

UE Vision, Réalité Augmentée et Applications

Matière : Vision

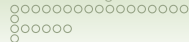
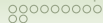
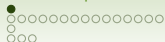
Partie II

(Cours)

Sylvie CHAMBON

schambon@toulouse-inp.fr

26 septembre et 2 octobre 2024



Plan

Détection de points d'intérêt

- Introduction

- Notations

- Construction d'une pyramide

DéTECTEURS du premier ordre

- DéTECTEURS du premier ordre : vers le multi-échelle

DéTECTEURS de régions

- Vers le multi-échelle : Détecteur de Kadir

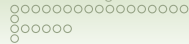
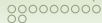
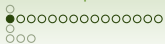
- Vers le multi-échelle : EBR, *Edge-Based Regions*

- Vers le multi-échelle : PCBR, *Principal Curvature-Based Regions*

DéTECTEURS du second ordre

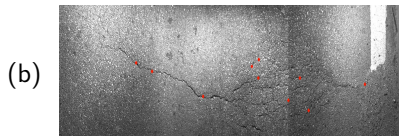
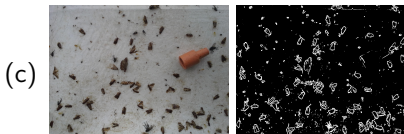
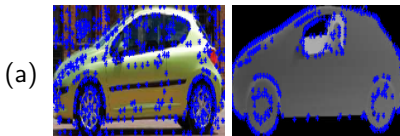
- Vers le multi-échelle : SIFT *Scale Invariant Feature Transform*

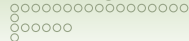
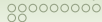
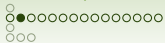
- Vers une variante accélérée de SIFT : SURF



Principe d'un détecteur

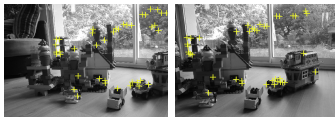
- **primitive** = élément pouvant être extrait d'une image afin d'obtenir une information caractéristique de la scène comme la présence d'un objet, une distance, un relief ...
- primitives = pixels, régions, contours, polygones, tout ensemble de points permettant d'obtenir l'information recherchée
- **primitive d'intérêt - points d'intérêt**



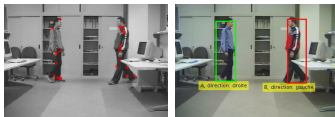


Applications

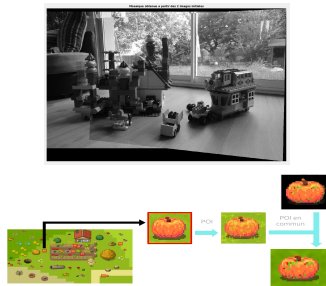
(e)

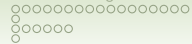
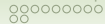
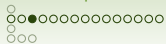


(f)



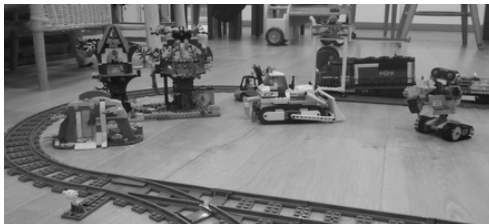
(g)

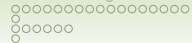
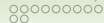
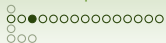




Fonctionnement d'un détecteur

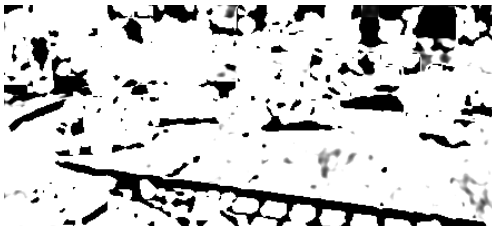
Image originale

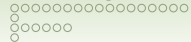
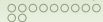
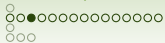




Fonctionnement d'un détecteur

(1) Réponse

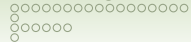
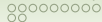
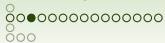




Fonctionnement d'un détecteur

(2) *Sélection des points après suppression non-maxima locaux*

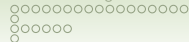
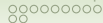
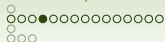




Fonctionnement d'un détecteur

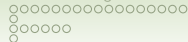
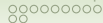
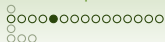
(3) Affinement de la sélection





Fonctionnement d'un détecteur : analyse

- Sans suppression des non maxima locaux
Nous avons un risque de points mal répartis
- Sans Affinement mais avec suppression des non maxima locaux
 - Points bien répartis sur toute l'image
 - Mais certains points n'auront pas des caractéristiques suffisamment saillantes pour être suivis correctement (risque élevé d'erreurs de suivi)
 - Certaines zones de l'image ne posséderont aucun point suivi



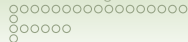
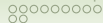
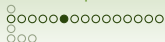
Qualité du résultat ?

La répétabilité d'une primitive correspond à sa capacité à être détectée quelle que soit l'image ou la donnée dans laquelle elle apparaît. Plus précisément, si une primitive est détectée dans une représentation donnée, alors, elle doit également être détectée dans une autre représentation.

Si nous notons :

- $\mathbf{p}$, un point dans une image I_1
- $\mathbf{p}'$ son correspondant théorique dans une autre image, I_2 alors ce point est répétable s'il existe un point $\mathbf{q}$, détecté dans I_2 et tel que :

$$\|\mathbf{p}' - \mathbf{q}\| \leq \epsilon.$$



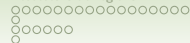
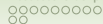
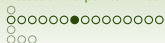
Qualité du résultat ?

Taux de répétabilité entre deux images, noté $R(l_1, l_2)$

$$R(\mathcal{P}_{D,1}, \mathcal{P}_{D,2}) = \frac{R(\mathcal{P}_{D,1} \rightarrow \mathcal{P}_{D,2}) + R(\mathcal{P}_{D,2} \rightarrow \mathcal{P}_{D,1})}{2} \quad (1)$$

avec

$$R(\mathcal{P}_{D,i} \rightarrow \mathcal{P}_{D,j}) = \frac{\#\{\text{Points répétables de } \mathcal{P}_{D,i} \text{ vers } \mathcal{P}_{D,j}\}}{\#\{\mathcal{P}_{D,i}\}} \quad (2)$$

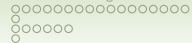
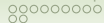
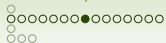


Qualité du résultat ?

- Détection exacte et précise : la localisation peut être imprécise, cela est dû aux différents opérateurs de lissage ou d'approximation utilisés.

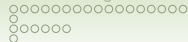
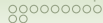
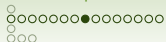
De nombreuses méthodes avec une étape de **relocalisation**.

- Robustesse à certaines difficultés :
 - Définition locale pour permettre la robustesse aux occultations
 - Invariant, ou co-variant, aux transformations d'images
 - Robuste à tous types de dégradations de l'image : bruits, flou, discrétisation ou encore compression.
 - Assurance d'une grande quantité de primitives d'intérêt
 - Détection rapide



Détection *versus* description

Quelle est le rôle du descripteur par rapport au détecteur ?



Détection *versus* description

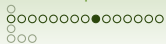
Quelle est le rôle du descripteur par rapport au détecteur ?

- Détecteur
- = Opérateur pour déterminer si c'est un point d'intérêt ou non
- Descripteur
- = Toutes les caractéristiques exploitées pour réaliser le suivi/la mise en correspondance
- Beaucoup plus de détecteurs que de descripteurs
- Exploiter la réponse pour faire le suivi est peu efficace

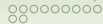
Deux solutions

1. Simple mesure de corrélation dans le voisinage
2. Définir des caractéristiques plus élaborées qu'un simple voisinage couleur (SIFT)

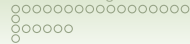
Détection de points d'intérêt



Détecteurs du premier ordre



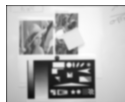
Détecteurs de régions



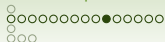
Détecteurs du second ordre



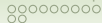
Notions d'analyse multi-échelle et multi-résolution



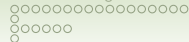
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

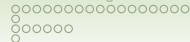
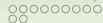
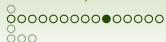


Détecteurs du second ordre



Notions d'analyse multi-échelle et multi-résolution

Est-ce que les points d'intérêt détectés sur des images à petites échelles se retrouvent sur des images à plus grande échelle et de la même manière avec des images multi-resolution



Notions d'analyse multi-échelle et multi-résolution

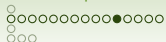
Est-ce que les points d'intérêt détectés sur des images à petites échelles se retrouvent sur des images à plus grande échelle et de la même manière avec des images multi-resolution

L'objectif de la multirésolution est d'obtenir des points différents dans les différentes échelles et résolution.

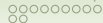
Donc, nous ne retrouvons pas les mêmes points.

Toutefois, en pratique, un même point d'intérêt peut être détecté sur quelques échelles consécutives et cet aspect peut même être utilisé pour sélectionner les points d'intérêt invariants sur plusieurs échelles.

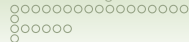
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

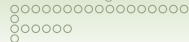
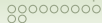
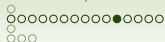


Détecteurs du second ordre



Notions d'analyse multi-échelle et multi-résolution

Comment les fusions multi-échelle sont elles réalisées ?



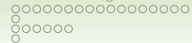
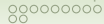
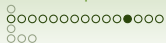
Notions d'analyse multi-échelle et multi-résolution

Comment les fusions multi-échelle sont elles réalisées ?

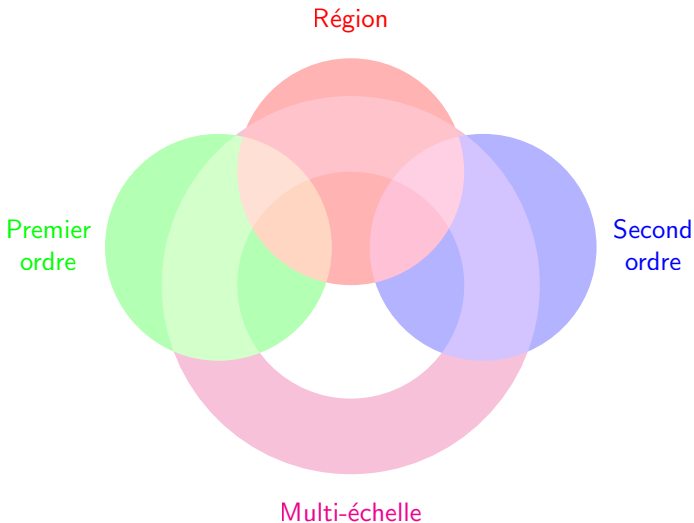
Cela dépend du détecteur.

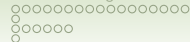
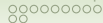
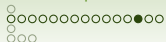
Dans SIFT, on va juste prendre toutes les détections à toutes les échelles.

Dans MSER, on va favoriser les détections qui ont lieu à plusieurs échelles successives.



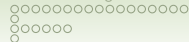
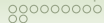
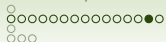
Classement des détecteurs





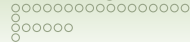
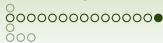
Classement des détecteurs

	Premier ordre	Région	Second ordre
MONO-ÉCHELLE	Moravec 1980, Harris 1988	SUSAN 1997, FAST 2006, MSER Matas2002, IBR 2004	Beaudet 1978, Kitchen 1982
MULTI-ÉCHELLE	Harris-laplace 2004	Kadir 2004, EBR 2004	Hessien- Laplace 2004, SIFT 2004, SURF 2008, CSS 1998, <i>MFC</i> 2017
			PCBR 2007



Invariance des détecteurs

Détecteur	Affines photométriques	Géométriques non affine	Affines géométriques	Changement d'échelle
Moravec, Harris				
Harris-laplace				×
SUSAN				
FAST				
MSER	×	×		
IBR	×	×	×	
Kadir	×	×	×	×
EBR	×	×	×	×
PCBR				×
Beaudet				
Kitchen1982				
Hessien-Laplace				×
SIFT		×		×
SURF				×
CSS	×	×	×	×
MFC				×

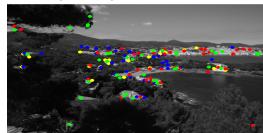


Complémentarité des détecteurs

Image



Multiples points d'intérêt



Harris



Beaudet

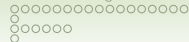
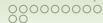
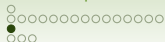


KR



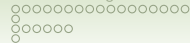
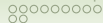
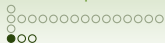
SURF





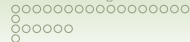
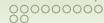
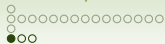
Notations

1. I : une image qui à chaque point/pixel $\mathbf{p}$ associe un niveau de gris noté $I(x, y)$
2. $I_x, I_y, I_{xx}, I_{yy}, I_{xy}$: images des dérivées premières et secondes de l'image
3. $I_x(x, y), \dots, I_{xy}(x, y)$: valeurs des dérivées premières et secondes au point $\mathbf{p}$
4. Pour simplifier, on notera I_x pour $I_x(x, y) \dots$



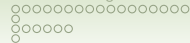
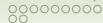
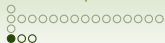
Construction d'une pyramide

- But : produire à partir d'une image, un ensemble d'images de résolution de plus en plus faibles



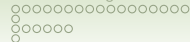
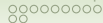
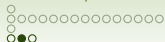
Construction d'une pyramide

- But : produire à partir d'une image, un ensemble d'images de résolution de plus en plus faibles
- Plusieurs techniques développées dans le domaine de la compression d'images



Construction d'une pyramide

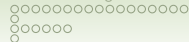
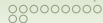
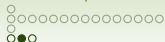
- But : produire à partir d'une image, un ensemble d'images de résolution de plus en plus faibles
- Plusieurs techniques développées dans le domaine de la compression d'images
- Différents types de pyramides : Pyramides **régulières** ou irrégulières



Construction d'une pyramide régulière

- Soit n la résolution courante à calculer (qui est plus faible que la résolution $n - 1$)

$$I^n(i, j) = \underset{(i', j') \in ZV(i, j)}{\text{reduction}} f_r(I^{n-1}(i', j')), \quad (3)$$

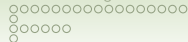
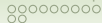
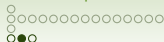


Construction d'une pyramide régulière

- Soit n la résolution courante à calculer (qui est plus faible que la résolution $n - 1$)

$$I^n(i, j) = \underset{(i', j') \in ZV(i, j)}{\text{reduction}} f_r(I^{n-1}(i', j')), \quad (3)$$

- $ZV(i, j) = \text{voisinage de réduction}$



Construction d'une pyramide régulière

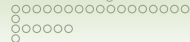
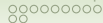
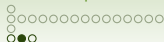
- Soit n la résolution courante à calculer (qui est plus faible que la résolution $n - 1$)

$$I^n(i, j) = \underset{(i', j') \in ZV(i, j)}{\text{reduction}} f_r(I^{n-1}(i', j')), \quad (3)$$

- $ZV(i, j)$ = voisinage de réduction

Fonction qui permet de retourner l'ensemble des pixels considérés au niveau $n - 1$ pour évaluer la valeur d'un pixel au niveau n

- f_r fonction à appliquer aux pixels pris en compte (pondération)



Construction d'une pyramide régulière

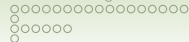
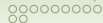
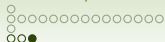
- Soit n la résolution courante à calculer (qui est plus faible que la résolution $n - 1$)

$$I^n(i, j) = \text{reduction}_{(i', j') \in \text{ZV}(i, j)} f_r(I^{n-1}(i', j')), \quad (3)$$

- $\text{ZV}(i, j) =$ voisinage de réduction

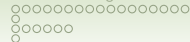
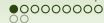
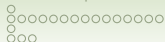
Fonction qui permet de retourner l'ensemble des pixels considérés au niveau $n - 1$ pour évaluer la valeur d'un pixel au niveau n

- f_r fonction à appliquer aux pixels pris en compte (pondération)
- reduction : fusion des valeurs des pixels voisins du niveau $n - 1$ pour obtenir le niveau n



Construction d'une pyramide régulière

- *Pyramides gaussiennes* – Lissage par un filtre gaussien, puis sous-échantillonnage, en récupérant un pixel sur quatre
- *Pyramides moyennes* – Moyenne sur un voisinage 2×2



Plan

Détection de points d'intérêt

Introduction

Notations

Construction d'une pyramide

DéTECTEURS du premier ordre

DéTECTEURS du premier ordre : vers le multi-échelle

DéTECTEURS de régions

Vers le multi-échelle : Détecteur de Kadir

Vers le multi-échelle : EBR, *Edge-Based Regions*

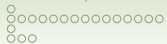
Vers le multi-échelle : PCBR, *Principal Curvature-Based Regions*

DéTECTEURS du second ordre

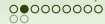
Vers le multi-échelle : SIFT *Scale Invariant Feature Transform*

Vers une variante accélérée de SIFT : SURF

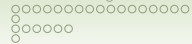
Détection de points d'intérêt



Détecteurs du premier ordre



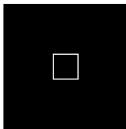
Détecteurs de régions



Détecteurs du second ordre

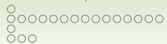


Moravec

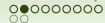


Point quelconque

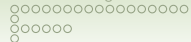
Détection de points d'intérêt



Détecteurs du premier ordre



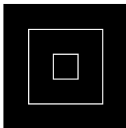
Détecteurs de régions



Détecteurs du second ordre

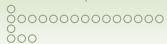


Moravec

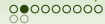


Point quelconque

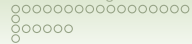
Détection de points d'intérêt



Détecteurs du premier ordre



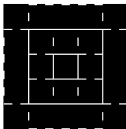
Détecteurs de régions



Détecteurs du second ordre

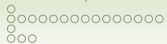


Moravec

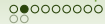


Point quelconque

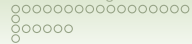
Détection de points d'intérêt



Détecteurs du premier ordre



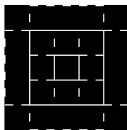
Détecteurs de régions



Détecteurs du second ordre

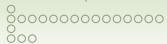


Moravec

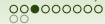


Point quelconque

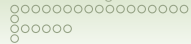
Détection de points d'intérêt



Détecteurs du premier ordre



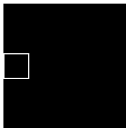
Détecteurs de régions



Détecteurs du second ordre

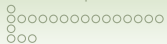


Moravec

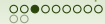


Point contour

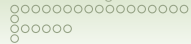
Détection de points d'intérêt



Détecteurs du premier ordre



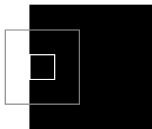
Détecteurs de régions



Détecteurs du second ordre

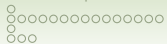


Moravec

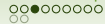


Point contour

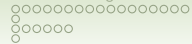
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



Détecteurs du second ordre

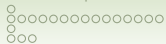


Moravec

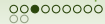


Point contour

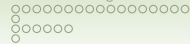
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



Détecteurs du second ordre



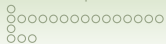
Moravec



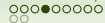
direction contour

Point contour

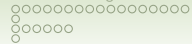
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



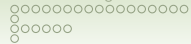
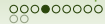
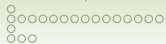
Détecteurs du second ordre



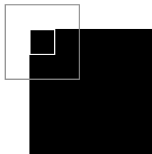
Moravec



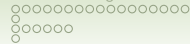
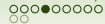
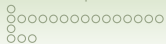
Point de type coin



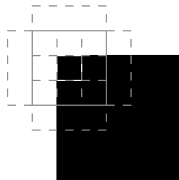
Moravec



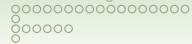
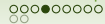
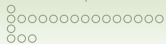
Point de type coin



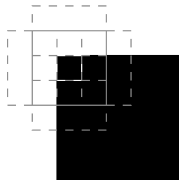
Moravec



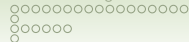
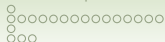
Point de type coin



Moravec



Point de type coin



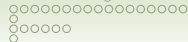
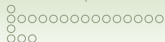
Moravec : définition

Soit (u, v) un déplacement, l'impact de ce déplacement se mesure par :

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

où :

1. w : poids
2. $E_{(u,v)}(x, y)$ = **moyenne du changement** si on déplace la fenêtre



Moravec : définition

Soit (u, v) un déplacement, l'impact de ce déplacement se mesure par :

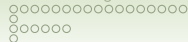
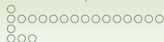
$$E_{(u,v)}(x, y) = \sum_{x,y} w(x, y) (I(x + u, y + v) - I(x, y))^2$$

où :

1. w : poids
2. $E_{(u,v)}(x, y)$ = **moyenne du changement** si on déplace la fenêtre

La réponse du détecteur est donc la valeur minimale parmi toutes les directions testées :

$$Mor(x, y) = \min_{(u,v)} E_{(u,v)}(x, y)$$



Moravec : définition

Soit (u, v) un déplacement, l'impact de ce déplacement se mesure par :

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

où :

1. w : poids

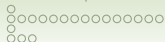
2. $E_{(u,v)}(x, y)$ = **moyenne du changement** si on déplace la fenêtre

La réponse du détecteur est donc la valeur minimale parmi toutes les directions testées :

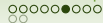
$$Mor(x, y) = \min_{(u,v)} E_{(u,v)}(x, y)$$

s'il existe une direction où le déplacement implique peu de changement alors, ce n'est pas un point d'intérêt

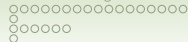
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

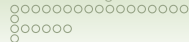
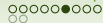
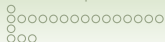


Détecteurs du second ordre



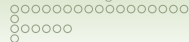
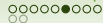
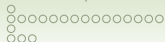
Harris : améliorations de Moravec

1. **La réponse est bruitée en raison des poids binaires** : utilisation des poids Gaussiens, de moyennes nulles



Harris : améliorations de Moravec

1. **La réponse est bruitée en raison des poids binaires** : utilisation des poids Gaussiens, de moyennes nulles
2. **La réponse du détecteur est anisotrope (dépendant de la direction)**



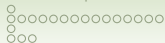
Harris : améliorations de Moravec

1. **La réponse est bruitée en raison des poids binaires** : utilisation des poids Gaussiens, de moyennes nulles
2. **La réponse du détecteur est anisotrope (dépendant de la direction)**

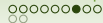
ET

3. **La réponse est trop forte/ambiguë au niveau des contours**

Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



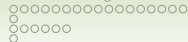
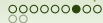
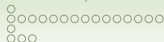
Détecteurs du second ordre



Harris : améliorations de Moravec

développement de Taylor à l'ordre 2 au voisinage de (x, y)

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



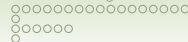
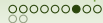
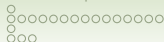
Harris : améliorations de Moravec

développement de Taylor à l'ordre 2 au voisinage de (x, y)

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

D'où, dans cette équation :

$$E_{(u,v)}^{\text{modifiée}}(x, y) = \sum_{x,y} w(x, y) \left(u \frac{\partial I}{\partial u} + v \frac{\partial I}{\partial v} + o(u^2, v^2) \right)^2$$



Harris : améliorations de Moravec

développement de Taylor à l'ordre 2 au voisinage de (x, y)

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

D'où, dans cette équation :

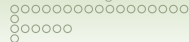
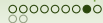
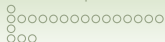
$$E_{(u,v)}^{\text{modifiée}}(x, y) = \sum_{x,y} w(x, y) \left(u \frac{\partial I}{\partial u} + v \frac{\partial I}{\partial v} + o(u^2, v^2) \right)^2$$

On obtient :

$$E_{\text{modifiée}} = Au^2 + 2Cuv + Bv^2,$$

avec $A = \frac{\partial^2 I}{\partial u^2} \otimes w$, $B = \frac{\partial^2 I}{\partial v^2} \otimes w$ et $C = \frac{\partial I}{\partial u} \frac{\partial I}{\partial v} \otimes w$

$\otimes$ = convolution

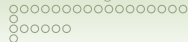
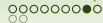
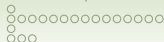


Harris : améliorations de Moravec

1. En repartant de l'équation (4)

$$E_{(u,v)}^{\text{modifiée}}(x, y) = (u, v)M(u, v)^T,$$

avec $M = \begin{pmatrix} A & C \\ C & B \end{pmatrix}.$



Harris : améliorations de Moravec

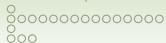
1. En repartant de l'équation (4)

$$E_{(u,v)}^{\text{modifiée}}(x, y) = (u, v)M(u, v)^T,$$

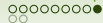
avec $M = \begin{pmatrix} A & C \\ C & B \end{pmatrix}.$

2. **Les valeurs propres de la matrice M correspondent aux courbures principales associées à E**
 - deux courbures faibles = région homogène
 - une des courbures est forte et l'autre est faible = un contour
 - deux courbures fortes = un coin, un point d'intérêt

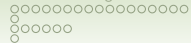
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



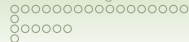
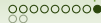
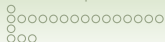
Détecteurs du second ordre



Harris : améliorations de Moravec

$$R(x, y) = \text{Det}(M) - k \text{Trace}(M)^2$$

avec $\text{Det}(M) = AB - C^2$ et $\text{Trace}(M) = A + B$

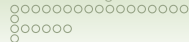
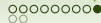
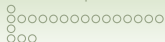


Harris : améliorations de Moravec

$$R(x, y) = \text{Det}(M) - k \text{Trace}(M)^2$$

avec $\text{Det}(M) = AB - C^2$ et $\text{Trace}(M) = A + B$

1. on s'est affranchi du choix de la direction !

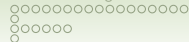
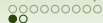
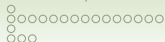


Harris : améliorations de Moravec

$$R(x, y) = \text{Det}(M) - k \text{Trace}(M)^2$$

avec $\text{Det}(M) = AB - C^2$ et $\text{Trace}(M) = A + B$

1. **on s'est affranchi du choix de la direction !**
2. Les valeurs de R sont
 - positives au voisinage d'un coin
 - négatives au voisinage d'un contour
 - faibles dans une région d'intensité constante

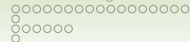
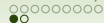
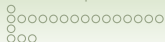


Harris multi-échelle : Harris-Laplace

1. Calcul de la réponse de Harris à plusieurs échelles

Extrema locaux estimées en (x, y, σ)

2. Affinement de la sélection des points d'intérêt dans l'espace échelle du laplacien



Harris multi-échelle : Harris-Laplace

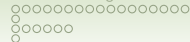
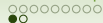
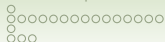
1. Calcul de la réponse de Harris à plusieurs échelles

Extrema locaux estimées en (x, y, σ)

2. Affinement de la sélection des points d'intérêt dans l'espace échelle du laplacien

3. Algorithme itératif appliqué pour chaque candidat retenu

4. Soient (i_0, j_0, σ_0) , les coordonnées et l'échelle initiale



Harris multi-échelle : Harris-Laplace

1. Calcul de la réponse de Harris à plusieurs échelles

Extrema locaux estimées en (x, y, σ)

2. Affinement de la sélection des points d'intérêt dans l'espace échelle du laplacien

3. Algorithme itératif appliqué pour chaque candidat retenu

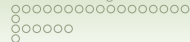
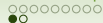
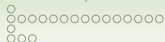
4. Soient (i_0, j_0, σ_0) , les coordonnées et l'échelle initiale

4.1 Recherche d'un extremum dans l'espace échelle du laplacien pour les échelles comprises dans $[c^{-1}\sigma_{l_k}; c\sigma_{l_k}]$,

4.2 Intervalle inter-échelle guidé par $c_2 < c$

Si un extremum n'est pas trouvé alors ce candidat est rejeté

Sinon, on note σ_u , l'échelle à laquelle cet extremum a été trouvé



Harris multi-échelle : Harris-Laplace

1. Calcul de la réponse de Harris à plusieurs échelles

Extrema locaux estimées en (x, y, σ)

2. Affinement de la sélection des points d'intérêt dans l'espace échelle du laplacien

3. Algorithme itératif appliqué pour chaque candidat retenu

4. Soient (i_0, j_0, σ_0) , les coordonnées et l'échelle initiale

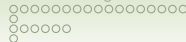
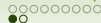
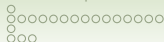
4.1 Recherche d'un extremum dans l'espace échelle du laplacien pour les échelles comprises dans $[c^{-1}\sigma_{l_k}; c\sigma_{l_k}]$,

4.2 Intervalle inter-échelle guidé par $c_2 < c$

Si un extremum n'est pas trouvé alors ce candidat est rejeté

Sinon, on note σ_u , l'échelle à laquelle cet extremum a été trouvé

4.3 Détection de la position (i_{k+1}, j_{k+1}) du maximum de la réponse de Harris le plus proche de (i_k, j_k) à l'échelle σ_u



Harris multi-échelle : Harris-Laplace

1. Calcul de la réponse de Harris à plusieurs échelles

Extrema locaux estimées en (x, y, σ)

2. Affinement de la sélection des points d'intérêt dans l'espace échelle du laplacien

3. Algorithme itératif appliqué pour chaque candidat retenu

4. Soient (i_0, j_0, σ_0) , les coordonnées et l'échelle initiale

4.1 Recherche d'un extremum dans l'espace échelle du laplacien pour les échelles comprises dans $[c^{-1}\sigma_{l_k}; c\sigma_{l_k}]$,

4.2 Intervalle inter-échelle guidé par $c_2 < c$

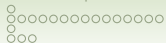
Si un extremum n'est pas trouvé alors ce candidat est rejeté

Sinon, on note σ_u , l'échelle à laquelle cet extremum a été trouvé

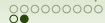
4.3 Détection de la position (i_{k+1}, j_{k+1}) du maximum de la réponse de Harris le plus proche de (i_k, j_k) à l'échelle σ_u

Nouvelle itération si $\sigma_k \neq \sigma_{k+1}$ et $(i_k, j_k) \neq (i_{k+1}, j_{k+1})$

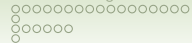
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

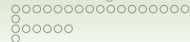
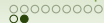
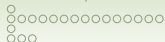


Détecteurs du second ordre



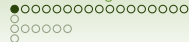
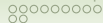
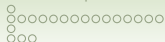
Remarques

- Une version moins coûteuse : seuillage de la réponse



Remarques

- Une version moins coûteuse : seuillage de la réponse
- Hessien-Laplace : même algorithme appliqué au détecteur de Beaudet (que nous verrons par la suite)



Plan

Détection de points d'intérêt

- Introduction

- Notations

- Construction d'une pyramide

DéTECTEURS du premier ordre

- DéTECTEURS du premier ordre : vers le multi-échelle

DéTECTEURS de régions

- Vers le multi-échelle : Détecteur de Kadir

- Vers le multi-échelle : EBR, *Edge-Based Regions*

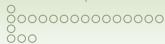
- Vers le multi-échelle : PCBR, *Principal Curvature-Based Regions*

DéTECTEURS du second ordre

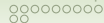
- Vers le multi-échelle : SIFT *Scale Invariant Feature Transform*

- Vers une variante accélérée de SIFT : SURF

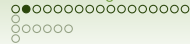
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

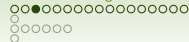
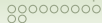
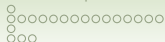


Détecteurs du second ordre



Détecteurs de régions

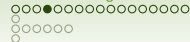
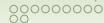
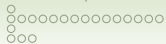
Les détecteurs de type régions ne détectent pas les coins ou des angles dans l'image. Ils considèrent le centre des régions détectées comme étant le point d'intérêt.



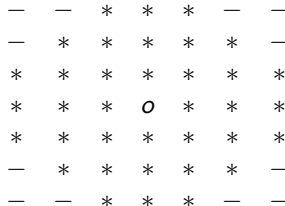
DéTECTEUR SUSAN (*Smallest UnivalUe Segment Assimilating Nucleus*)

Les principes

1. Notion de **noyau**, *nucleus* = pixel étudié
2. Notion de région d'intérêt autour de ce noyau = **masque** circulaire
3. Notion de sous-région, *segment*, autour du noyau, photométriquement proche du noyau
4. Forme et taille de cette sous-région = caractérise le point et indique sa particularité



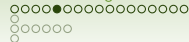
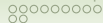
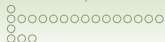
SUSAN : masque circulaire



o : pixel étudié

* : pixel pris en compte dans le voisinage

— : pixel non pris en compte dans le voisinage



SUSAN : définition de l'USAN

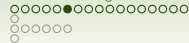
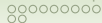
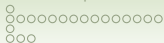
$$f : \mathbb{N}^2 \rightarrow \mathbb{R}$$

$$\mathbf{p} \rightarrow f(\mathbf{p}) = \sum_{\mathbf{p}' \in \mathcal{V}(\mathbf{p})} e^{-\left(\frac{I(\mathbf{p}) - I(\mathbf{p}')}{T_u}\right)^6}$$

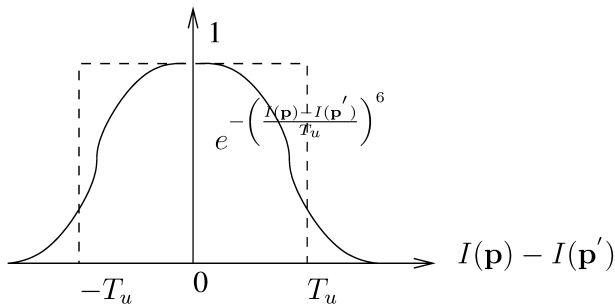
T_u correspond au contraste maximal autorisé

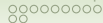
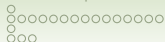
USAN = région d'intérêt/région d'appartenance du pixel étudié

Le barycentre de l'USAN = le centre



SUSAN : Comportement de la fonction de similarité



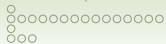


SUSAN : définition de la réponse

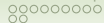
$$SU(\mathbf{p}) = \begin{cases} T_g - f(\mathbf{p}) & \text{si } f(\mathbf{p}) < T_g \\ 0 & \text{sinon} \end{cases}$$

T_g est un seuil géométrique.

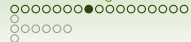
Détection de points d'intérêt



Détecteurs du premier ordre



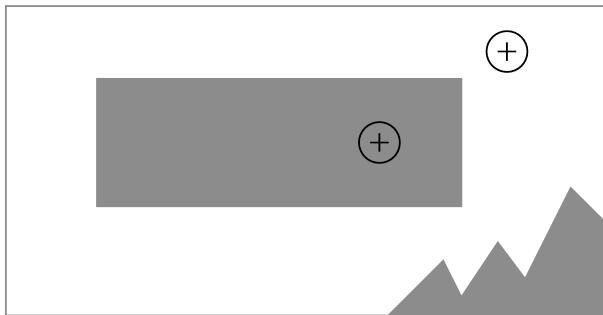
Détecteurs de régions

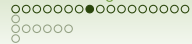
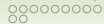
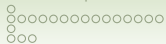


Détecteurs du second ordre

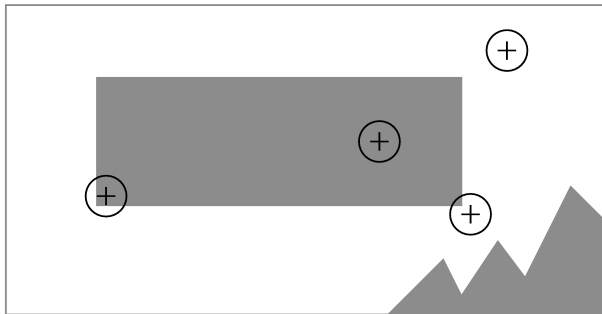


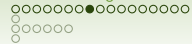
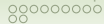
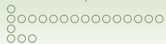
Analyse de la surface USAN



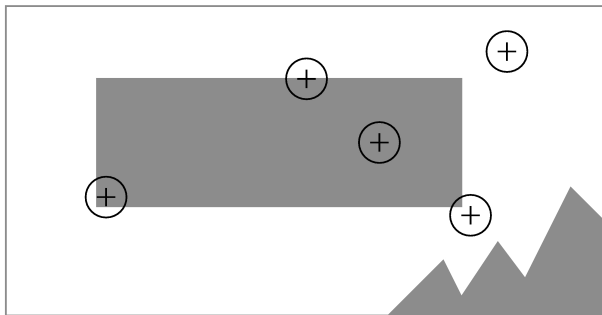


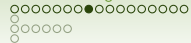
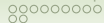
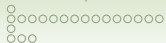
Analyse de la surface USAN



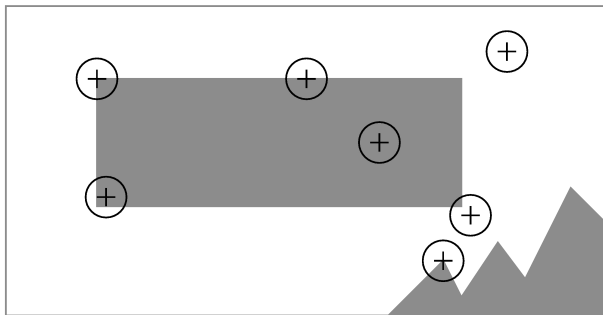


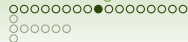
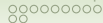
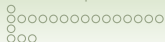
Analyse de la surface USAN





Analyse de la surface USAN

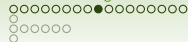
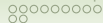
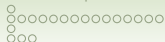




Comportement du détecteur

Comportement “idéal”

- Surface de l'USAN proche de la surface du voisinage = zone homogène
- Surface de l'USAN proche de la moitié de la surface du voisinage = contour
- Surface de l'USAN plus petite que la moitié de la surface du voisinage = point d'intérêt

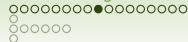
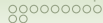
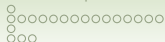


Comportement du détecteur

Comportement “idéal”

- Surface de l'USAN proche de la surface du voisinage = zone homogène
- Surface de l'USAN proche de la moitié de la surface du voisinage = contour
- Surface de l'USAN plus petite que la moitié de la surface du voisinage = point d'intérêt

Ainsi, souvent, $T_g = \frac{s_{max}}{2}$ où s_{max} est la surface maximale de l'USAN étant donnée la taille du voisinage considéré.



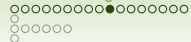
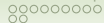
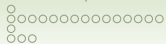
Comportement du détecteur

Comportement “idéal”

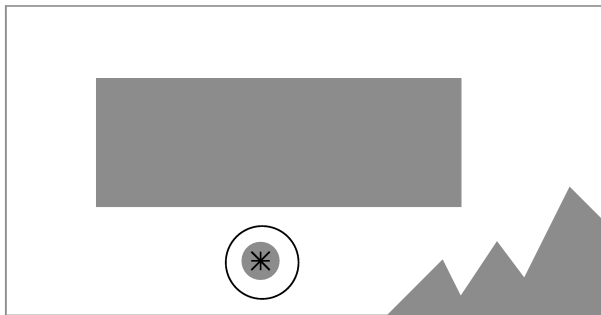
- Surface de l'USAN proche de la surface du voisinage = zone homogène
- Surface de l'USAN proche de la moitié de la surface du voisinage = contour
- Surface de l'USAN plus petite que la moitié de la surface du voisinage = point d'intérêt

Ainsi, souvent, $T_g = \frac{s_{max}}{2}$ où s_{max} est la surface maximale de l'USAN étant donnée la taille du voisinage considéré.

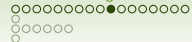
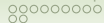
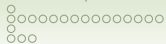
Bien sûr, il y a des cas où le respect de cette contrainte n'est pas corrélé au fait que le point étudié est un point d'intérêt ← nécessité d'ajouter des contraintes



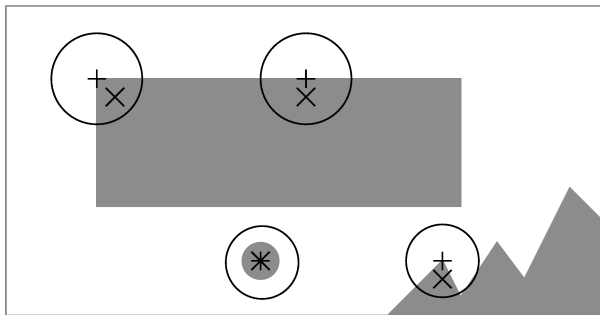
SUSAN : Contrainte du centre



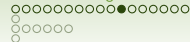
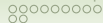
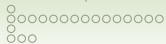
+ Noyaux
× Centres



SUSAN : Contrainte du centre

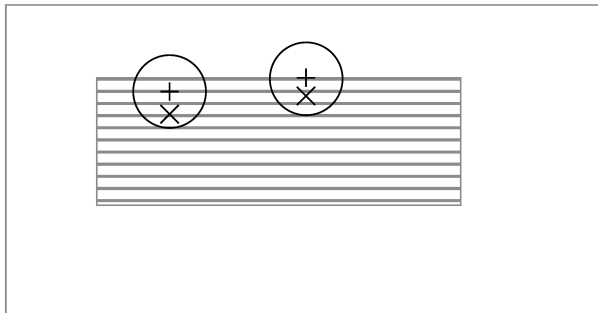


+ Noyaux
× Centres

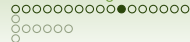
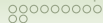
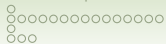


SUSAN : Contrainte de continuité

Quelle est la différence entre le noyau et le centre d'une région ?

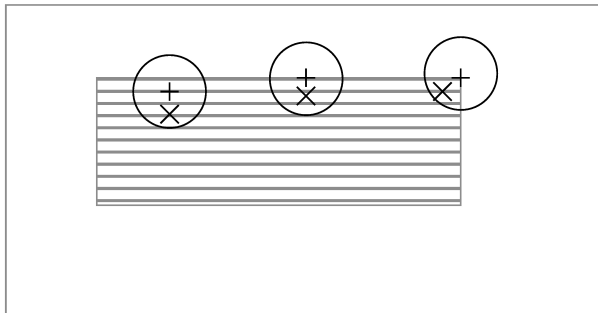


+ Noyaux
× Centres



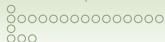
SUSAN : Contrainte de continuité

Quelle est la différence entre le noyau et le centre d'une région ?

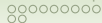


+ Noyaux
× Centres

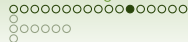
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

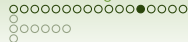
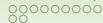
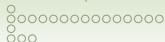


Détecteurs du second ordre



SUSAN : Conclusion

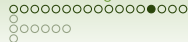
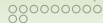
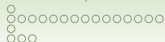
Plus on ajoute de contraintes, plus on élimine de “mauvais” points d’intérêt mais plus on risque d’éliminer de “bons” points d’intérêt !



SUSAN FAST, Features from Accelerated Segment Test

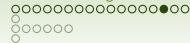
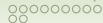
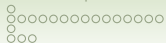
- Proche de SUSAN
- Zone circulaire : Algorithme du cercle de Bresenham de rayon 3
→ 16 points sur le périmètre de ce cercle
- Ensembles de pixels de même intensité qui forment une région connexe
- Vérification d'une de ces hypothèses, pour T à choisir :
 1. $\exists Np = 12$ pixels connexes tels que $I(x', y') < I(x, y) - T$
 2. $\exists Np = 12$ pixels connexes tels que $I(x', y') > I(x, y) + T$
- Version plus rapide avec 4 pixels et vérification de 3/4 pixels
- Inconvénient : comment choisir les n pixels significatives ?

Solution **Apprentissage**



Maximally Stable Extremal Regions (MSER)

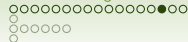
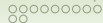
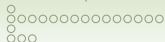
- **Principe** : intensité varie rapidement aux bords des objets
régions = composantes connexes de l'image seuillée (ligne de partage des eaux)
- **Propriétés** : invariant aux transformations monotones d'intensité et aux transformations géométriques homographiques ou non linéaires mais continues
- **Inconvénient** : sensible au flou (intensité sur les frontières mal définie)



Intensity Based Regions (IBR)

- **Étapes**

1. Détection des extrema locaux de l'intensité lumineuse de l'image /
2. Caractérisation de la région autour de ces extrema



Intensity Based Regions (IBR)

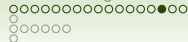
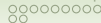
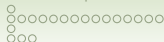
- Étapes

1. Détection des extrema locaux de l'intensité lumineuse de l'image I
2. Caractérisation de la région autour de ces extrema

- Caractérisation

1. Fonction d'intensité le long des rayons issus de cet extremum

$$f_I(t) = \frac{|I(t) - I(0)|}{\max(d, \frac{1}{t} \int_0^t |I(x) - I(0)| dx)} \quad (4)$$



Intensity Based Regions (IBR)

- Étapes

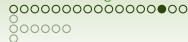
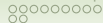
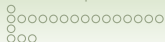
1. Détection des extrema locaux de l'intensité lumineuse de l'image I
2. Caractérisation de la région autour de ces extrema

- Caractérisation

1. Fonction d'intensité le long des rayons issus de cet extremum

$$f_I(t) = \frac{|I(t) - I(0)|}{\max(d, \frac{1}{t} \int_0^t |I(x) - I(0)| dx)} \quad (4)$$

2. Extremum chaînés entre eux



Intensity Based Regions (IBR)

- Étapes

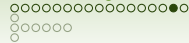
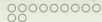
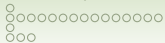
- Détection des extrema locaux de l'intensité lumineuse de l'image I
- Caractérisation de la région autour de ces extrema

- Caractérisation

- Fonction d'intensité le long des rayons issus de cet extremum

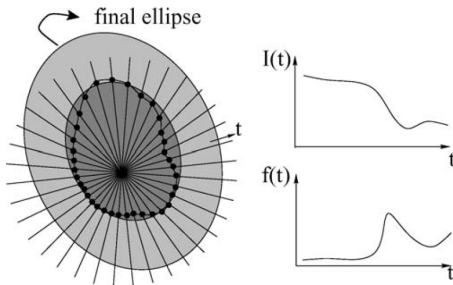
$$f_I(t) = \frac{|I(t) - I(0)|}{\max(d, \frac{1}{t} \int_0^t |I(x) - I(0)| dx)} \quad (4)$$

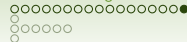
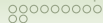
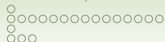
- Extremum chaînés entre eux
- Détermination de l'ellipse



Intensity Based Regions (IBR)

Illustration de la région

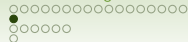
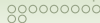
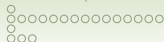




Intensity Based Regions (IBR)

Analyse

- Notion de blob : proche de l'approche définie dans SUSAN
- Taille de l'ellipse doublée
- Robuste sur les documents imprimés
- Invariant par les transformations affines géométriques et photométriques



DéTECTEUR de Kadir

- Entropie de la distribution des valeurs des pixels au voisinage du point considéré
- = Plus le voisinage est désordonné plus il est discriminant

$$\mathcal{H} : \mathbb{N}^2 \rightarrow \mathbb{R}$$

$$\mathbf{p} \mapsto - \sum_{k,l} p(\mathbf{p}) \log(p(\mathbf{p})) \quad (5)$$

avec $(k, l) \in \mathbb{N}^2$

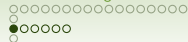
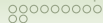
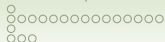
et $(k - i)^2 + (l - j)^2 < r^2$ où r est le rayon du voisinage considéré

et probabilité $p(\mathbf{p})$ = distribution des valeurs des pixels du voisinage

- La réponse est donnée par :

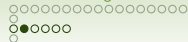
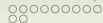
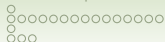
$$\text{KA}(\mathbf{p}) = \mathcal{H}(\mathbf{p}) \frac{\sigma^2}{2\sigma - 1} \sum_{k,l} \left| \frac{\partial p(\mathbf{p})}{\partial \sigma} \right|. \quad (6)$$

- Invariant aux transformations photométriques (biais), géométriques affines
- = indépendant du point de vue



Edge-Based Regions (EBR)

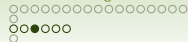
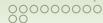
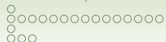
- Contours stables aux changement de point de vue et de conditions d'éclairage
- **Conséquence** : Utilisation de la géométrie locale : droites et parallélogrammes



Edge-Based Regions (EBR)

Étapes

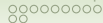
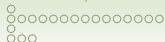
1. Coins localisés avec Harris
2. Contours détectés avec Deriche
3. Pour chaque coin, analyse des contours passant par ce coin



Edge-Based Regions (EBR)

Principe

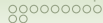
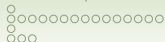
- Un coin p et deux points $p_1(l_1)$ et $p_2(l_2)$ se déplaçant le long des contours partant de ce coin



Edge-Based Regions (EBR)

Principe

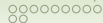
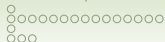
- Un coin p et deux points $p_1(l_1)$ et $p_2(l_2)$ se déplaçant le long des contours partant de ce coin
- Distance curviligne $l_{1,2}$ de chaque point $p_{1,2}$ au coin p = surface entre le contour et la droite joignant le coin et le point $p_{1,2}$



Edge-Based Regions (EBR)

Principe

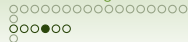
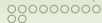
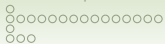
- Un coin p et deux points $p_1(l_1)$ et $p_2(l_2)$ se déplaçant le long des contours partant de ce coin
- Distance curviligne $l_{1,2}$ de chaque point $p_{1,2}$ au coin p = surface entre le contour et la droite joignant le coin et le point $p_{1,2}$
- Distance invariante aux transformations affines



Edge-Based Regions (EBR)

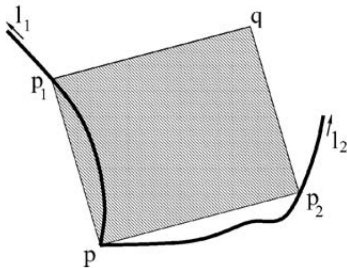
Principe

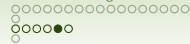
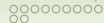
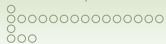
- Un coin p et deux points $p_1(l_1)$ et $p_2(l_2)$ se déplaçant le long des contours partant de ce coin
- Distance curviligne $l_{1,2}$ de chaque point $p_{1,2}$ au coin p = surface entre le contour et la droite joignant le coin et le point $p_{1,2}$
- Distance invariante aux transformations affines
- Un coin p et deux points $p_1(l_1)$ et $p_2(l_2)$ = une famille de parallélogrammes



Edge-Based Regions (EBR)

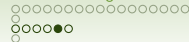
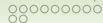
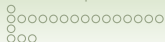
Analyse des coins





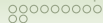
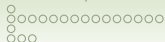
Edge-Based Regions (EBR)

- Dans cette famille de parallélogrammes, sélection de celui qui a une valeur minimale d'une fonction photométrique



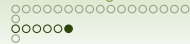
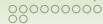
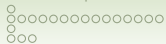
Edge-Based Regions (EBR)

- Dans cette famille de parallélogrammes, sélection de celui qui a une valeur minimale d'une fonction photométrique
- Fonction : distance entre le centre de **gravité et les diagonales du parallélogramme**



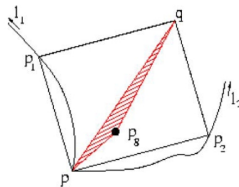
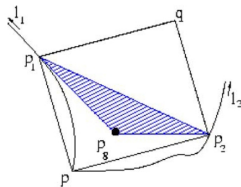
Edge-Based Regions (EBR)

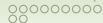
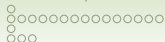
- Dans cette famille de parallélogrammes, sélection de celui qui a une valeur minimale d'une fonction photométrique
- Fonction : distance entre le centre de **gravité et les diagonales du parallélogramme**
- Premiers pas vers une réponse invariante à l'échelle ...



Edge-Based Regions (EBR)

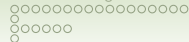
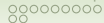
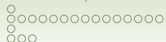
Illustration du critère associé





PCBR Principal Curvature-Based Regions

- Estimer les courbures principales estimées à différentes échelles
- Améliorer cette détection en utilisant
 1. des opérateurs morphologiques pour fermer les contours
 2. un seuillage par hystérésis des courbures
 3. la conservation uniquement si la courbure est élevée sur différentes échelles
- Détecter des lignes de partages des eaux (à partir des courbures)



Plan

Détection de points d'intérêt

Introduction

Notations

Construction d'une pyramide

DéTECTEURS du premier ordre

DéTECTEURS du premier ordre : vers le multi-échelle

DéTECTEURS de régions

Vers le multi-échelle : Détecteur de Kadir

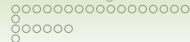
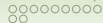
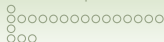
Vers le multi-échelle : EBR, *Edge-Based Regions*

Vers le multi-échelle : PCBR, *Principal Curvature-Based Regions*

DéTECTEURS du second ordre

Vers le multi-échelle : SIFT *Scale Invariant Feature Transform*

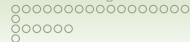
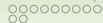
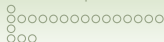
Vers une variante accélérée de SIFT : SURF



DéTECTEUR de Beaudet

Courbure de la surface : minimiser le déterminant de la matrice hessienne

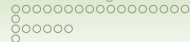
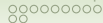
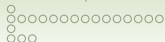
$$\text{Beaudet}(x, y) = \det \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix} = I_{xx}I_{yy} - I_{xy}^2$$



DéTECTEUR de Kitchen et Rosenfeld

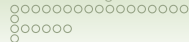
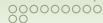
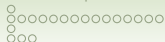
Courbure du contour pondéré par la norme du gradient

$$\text{KR}(x, y) = \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{I_x^2 + I_y^2}$$



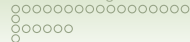
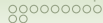
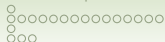
Courbure et espace d'échelle : *Curvature Scale Space* (CSS)

- **Principe :** recherche les extrema de la fonction courbure le long des contours en utilisant une approche multi-échelle



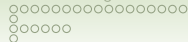
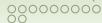
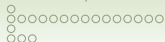
Courbure et espace d'échelle : *Curvature Scale Space* (CSS)

- **Principe :** recherche les extrema de la fonction courbure le long des contours en utilisant une approche multi-échelle
- **Étapes**
 1. Détection par Canny
 2. Fusion des contours proches
 3. Calcul de la courbure des contours avec lissage par une gaussienne
 4. Détection des extrema dans cet espace



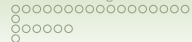
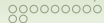
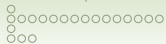
Courbure et espace d'échelle : *Curvature Scale Space* (CSS)

- **Principe :** recherche les extrema de la fonction courbure le long des contours en utilisant une approche multi-échelle
- **Étapes**
 1. Détection par Canny
 2. Fusion des contours proches
 3. Calcul de la courbure des contours avec lissage par une gaussienne
 4. Détection des extrema dans cet espace
- **Variante :** Distinction entre contours courts et longs (détermination automatique du seuil de décision)



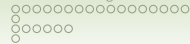
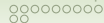
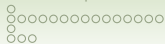
Courbure et espace d'échelle : *Curvature Scale Space* (CSS)

- **Principe** : recherche les extrema de la fonction courbure le long des contours en utilisant une approche multi-échelle
- **Étapes**
 1. Détection par Canny
 2. Fusion des contours proches
 3. Calcul de la courbure des contours avec lissage par une gaussienne
 4. Détection des extrema dans cet espace
- **Variante** : Distinction entre contours courts et longs (détermination automatique du seuil de décision)
- **Invariance** : Transformations affines géométriques et photométriques



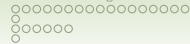
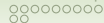
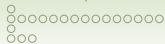
SIFT *Scale Invariant Feature Transform*

1. Détection des extremums en **échelle** et en **espace**



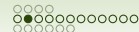
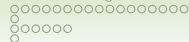
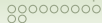
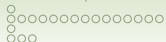
SIFT *Scale Invariant Feature Transform*

1. Détection des extremums en **échelle et en espace**
2. Caractérisation par une **orientation principale** = gradient dominant dans le voisinage du point considéré, avec prise en compte du bruit



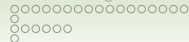
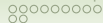
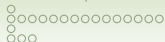
SIFT *Scale Invariant Feature Transform*

1. Détection des extremums en **échelle et en espace**
2. Caractérisation par une **orientation principale** = gradient dominant dans le voisinage du point considéré, avec prise en compte du bruit
3. Ajout d'un descripteur associé prenant en compte un **histogramme local de ces orientations**



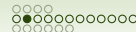
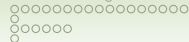
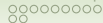
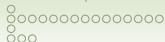
Propriétés intéressantes

1. **Détecteur complet** : un descripteur associé permettant mettre en correspondance de manière robuste



Propriétés intéressantes

1. **Détecteur complet** : un descripteur associé permettant mettre en correspondance de manière robuste
2. De nombreuses propriétés
 - 2.1 Localisation précise
 - 2.2 robustes aux occultations et aux bruits
 - 2.3 invariance aux changements de luminosité
 - 2.4 robustesse aux changements de points de vue

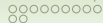
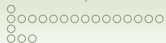


Propriétés intéressantes

1. **Détecteur complet** : un descripteur associé permettant mettre en correspondance de manière robuste
2. De nombreuses propriétés
 - 2.1 Localisation précise
 - 2.2 robustes aux occultations et aux bruits
 - 2.3 invariance aux changements de luminosité
 - 2.4 robustesse aux changements de points de vue

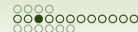
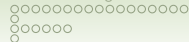
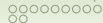
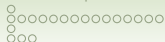
MAIS

3. **Non invariant aux transformations affines**



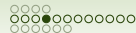
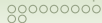
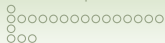
Étapes

1. Détection des extremums en échelle et en espace
2. Localisation des points
3. Affectation d'une orientation
4. Estimation des descripteurs



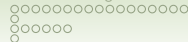
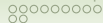
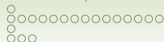
Étapes

1. Détection des extremums en échelle et en espace : Détection
2. Localisation des points : Affinement de la détection
3. Affectation d'une orientation : Caractérisation
4. Estimation des descripteurs : Pour la mise en correspondance



Détection des extremums en échelle et en espace

- Approche par **filtres en cascade**

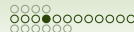
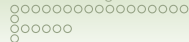
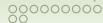
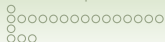


Détection des extremums en échelle et en espace

- Approche par **filtres en cascade**
- Estimation de différences de gaussiennes = **approximation du laplacien**

$$\begin{aligned} DoG(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

où $G(x, y, \sigma)$ est la fonction gaussienne



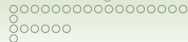
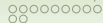
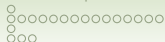
Détection des extremums en échelle et en espace

- Approche par **filtres en cascade**
- Estimation de différences de gaussiennes = **approximation du laplacien**

$$\begin{aligned} DoG(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

où $G(x, y, \sigma)$ est la fonction gaussienne

- Pour des valeurs successives de σ : laplacien normalisé en échelle

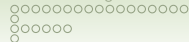
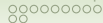
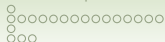


Détection des extremums en échelle et en espace

- Définition d'une octave** = Ensemble d'échelles
ensemble des images lissées par un filtre gaussien à différentes échelles pour une résolution donnée

Pour passer d'une octave à l'autre = double la valeur de σ

Donc : si on souhaite n échelles par octave, alors il faut fixer
 $k = 2^{1/n}$



Détection des extremums en échelle et en espace

- Définition d'une octave** = Ensemble d'échelles
ensemble des images lissées par un filtre gaussien à différentes échelles pour une résolution donnée

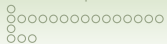
Pour passer d'une octave à l'autre = double la valeur de σ

Donc : si on souhaite n échelles par octave, alors il faut fixer
 $k = 2^{1/n}$

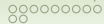
- maxima/minima locaux pour chaque octave**

Voisinage 3×3 + échelles inférieure et supérieure

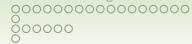
Détection de points d'intérêt



Détecteurs du premier ordre



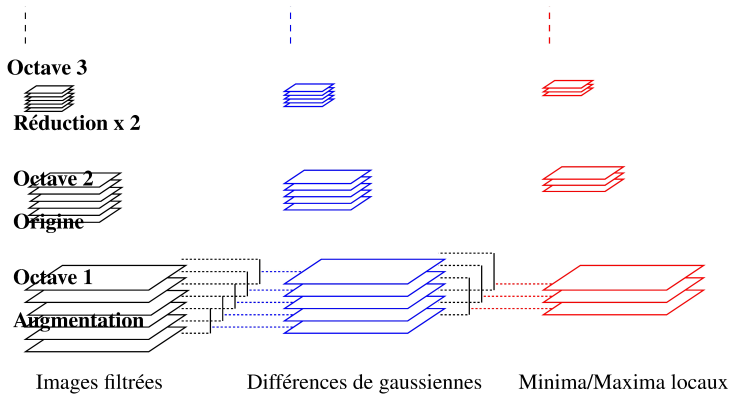
Détecteurs de régions



Détecteurs du second ordre



Détection



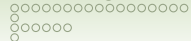
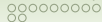
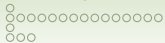
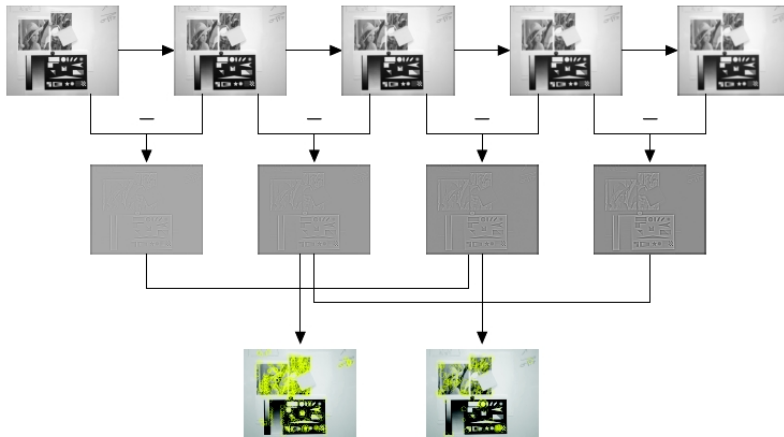
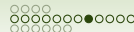
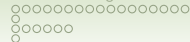
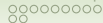
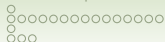


Illustration sur une octave

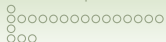




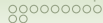
Algorithme complet de la détection (sans affinement)

1. Calculer une pyramide gaussienne d'images
2. Pour chaque octave o (chaque image de la pyramide) faire
 - Pour chaque valeur de k faire
 - Calculer en chaque point (x, y) , $G(x, y, k\sigma)$
3. Pour chaque octave o faire
 - Pour chaque échelle $s \in [2, n]$ faire
 - Calculer en chaque point (x, y) , $DoG(x, y, \sigma)$
4. Pour chaque octave o faire
 - Pour chaque échelle $s \in [2, n - 1]$ faire
 - Extraire les minima/maxima locaux

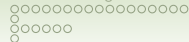
Détection de points d'intérêt



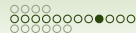
Détecteurs du premier ordre



Détecteurs de régions



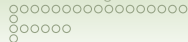
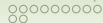
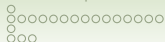
Détecteurs du second ordre



Affinement et correction de la localisation des points

1. Calcul de la position au sous-pixel

Estimation du décalage et re-calage si nécessaire



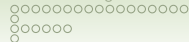
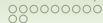
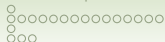
Affinement et correction de la localisation des points

1. Calcul de la position au sous-pixel

Estimation du décalage et re-calage si nécessaire

2. Suppression des points faiblement contrastés

Différence (après filtrage) entre le point initial et le point re-localisé



Affinement et correction de la localisation des points

1. Calcul de la position au sous-pixel

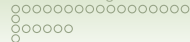
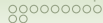
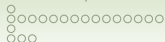
Estimation du décalage et re-calage si nécessaire

2. Suppression des points faiblement contrastés

Différence (après filtrage) entre le point initial et le point re-localisé

3. Suppression des points contours

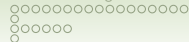
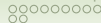
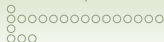
Utilisation de la matrice Hessienne (comme Harris)



Position sub-pixellique ou relocalisation

En quoi consiste la relocalisation lorsque la détection n'est pas optimale au cours du test de la répétabilité ?

- **But :** Contrecarrer les effets des lissages

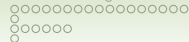
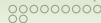
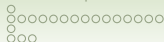


Position sub-pixellique ou relocalisation

En quoi consiste la relocalisation lorsque la détection n'est pas optimale au cours du test de la répétabilité ?

- **But :** Contrecarrer les effets des lissages
- **Position calculée au sous-pixel :** Développement de Taylor à l'ordre 2

$$D(\hat{\mathbf{p}}) = D(\mathbf{p}) + \frac{\partial D(\mathbf{p})^T}{\partial \mathbf{p}} + \frac{1}{2} \mathbf{p}^T \frac{\partial^2 D(\mathbf{p})}{\partial \mathbf{p}^2} \mathbf{p}, \quad (7)$$



Position sub-pixellique ou relocalisation

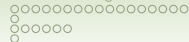
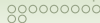
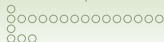
En quoi consiste la relocalisation lorsque la détection n'est pas optimale au cours du test de la répétabilité ?

- **But :** Contrecarrer les effets des lissages
- **Position calculée au sous-pixel :** Développement de Taylor à l'ordre 2

$$D(\hat{\mathbf{p}}) = D(\mathbf{p}) + \frac{\partial D(\mathbf{p})^T}{\partial \mathbf{p}} + \frac{1}{2} \mathbf{p}^T \frac{\partial^2 D(\mathbf{p})}{\partial \mathbf{p}^2} \mathbf{p}, \quad (7)$$

- Localisation précise de l'extremum = passage par zéro de la dérivée suivant $\mathbf{p}$

$$\hat{\mathbf{p}} = - \frac{\partial^2 D(\mathbf{p})^{-1}}{\partial \mathbf{p}^2} \frac{\partial D(\mathbf{p})}{\partial \mathbf{p}}. \quad (8)$$



Position sub-pixellique ou relocalisation

En quoi consiste la relocalisation lorsque la détection n'est pas optimale au cours du test de la répétabilité ?

- **But :** Contrecarrer les effets des lissages
- **Position calculée au sous-pixel :** Développement de Taylor à l'ordre 2

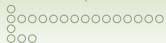
$$D(\hat{\mathbf{p}}) = D(\mathbf{p}) + \frac{\partial D(\mathbf{p})^T}{\partial \mathbf{p}} + \frac{1}{2} \mathbf{p}^T \frac{\partial^2 D(\mathbf{p})}{\partial \mathbf{p}^2} \mathbf{p}, \quad (7)$$

- Localisation précise de l'extremum = passage par zéro de la dérivée suivant $\mathbf{p}$

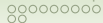
$$\hat{\mathbf{p}} = - \frac{\partial^2 D(\mathbf{p})^{-1}}{\partial \mathbf{p}^2} \frac{\partial D(\mathbf{p})}{\partial \mathbf{p}}. \quad (8)$$

- Si le décalage est plus grand que 0.5 dans chaque direction = relocalisation

Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

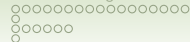
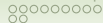
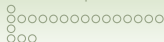


Détecteurs du second ordre



Filtrage des minima/maxima locaux

Première contrainte pour la sélection des minima/maxima locaux
Élimination des points dont le contraste est trop faible

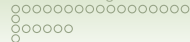
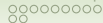
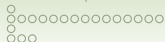


Filtrage des minima/maxima locaux

Première contrainte pour la sélection des minima/maxima locaux
Élimination des points dont le contraste est trop faible

- Si on utilise :

$$D(\hat{\mathbf{p}}) = D(\mathbf{p}) + \frac{1}{2} \frac{\partial D(\mathbf{p})^T}{\partial \mathbf{p}} \hat{\mathbf{p}} \quad (9)$$



Filtrage des minima/maxima locaux

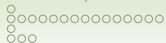
Première contrainte pour la sélection des minima/maxima locaux Élimination des points dont le contraste est trop faible

- Si on utilise :

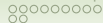
$$D(\hat{\mathbf{p}}) = D(\mathbf{p}) + \frac{1}{2} \frac{\partial D(\mathbf{p})^T}{\partial \mathbf{p}} \hat{\mathbf{p}} \quad (9)$$

- Expérimentalement : élimination des points tels que $|D(\hat{\mathbf{p}})| < 0.03$
En supposant une variation entre 0 et 1

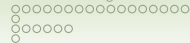
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions

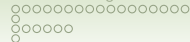
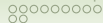
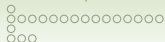


Détecteurs du second ordre



Filtrage des minima/maxima locaux

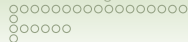
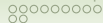
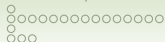
Deuxième contrainte pour la sélection des minima/maxima locaux
Points sur les contours



Filtrage des minima/maxima locaux

Deuxième contrainte pour la sélection des minima/maxima locaux Points sur les contours

- Même chose que pour l'approche de Harris



Filtrage des minima/maxima locaux

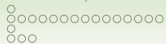
Deuxième contrainte pour la sélection des minima/maxima locaux Points sur les contours

- Même chose que pour l'approche de Harris
- Seuls les points où le rapport entre la plus grande et la plus petite valeur propre est inférieur à un seuil r sont conservés

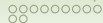
$$\frac{\text{Trace}(M)^2}{\text{Det}(M)} < \frac{(r+1)^2}{r} \quad (10)$$

- $r = 10$

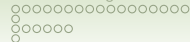
Détection de points d'intérêt



Détecteurs du premier ordre



Détecteurs de régions



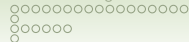
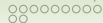
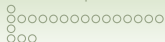
Détecteurs du second ordre



Speeded Up Robust Features, SURF

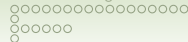
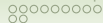
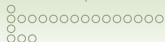
- **Intérêt :** Multi-échelle moins coûteux mais équivalent à SIFT

- **Intérêt :** Multi-échelle moins coûteux mais équivalent à SIFT
- **Réponse :** approximation du déterminant de la matrice hessienne à plusieurs échelles



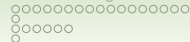
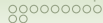
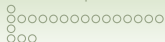
Speeded Up Robust Features, SURF

- **Intérêt** : Multi-échelle moins coûteux mais équivalent à SIFT
- **Réponse** : approximation du déterminant de la matrice hessienne à plusieurs échelles
- **Outils**
 - Lissage gaussien réalisé par des **box filter**
 - Utilisation des **images intégrales**

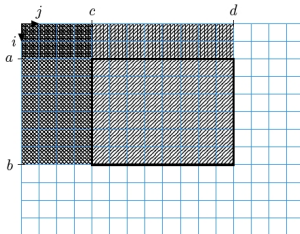


Speeded Up Robust Features, SURF

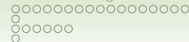
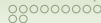
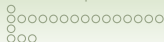
- **Intérêt** : Multi-échelle moins coûteux mais équivalent à SIFT
 - **Réponse** : approximation du déterminant de la matrice hessienne à plusieurs échelles
 - **Outils**
 - Lissage gaussien réalisé par des **box filter**
 - Utilisation des **images intégrales**
 - **En pratique**
 - Calcul par convolution est déjà une approximation
- MAIS** Résultats exploitables
- Utilisation d'une implémentation moins coûteuses



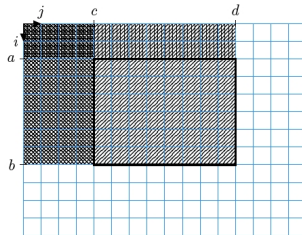
Notion d'image intégrale



$$I_{int}(i, j) = \sum_{k=0}^{i-1} \sum_{l=0}^{j-1} I(k, l). \quad (11)$$



Notion d'image intégrale

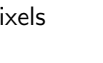


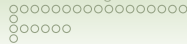
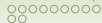
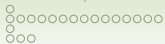
$$I_{int}(i, j) = \sum_{k=0}^{i-1} \sum_{l=0}^{j-1} I(k, l). \quad (11)$$

Permet de rendre rapide le calcul de sommes dans un rectangle donné

$$\sum_{i=a}^b \sum_{j=c}^d I(i, j) = I_{int}(b, d) - I_{int}(a, d) - I_{int}(b, c) + I_{int}(a, c). \quad (12)$$

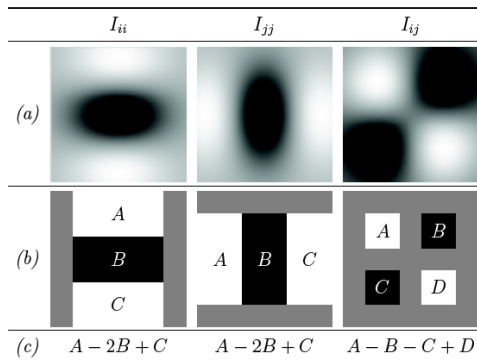
- pixels



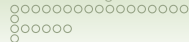
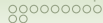
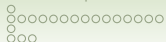


Box filtering

- Masque des dérivées = combinaisons de sommes de valeurs de pixels situés dans différents rectangles (*box*)

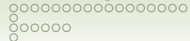
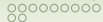
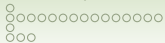


- L'utilisation de l'image intégrale réduit le nombre d'opérations arithmétiques

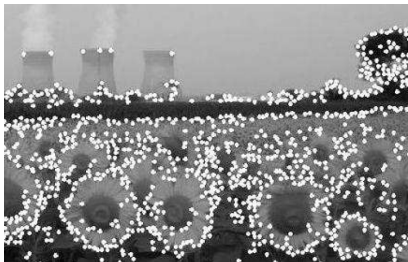


BRISK

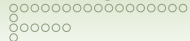
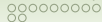
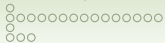
- Considération d'un voisinage circulaire
- Sélection d'un ensemble de points bien répartis dans ce voisinage
- Comparaison des intensités de ces points avec le point étudié
- Le descripteur = chaîne de bits indiquant le signe des comparaisons



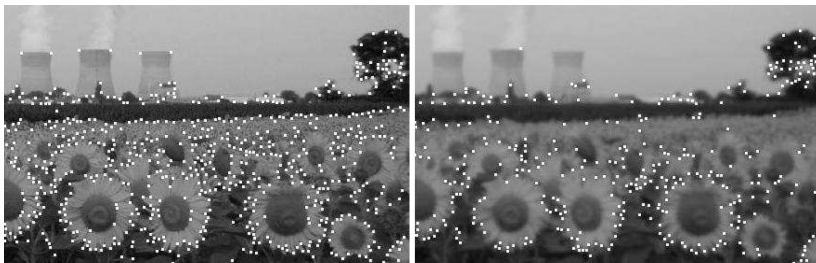
Comparaison des différents détecteurs



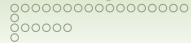
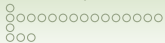
Beaudet



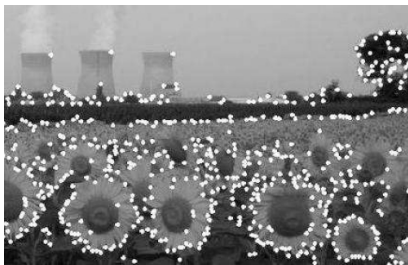
Comparaison des différents détecteurs



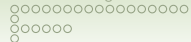
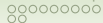
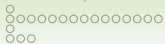
SUSAN



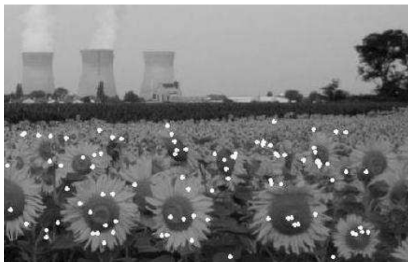
Comparaison des différents détecteurs



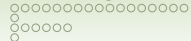
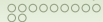
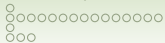
Harris



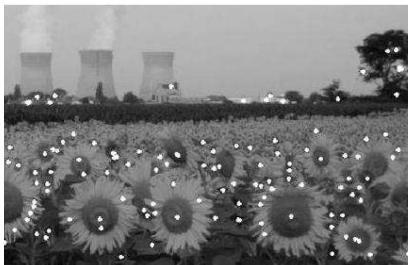
Comparaison des différents détecteurs



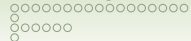
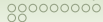
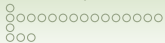
EBR



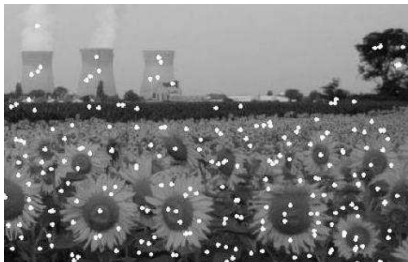
Comparaison des différents détecteurs



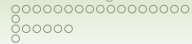
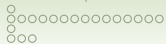
MSER



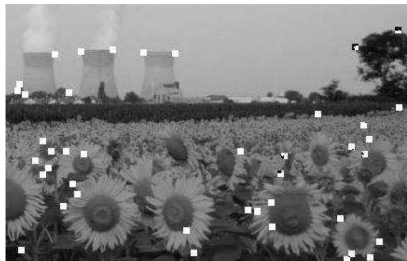
Comparaison des différents détecteurs



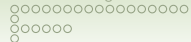
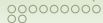
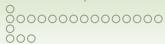
IBR



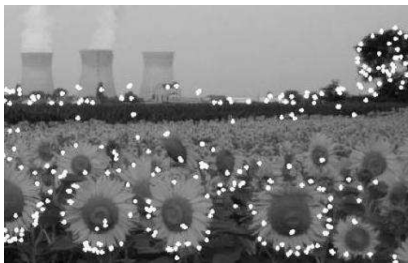
Comparaison des différents détecteurs



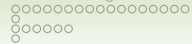
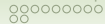
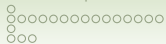
CSS



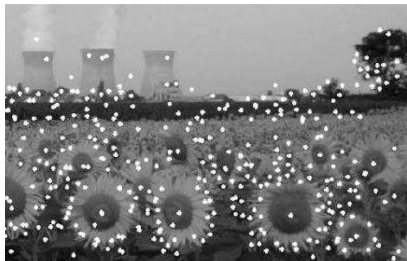
Comparaison des différents détecteurs



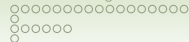
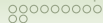
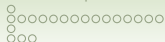
SIFT



Comparaison des différents détecteurs



SURF



Est-ce qu'il y a des détecteurs plus adaptés pour ensuite faire correspondre des points d'intérêt ou est-ce indépendant? Par exemple SIFT ?

- *Certains détecteurs n'ont pas de descripteur et/ou méthode d'appariement spécifiques associés.*
- *On peut donc choisir toutes les techniques que l'on souhaite parmi celles présentées dans la partie IV, sur la mise en correspondance.*