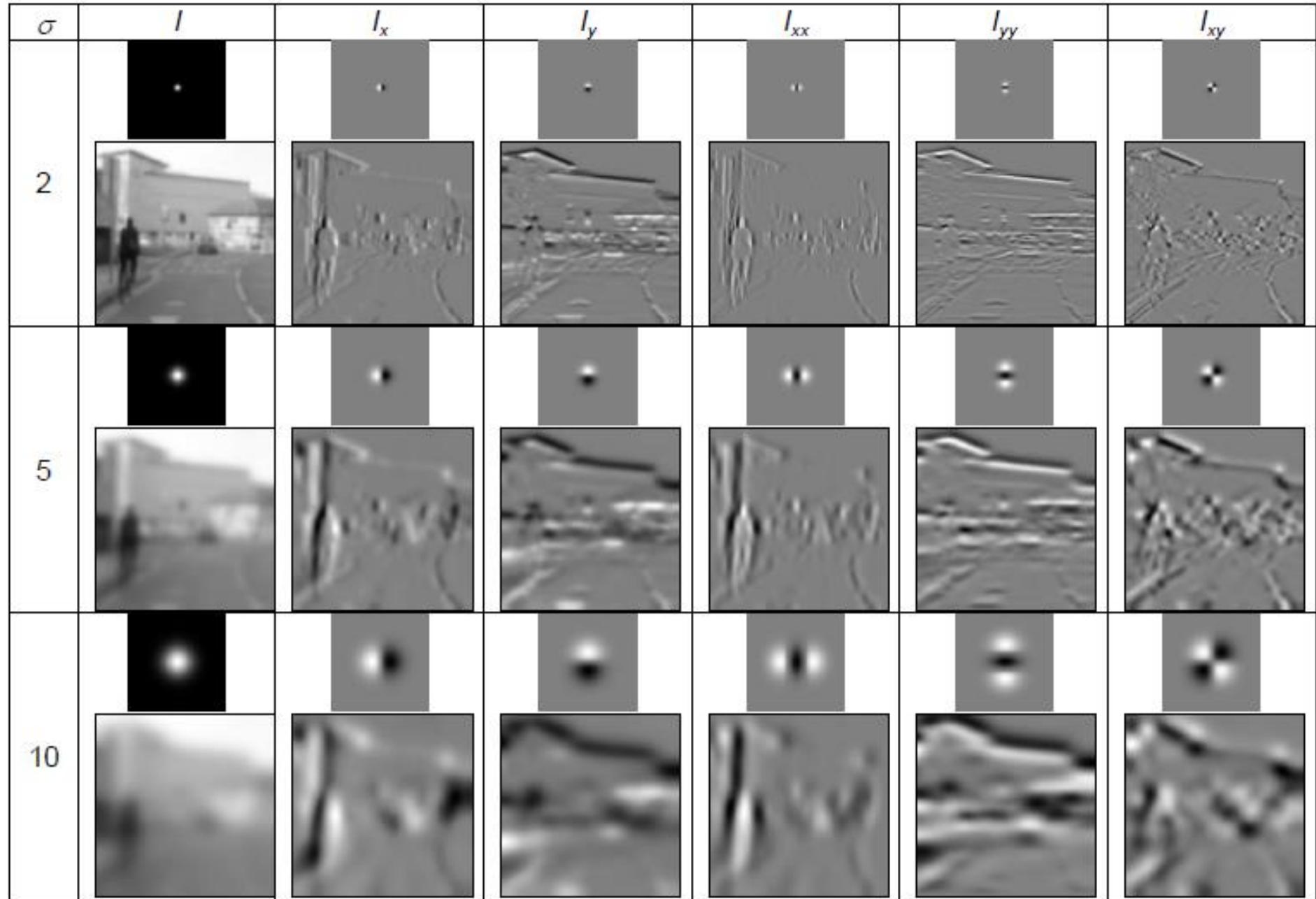
Dans le cadre des espaces d'échelles gaussien, le noyau de convolution g s'identifie au noyau gaussien 2d d'écart-type σ :

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

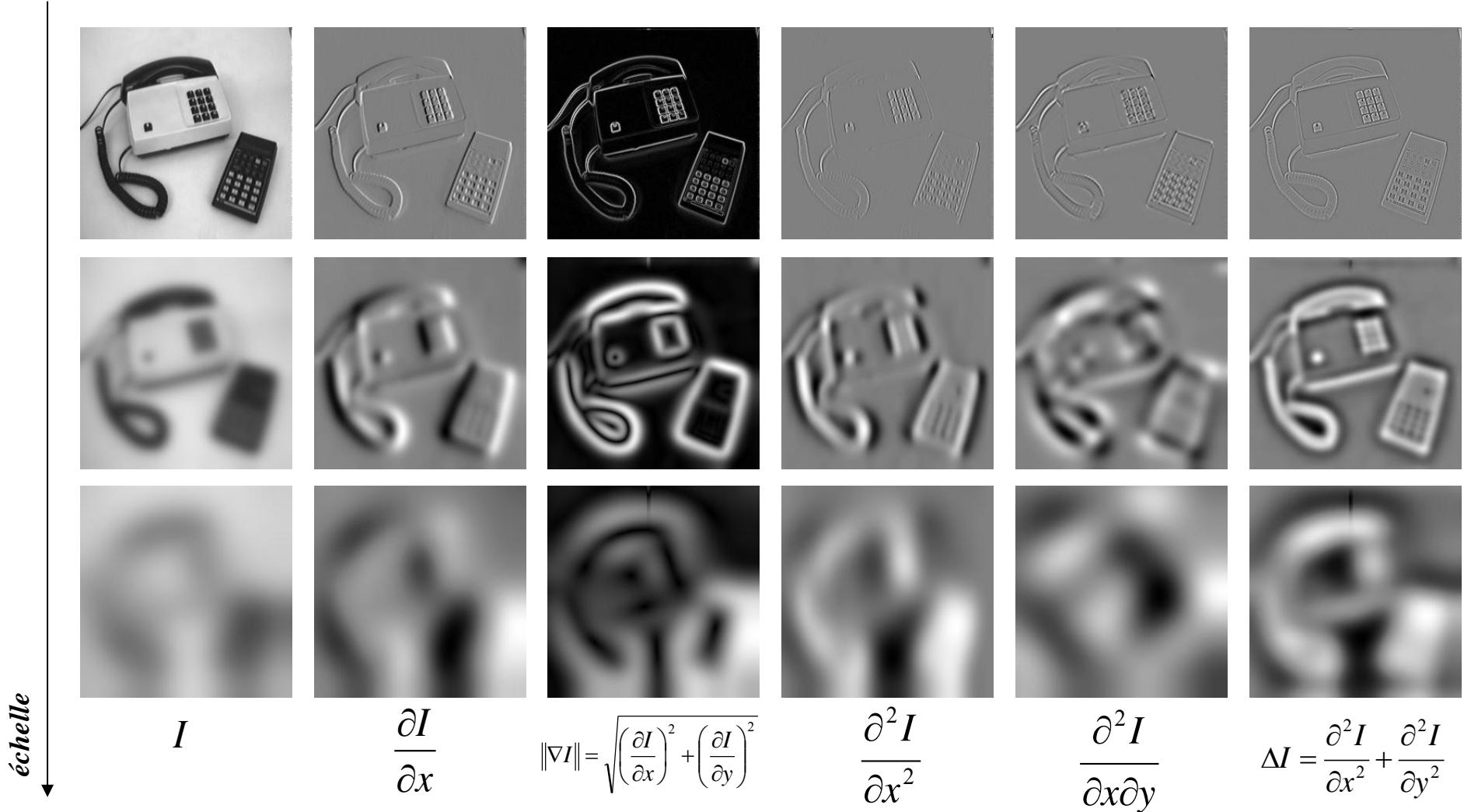
Les dérivées de l'image I estimées à l'échelle σ sont donc définies par les convolutions avec la dérivée de gaussienne correspondante :

$$\left(\frac{\partial^{i+j} I}{\partial x^i \partial y^j} \right)_\sigma \stackrel{\text{déf.}}{=} I \star \left(\frac{\partial^{i+j} G_\sigma}{\partial x^i \partial y^j} \right)$$

DERIVEES MULTI-ECHELLES ET NOYAUX DE CONVOLUTION ASSOCIES

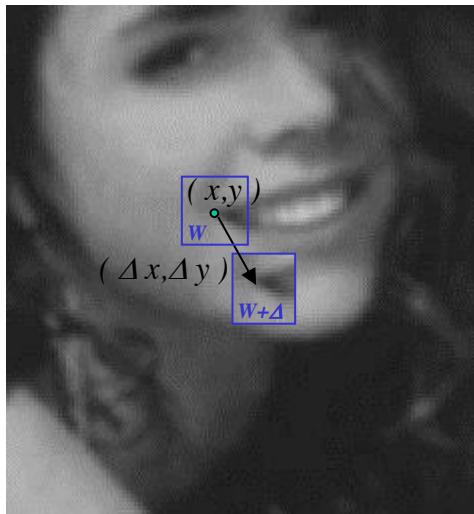


ESPACE D'ECHELLES GAUSSIEN ET GRANDEURS DIFFERENTIELLES



POINTS ANGULEUX ET MATRICE D'AUTOCORRELATION

Les points anguleux (ou points d'intérêt, points saillants,...) sont des points « qui contiennent beaucoup d'information » relativement à l'image. Ce sont des points aux voisinages desquels l'image *varie significativement dans plusieurs directions*.



Une mesure des variations locales de l'image I au point (x,y) associée à un déplacement $(\Delta x, \Delta y)$ est fournie par la *fonction d'autocorrélation* :

$$\chi(x, y) = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2$$

Où W est une fenêtre centrée au point (x,y) .

Or, en utilisant une approximation du premier ordre :

$$I(x_k + \Delta x, y_k + \Delta y) \approx I(x_k, y_k) + \left(\frac{\partial I}{\partial x}(x_k, y_k) - \frac{\partial I}{\partial y}(x_k, y_k) \right) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Et donc :

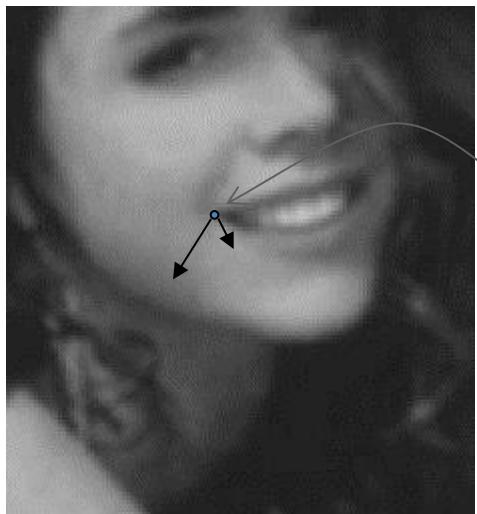
$$\chi(x, y) = \sum_{(x_k, y_k) \in W} \left(\left(\frac{\partial I}{\partial x}(x_k, y_k) - \frac{\partial I}{\partial y}(x_k, y_k) \right) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right)^2 = (\Delta x \quad \Delta y) \underbrace{\begin{pmatrix} \sum_{(x_k, y_k) \in W} \left(\frac{\partial I}{\partial x}(x_k, y_k) \right)^2 & \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) & \sum_{(x_k, y_k) \in W} \left(\frac{\partial I}{\partial y}(x_k, y_k) \right)^2 \end{pmatrix}}_{\Xi(x, y)} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Matrice d'autocorrélation de l'image I en (x,y)

MATRICE D'AUTOCORRELATION ET DETECTEUR DE HARRIS

$$\Xi(x, y) = \begin{pmatrix} \sum_{(x_k, y_k) \in W} (\frac{\partial I}{\partial x}(x_k, y_k))^2 & \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) & \sum_{(x_k, y_k) \in W} (\frac{\partial I}{\partial y}(x_k, y_k))^2 \end{pmatrix}$$

La matrice d'autocorrélation Ξ représente la variation locale de l'image I en (x,y) . (x,y) sera considéré comme un point anguleux de I si pour tous les déplacements $(\Delta x, \Delta y)$, la quantité $(\Delta x, \Delta y) \cdot \Xi(x, y) \cdot (\Delta x, \Delta y)^t$ est grande.



Les points anguleux sont les points (x,y) pour lesquels la matrice d'autocorrélation $\Xi(x,y)$ a *deux valeurs propres grandes*.

Cela correspond aux points pour lesquels il existe localement une base de vecteurs propres de Ξ décrivant des variations locales importantes de l'image.

Le *détecteur de Harris* calcule une *fonction d'intérêt* $\Theta(x,y)$:

$$\Theta(x, y) = \det \Xi - \alpha \operatorname{trace} \Xi$$

Le premier terme correspond au produit des valeurs propres, le second terme pénalise les points de contours avec une seule forte valeur propre.

Les points d'intérêt correspondent aux maxima locaux de la fonction Θ qui sont au delà d'un certain seuil (typiquement 1% de la valeur max de Θ).

[Harris 88]

CALCUL DE LA FONCTION D'INTERET DE HARRIS

1. On calcule les dérivées premières à partir des dérivées de gaussienne (écart-type σ_1)
2. On calcule les termes de la matrice d'autocorrélation Ξ en calculant une moyenne locale des dérivées sous la forme d'une gaussienne (écart-type σ_2 , typiquement $\sigma_2 = 2 \sigma_1$)
3. On calcule la fonction d'intérêt : $\Theta = \det(\Xi) - \alpha \operatorname{trace}(\Xi)$ (typiquement $\alpha = 0,06$).
4. On calcule les maxima locaux de Θ supérieurs à un certain seuil (typiquement 1% de Θ_{\max}).



POINTS DE HARRIS MULTI-ECHELLES



$\sigma = 1$



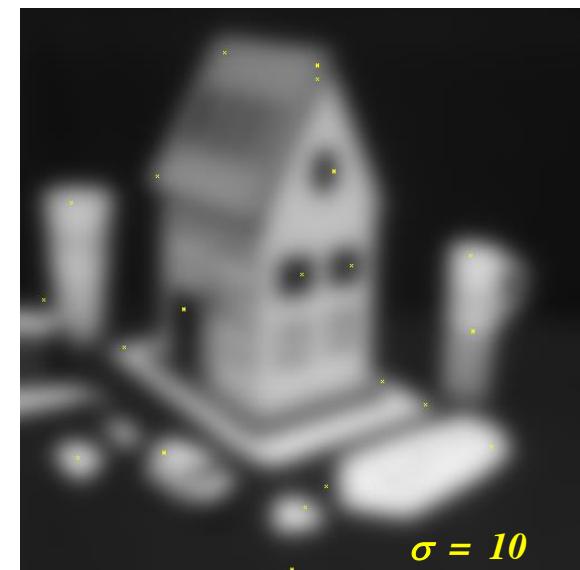
$\sigma = 2$



$\sigma = 3$



$\sigma = 5$



$\sigma = 10$

Points de Harris obtenus en calculant les dérivées premières par convolution avec une dérivée de gaussienne d'écart-type σ .

DETECTEUR SIFT : EXTREMA DANS L'ESPACE D'ECHELLES

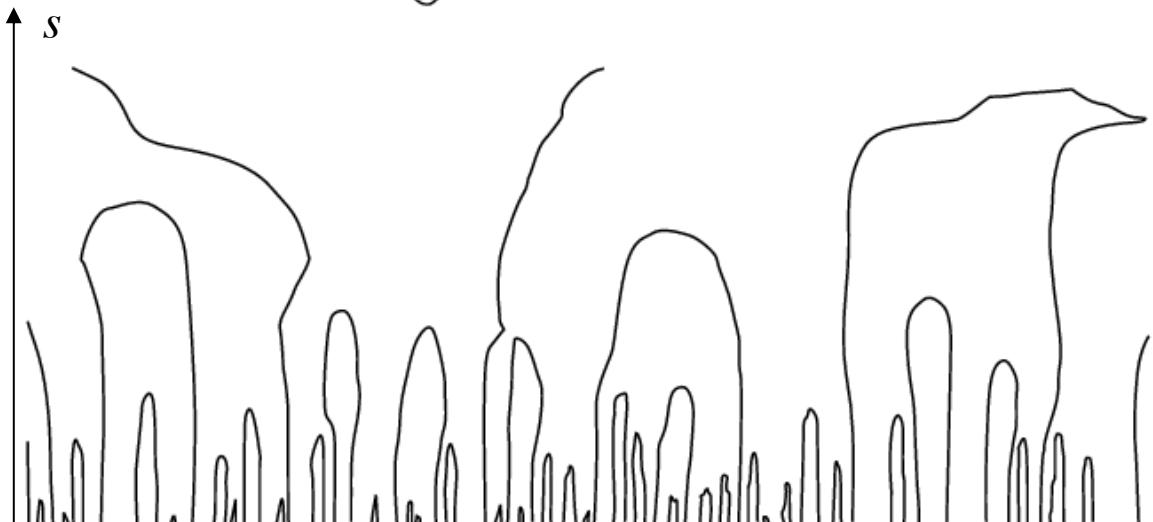
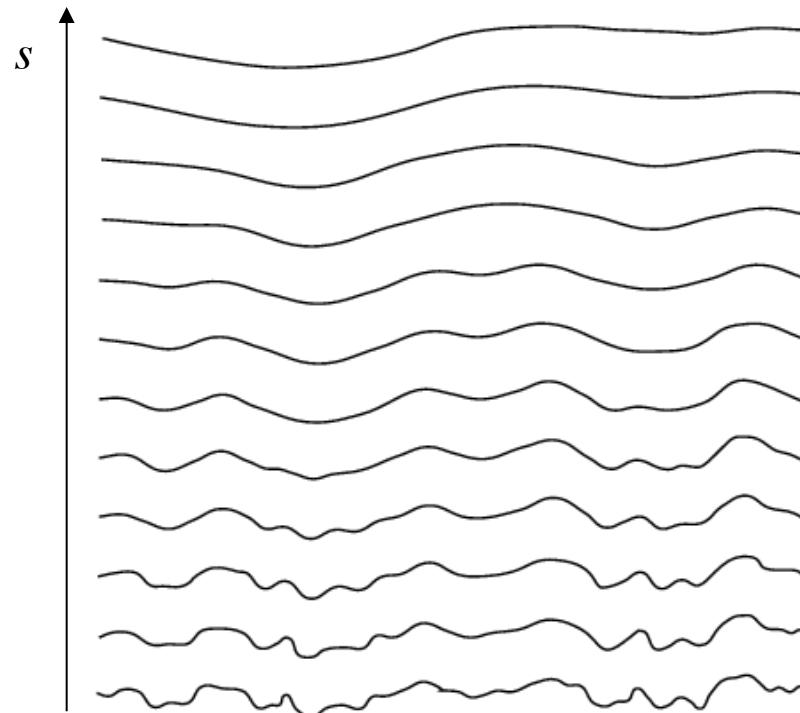
Le détecteur SIFT (Scale Invariant Feature Transform) utilise une approche de la notion de structure / point d'intérêt mieux adaptée aux grandes échelles que celle de point anguleux :

Le *blob* (structure elliptique)

Cette structure peut se caractériser à toutes les échelles et correspond au point de l'espace-échelle (x,y,s) où un extremum local disparaît.

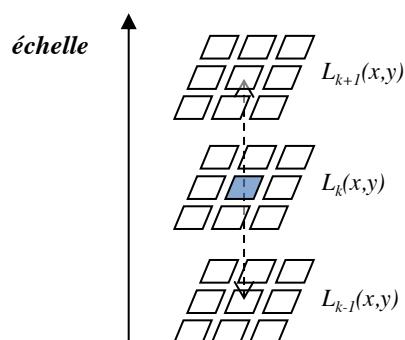
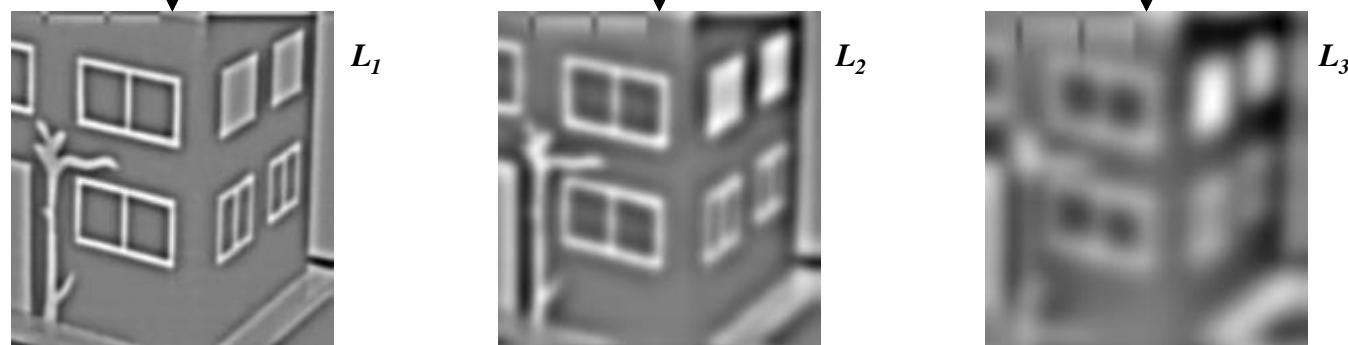
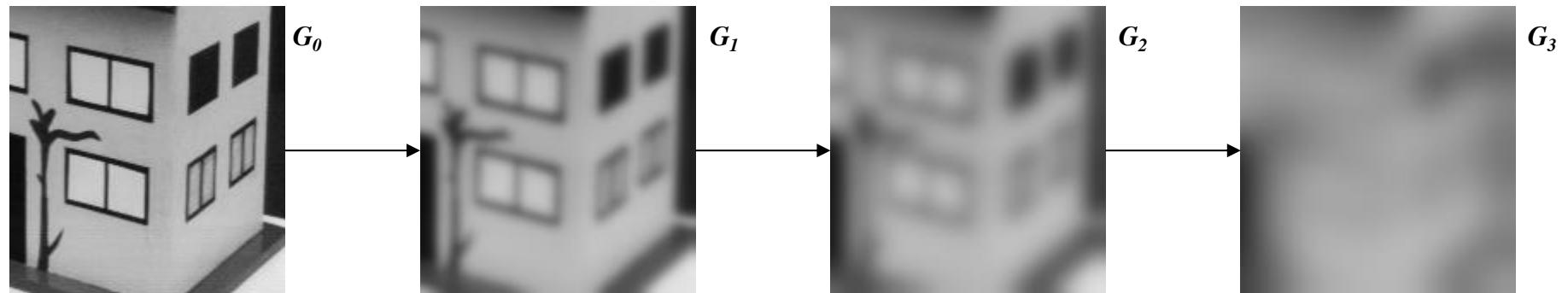
Cf principe de causalité des espaces d'échelles.

En 1d (ci-contre) : point d'échelle s maximum sur chaque courbe de l'empreinte (fingerprint) de l'espace d'échelles.



Espace d'échelle gaussien d'un signal 1d (en haut) et position des extrema dans l'espace-échelle (x,s) (en bas).

DETECTEUR SIFT : EXTREMA DANS L'ESPACE D'ECHELLES



La fonction $G_k(x,y) = G(x,y,k\sigma)$ est l'image convoluée par une gaussienne d'écart-type $k\sigma$. Les fonctions $L_k(x,y)$ correspondent à la différence (ici normalisée) entre 2 gaussiennes adjacentes.

La fonction $L_k(x,y)$ est une représentation laplacienne de l'image, qui correspond à une décomposition fréquentielle localisée : contribution des structures contrastées d'échelle (de « taille ») $k\sigma$ au point (x,y) .

Les points sélectionnés par SIFT sont les maxima et les minima locaux de la fonction $L_k(x,y)$, à la fois dans l'échelle courante et dans les échelles adjacentes (voir ci-contre).

[Lowe 04]

POINTS D'INTERET SIFT

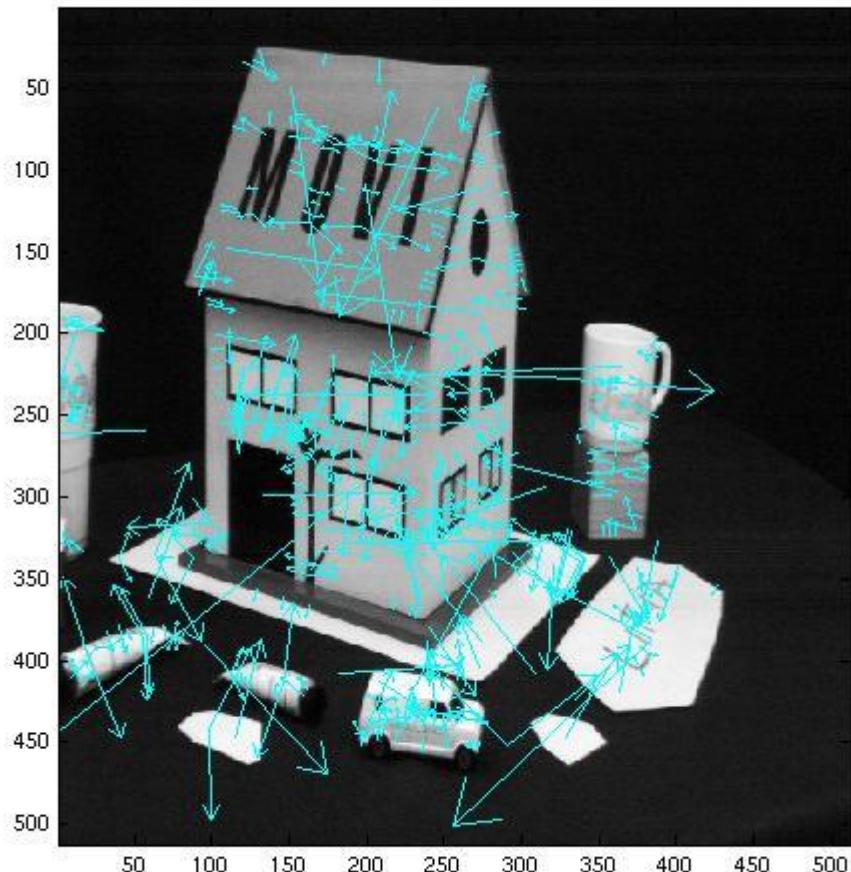


Image 1 : 589 points détectés.

[Lowe 04]

Pour chaque extrema de l'espace d'échelle des différences de gaussiennes (point d'intérêt SIFT), on calcule la direction associée par :

$$\theta(x, y) = \arctan\left(\frac{G_y^\sigma(x, y)}{G_x^\sigma(x, y)}\right)$$

avec $G_x^\sigma(x, y) = \frac{\partial}{\partial x} G(x, y, \sigma) = I(x, y) * \frac{\partial}{\partial x} g_\sigma(x, y)$

(où σ est l'échelle sélectionnée)

Ci-contre, point d'intérêt SIFT : la direction de la flèche représente la direction θ et sa longueur l'échelle σ associée.

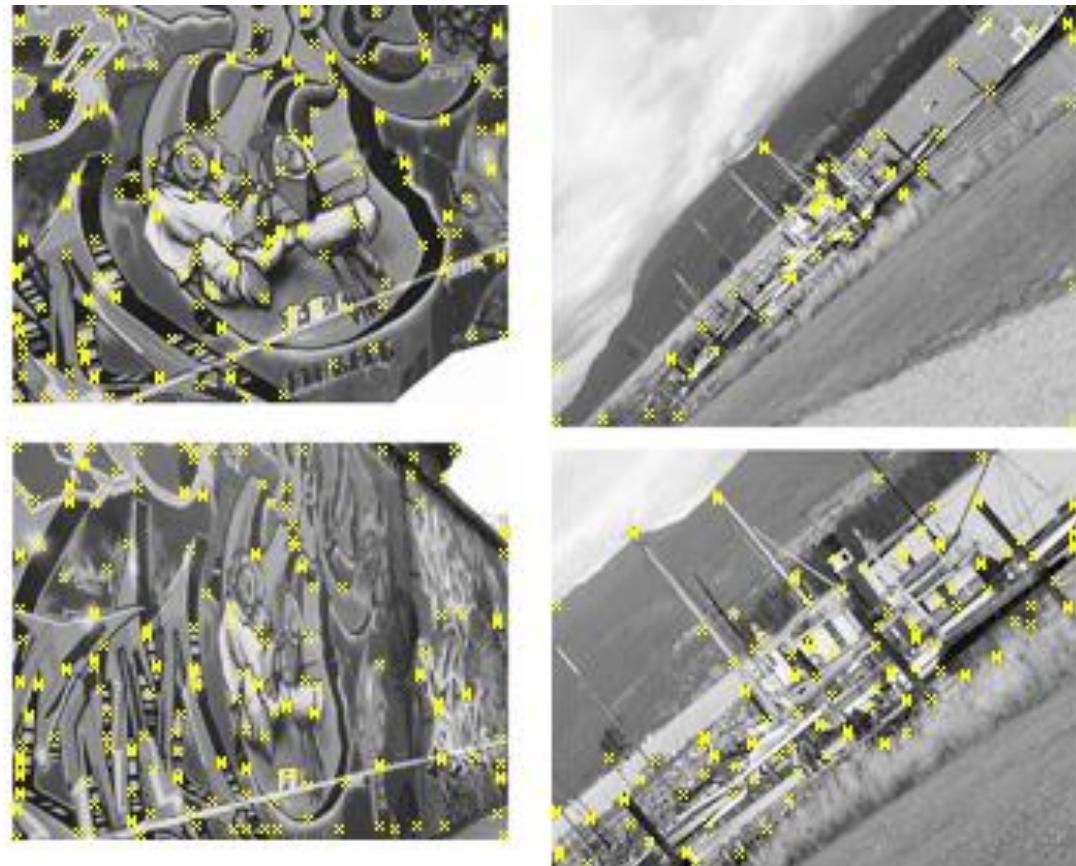
EVALUATION DES DETECTEURS DE POINTS D'INTERET

La plupart des détecteurs de point d'intérêt sont définis indépendamment des descripteurs avec lesquels on les utilise. Il est donc nécessaire de pouvoir les évaluer en eux-mêmes.

Les propriétés recherchées d'un bon détecteur :

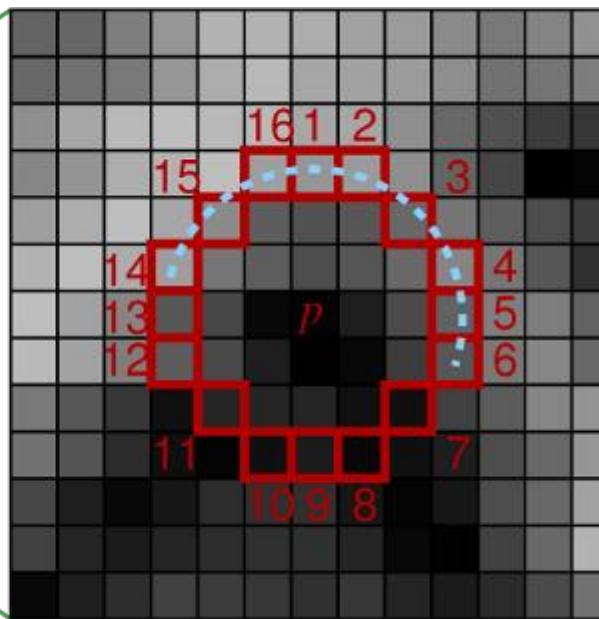
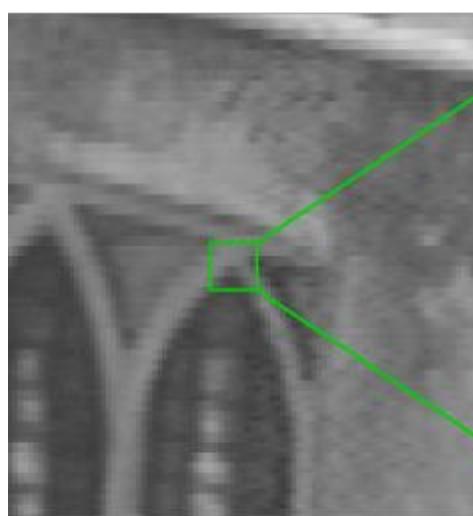
- **Répétabilité** : le point doit apparaître aux mêmes endroits quelque soit la déformation.
- **Représentativité** : les points doivent être le plus nombreux possible.
- **Efficacité** : le détecteur doit être rapide à calculer (cf SURF, FAST)

(Rq : répétabilité et représentativité ne sont pas indépendants !)



[Schmid 2000]

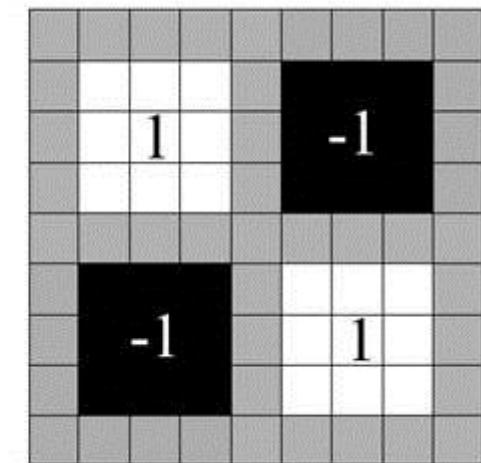
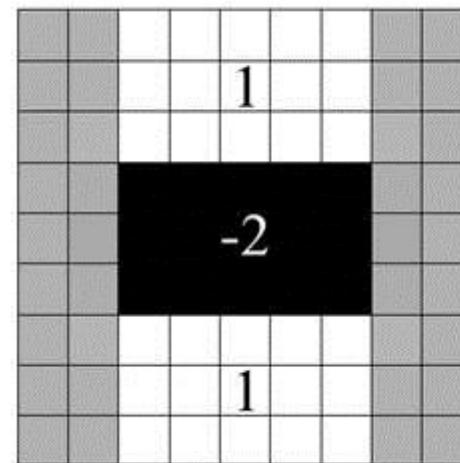
DES DETECTEURS RAPIDES A CALCULER ?



Le détecteur FAST sélectionne les points dont le voisinage circulaire présente des plages contiguës assez longues de points significativement plus clairs (resp. plus sombres).

[Rosten 05]

Le détecteur SURF approxime le calcul des dérivées secondes dans des voisinages rectangulaires à l'aide d'images intégrales, et sélectionne les maxima locaux du déterminant de la matrice hessienne.



[Bay 06]

Introduction sur les descripteurs (feature descriptors)

- Quelles propriétés escomptées ?
 - Des Patches avec un contenu similaire doivent avoir des descripteurs proches (similaires).
 - Les exigences d'invariance (ou de robustesse) par rapport à des transformations sont multiples : rotation d'images, transformation d'échelle, changement de point de vue, changement de luminosité, changement d'objets dans la catégorie, ...



DESCRIPTEURS : INVARIANTS DIFFÉRENTIELS

Objectif: représenter les points d'intérêt par des *indices* qui soient *invariants* par *rotation* et par *changement d'échelle*.

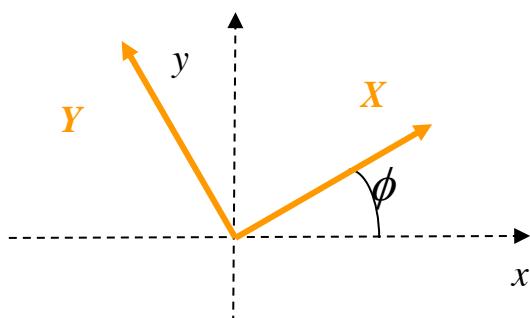
Le principe utilisé ici est basé sur l'utilisation des dérivées spatiales multi-échelle :

Le « *jet local* » de I : $L_{ij}^\sigma = I * G_{ij}^\sigma$ avec : $G_{ij}^\sigma = \frac{\partial^{i+j}}{\partial x^i \partial y^j} G^\sigma$ et : $G^\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$

On notera : $\{L_{ij}^\sigma; 0 \leq i + j \leq 3\} = \{L, L_x, L_y, L_{xx}, L_{xy}, L_{yy}, L_{xxx}, L_{xxy}, L_{xyy}, L_{yyy}\}$ (dérivées jusqu'au 3e ordre)

L'idée est de *combiner ces dérivées* pour obtenir des grandeurs *invariantes par rotation* :

Par exemple, le laplacien $I_{xx} + I_{yy}$ est invariant par rotation :



$$\begin{aligned} &\begin{cases} x = X \cos \phi + Y \sin \phi \\ y = X \sin \phi - Y \cos \phi \end{cases} \quad \begin{cases} X = x \cos \phi + y \sin \phi \\ Y = -x \sin \phi + y \cos \phi \end{cases} \\ &\begin{cases} I_X = I_x \cos \phi + I_y \sin \phi \\ I_Y = I_x \sin \phi - I_y \cos \phi \end{cases} \quad \text{et : } \begin{cases} I_{XX} = I_{xx} \cos^2 \phi + 2I_{xy} \cos \phi \sin \phi + I_{yy} \sin^2 \phi \\ I_{YY} = I_{xx} \sin^2 \phi - 2I_{xy} \cos \phi \sin \phi + I_{yy} \cos^2 \phi \end{cases} \end{aligned}$$

et donc : $I_{XX} + I_{YY} = I_{xx} + I_{yy}$

DESCRIPTEURS : INVARIANTS DIFFÉRENTIELS

Invariance aux rotations image : invariants différentiels

On peut ainsi construire toute une famille de grandeurs invariantes par rotation : les invariants différentiels de Hilbert.

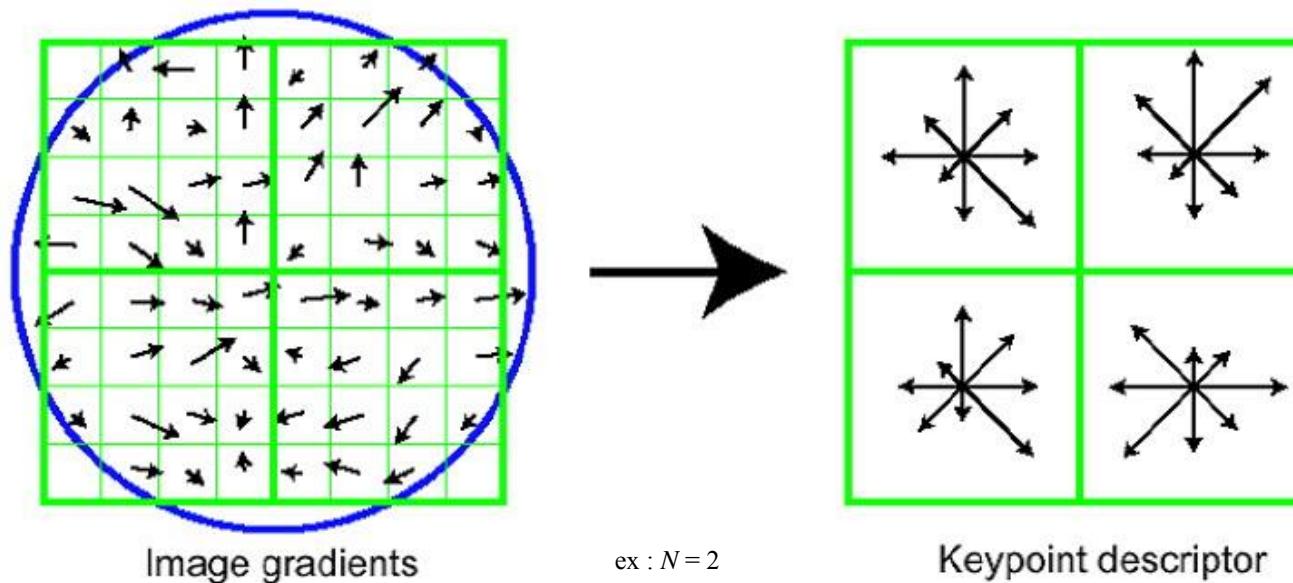
The diagram illustrates the construction of differential invariants from the gradient norm and Laplacian. It features two horizontal arrows pointing towards a vertical stack of terms enclosed in brackets. The top arrow is labeled "Norme du gradient" and points to the first term, which is the gradient norm squared, $L_x^2 + L_y^2$. The bottom arrow is labeled "Laplacien" and points to the second term, which is the Laplacian squared, $L_{xx}^2 + 2L_{xy}L_{yx} + L_{yy}^2$. To the right of these terms, a vertical bracket groups them under the label L .

$$L \left[\begin{array}{l} L_x^2 + L_y^2 \\ L_{xx}^2 + 2L_{xy}L_{yx} + L_{yy}^2 \end{array} \right]$$

DESCRIPTEURS SIFT : HISTOGRAMMES LOCAUX D'ORIENTATION

Les descripteurs associés aux points d'intérêt SIFT sont des histogrammes des orientations locales autour du point d'intérêt.

- On divise l'espace autour de chaque point d'intérêt (x,y) en N^2 carrés 4x4.
- On calcule le gradient ($G_x(a,b,\sigma), G_y(a,b,\sigma)$) pour les $4 \times 4 \times N^2$ points (a,b) .
- Pour chaque carré 4x4, on calcule un histogramme des orientations quantifiées en 8 directions, en pondérant par : (1) le module du gradient (2) l'inverse de la distance au point d'intérêt (x,y) .
- Pour être invariant en rotation : l'orientation locale du point d'intérêt $\theta(x,y)$ est utilisée comme *origine* (orientation nulle) des histogrammes.

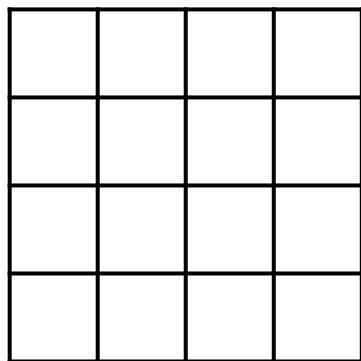


Les descripteurs formés sont donc des vecteurs de taille $8 \times N^2$, qui seront appariés en utilisant une distance (e.g. distance euclidienne)

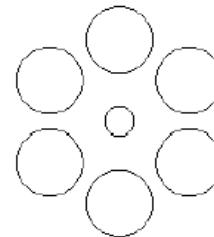
[Lowe 04]

Extension: Daisy

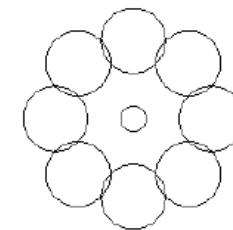
Circular gradient binning



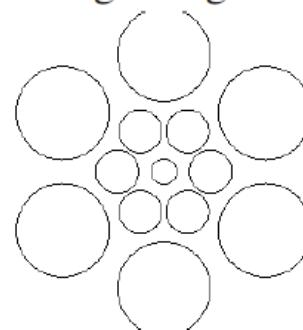
SIFT



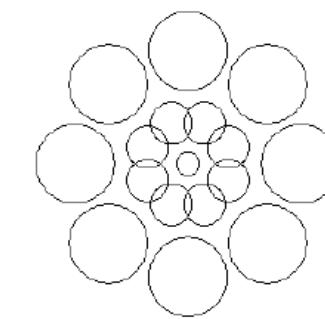
1 Ring 6 Segments



1 Ring 8 Segments



2 Rings 6 Segments



2 Rings 8 Segments

Daisy

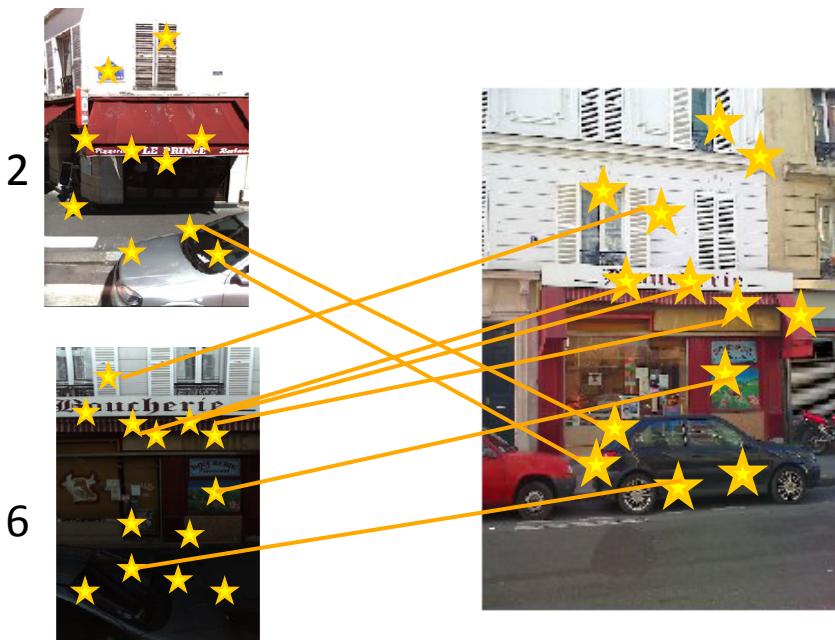
DU LOCAL AU GLOBAL : CONSENSUS DES DESCRIPTEURS LOCAUX

Souvent l'utilisation des primitives visuelles aboutit à une décision prise à un niveau global sur l'image ou l'ensemble de points : étiquette de classe (reconnaissance, catégorisation), paramètres de déplacement (odométrie visuelle).

Comment réaliser cette décision collective à partir de l'ensemble des descripteurs ?

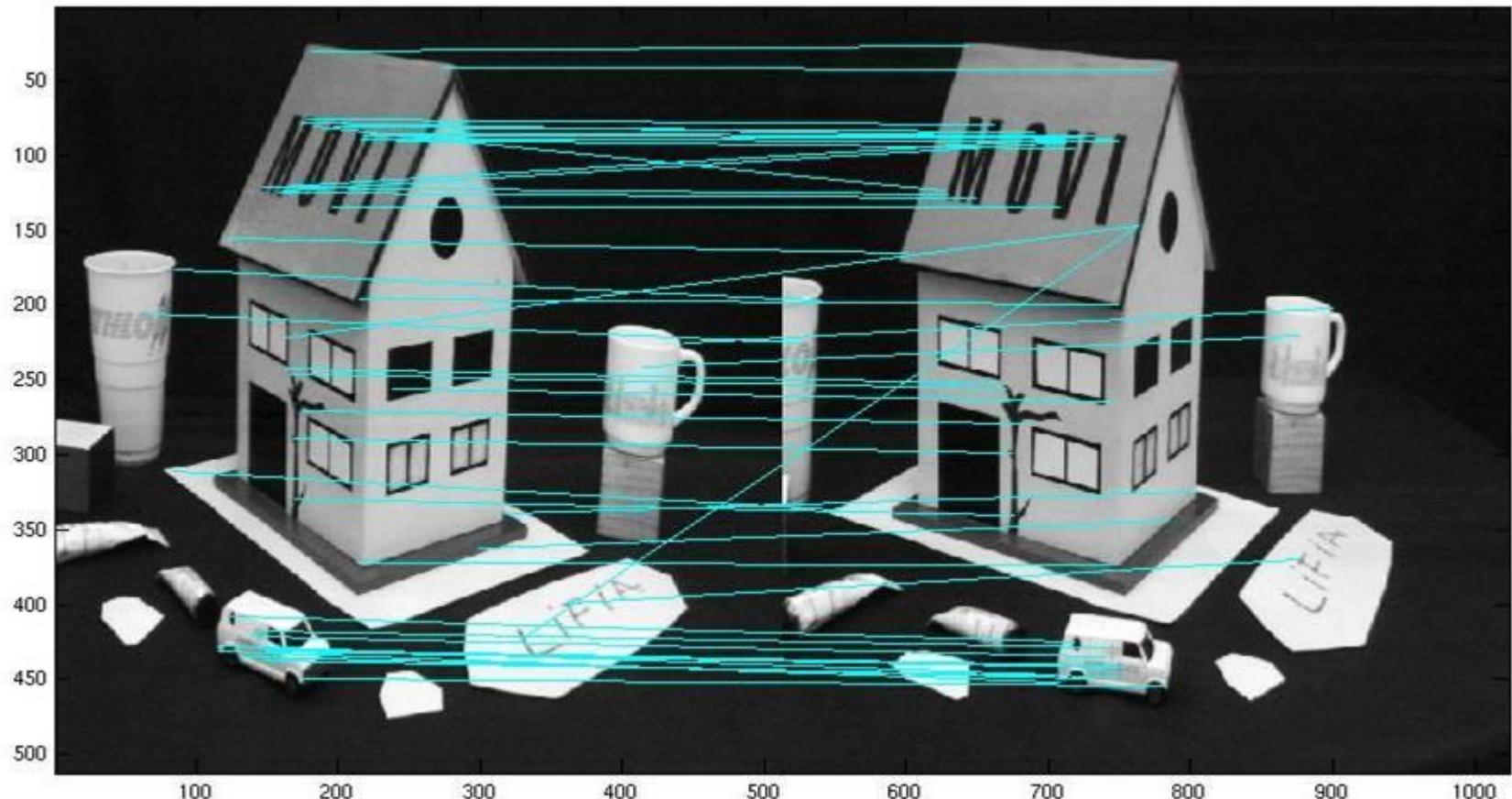
Principe de vote : chaque descripteur local est classifié et la classe globale est attribuée selon un critère majoritaire (ex : reconnaissance de pièces, catégories d'images...)

Sélection par cohérence : un sous-ensemble des appariements locaux est (itérativement) sélectionné pour établir une décision cohérente (ex : odométrie visuelle...)



Alternative: BoW et distances classiques

APPARIEMENT DES POINTS SIFT



Résultat d'appariement par SIFT entre l'image (2) à gauche, 510 points détectés, et l'image (1) à droite, 589 points détectés. 51 points ont été appariés, ce qui correspond à une distance euclidienne entre les descripteurs en deçà d'un certain seuil.

[Lowe 04]

APPARIEMENT DE PRIMITIVES : METRIQUES

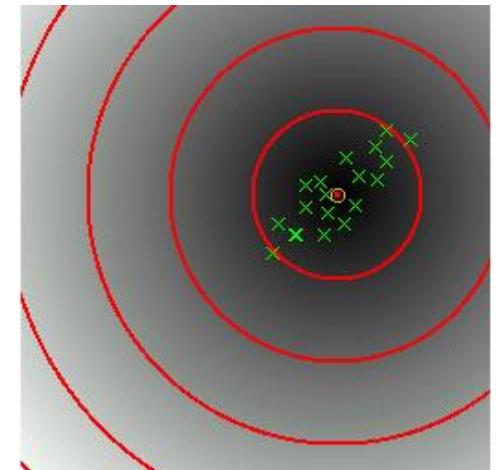
L'appariement entre primitives repose donc en général sur la comparaison de descripteurs locaux deux à deux.

Si le couple détecteur / descripteur possède l'invariance voulue, la comparaison peut être réalisée par une métrique simple :

La distance euclidienne :

$$\delta_e(x, x')^2 = (x - x')^T (x - x')$$

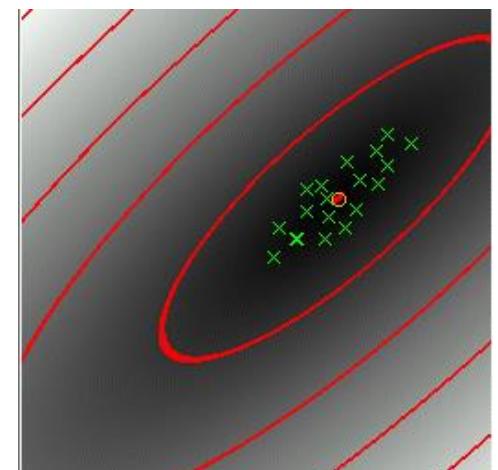
Cette distance ne tient compte ni des différences d'amplitude ni des éventuelles corrélations entre les différentes composantes des descripteurs.



La distance de Mahalanobis :

$$\delta_m(x, x')^2 = (x - x')^T C^{-1} (x - x')$$

Avec $C = (\text{cov}(x_i, x_j))_{i,j}$ matrice de covariance de la base des descripteurs.



APPARIEMENT DE PRIMITIVES : METRIQUES

Dans le cas de grande base de descripteurs (indexation), la matrice de covariance est calculée et mise à jour off-line. Si l'on diagonalise C^{-1} , on se ramène à un calcul de distance euclidienne sur des descripteurs normalisés :

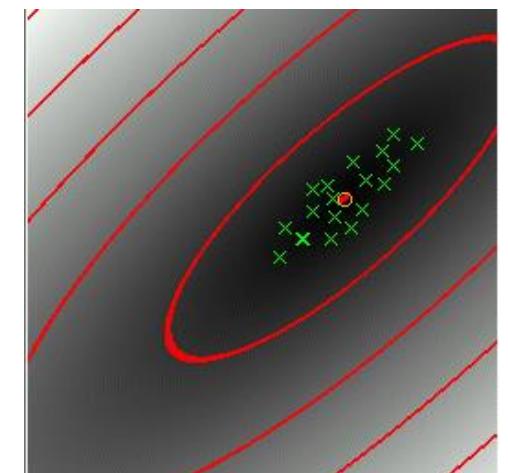
$$C^{-1} = P^T D P$$

$$\delta_m(x, x') = \sqrt{(x - x')^T C^{-1} (x - x')} = \underbrace{\|\sqrt{D}Px - \sqrt{D}Px'\|}_{\text{distance ellipsoïdale}}$$

A chaque mise à jour de la base d'index, on doit donc :

distance ellipsoïdale

- Mettre à jour la matrice de covariance : C
- Calculer et diagonaliser : C^{-1}
- Normaliser tous les vecteurs : $x \rightarrow \sqrt{D}Px$



REFERENCES

- **C. Harris & M. Stephens 1988** « *A combined corner and edge detector* » *Alvey Vision Conference pp 147-151*
- **D.G. Lowe 2004** « *Distinctive Image Features from Scale-Invariant Keypoints* » *International Journal of Computer Vision 60(2) pp 91-110*
- **C. Schmid, R. Mohr & C. Bauckhage 2000** « *Evaluation of Interest Point Detectors* » *Int. Jornal of Computer Vision 37(2) pp 151-172*
- **C. Schmid & R. Mohr 1997** « *Local grayvalue invariants for image retrieval* » *IEEE Transactions on Pattern Analysis and Machine Intelligence 19(5) pp 530-534*
- **E. Rosten & T. Drummond** “Fusing points and lines for high performance tracking” *Int. Conf. on Computer Vision (ICCV 2005), 1508—1511, 2005.*
- **H. Bay, T. Tuytelaars & L. Van Gool** “SURF: Speeded up robust features”, *Computer Vision and Image Understanding, 110 (3), June, 2008, 346-359*

REFERENCES

- **N. Dalal & B. Triggs 2005** « *Histogram of oriented gradients for human detection* », *Int. Conf. Of Computer Vision and Pattern recognition (CVPR), 2005*
- **G. Csurka, C.R. Dance, L. Fan, J. Willamowski & C. Bray**, "Visual categorization with bags of keypoints", In Workshop on Statistical Learning in Computer Vision, ECCV, 2004.
- **B. Tomasik, P. Thiha & D. Turnbull 2009** « *Tagging products using image classification* », *SIGIR 2009*.
- **H. Foroosh, J. Zerubia & M. Berthod 2002** « *Extension of phase correlation to subpixel registration* » *IEEE Transactions on Image Processing 11(3) pp 188-200*
- **Q. Chen, M. Defrise & F. Deconinck 1994** « *Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition* » *IEEE Transactions on Pattern Analysis and Machine Intelligence 16(12) pp 1156-1168*