Bases du traitement des images

Détecter un objet

Dufrenois Franck
LISIC

Université du littoral

Objectif

Introduction

- Détecter une structure image: voiture – personne - objet dans une scène
- Utiliser une mesure de similarité entre l'image et une structure référence
- Détection: rechercher le minimum ou le maximum de cette mesure

Détection d'objets

Les mesures de similarité

- La mesure de corrélation
- La somme des différences au carré: SSD
- La mesure de Bhattacharya

Détection d’objets

Rappel: corrélation 1D

Soient deux signaux x(t) et y(t),mesurer la corrélation entre les deux consiste à réaliser:

$$\boxed{\begin{aligned}\Gamma_{xy}(\tau) = (x \circ y)(\tau) &= \int_{-\infty}^{+\infty} x(u) y^*(u+\tau) du \\ &= \int_{-\infty}^{+\infty} x(u) y^*(u-\tau) du = x(\tau) \circ y(\tau)\end{aligned}}$$

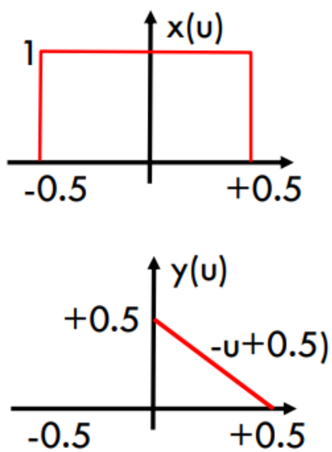
$$\boxed{\Gamma_{xy}(\tau)} = \int_{-\infty}^{+\infty} x(u) y^*(u-\tau) du \overset{t=u-\tau}{=} \int_{-\infty}^{+\infty} x(t+\tau) y^*(t) dt = \left(\int_{-\infty}^{+\infty} y\ (t) x^*(t+\tau) dt\right)^* = \boxed{\Gamma_{yx}(-\tau)^*}$$

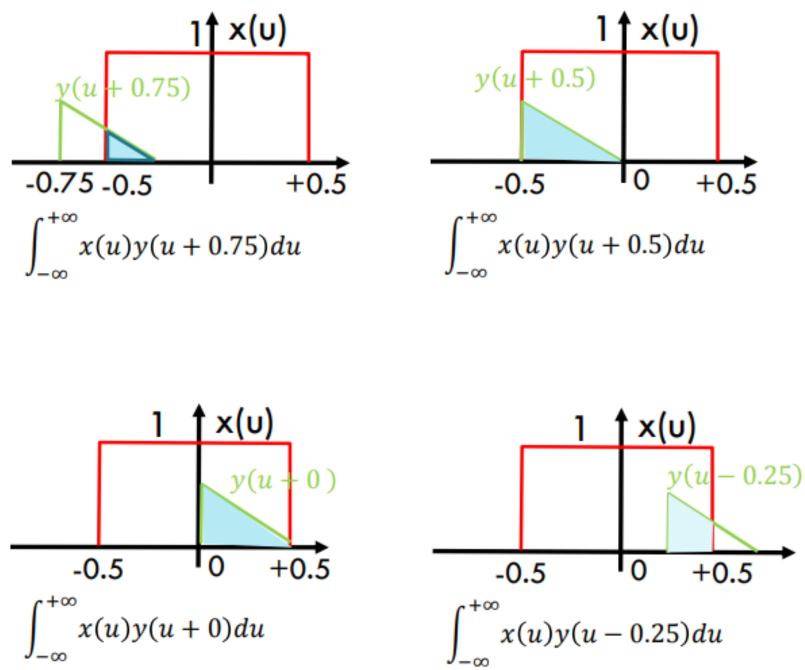
Symétrie hermitienne

Détection d’objets

Illustration

$$(x \circ y)(\tau) = \int_{-\infty}^{+\infty} x(u) y^*(u + \tau) du$$





<u>Mesure de similarité</u>: si le signaux sont identiques, l’aire d’intersection sera maximale lors que les signaux seront superposés!

Détection d'objets

Corrélation versus Convolution

Produit de corrélation

$$\Gamma_{xy}(\tau) = \int_{-\infty}^{+\infty} x(u)y^*(u-\tau)du$$

?

Produit de convolution

$$x(\tau) * y(\tau) = \int_{-\infty}^{+\infty} x(u)y^*(\tau-u)du$$

$$= \int_{-\infty}^{+\infty} x(u)y^*(-(\tau-u))du$$

$$\boxed{x(\tau) \circ y(\tau) = x(\tau) * y^*(-\tau)}$$

Produit de corrélation = produit de convolution lorsque le filtre h est paire

Détection d'objets

Corrélation: propriétés

- **Parité :** Considérons que x(t) soit un signal complexe, alors on peut écrire que:

$$\Gamma_{xx}(\tau) = R(\tau) + jI(\tau)$$

Puisque par symétrie hermitienne on a:

$$\Gamma_{xx}(\tau) = \Gamma_{xx}{}^{*}(-\tau)$$

$$R(\tau) + jI(\tau) = R(-\tau) - jI(-\tau)$$

Et donc on en déduit que :

$R(\tau)$ **est une fonction paire**

$I(\tau)$ **est une fonction impaire**

- **Borné :** On montre que la fonction d'autocorrélation vérifie:

$$|\Gamma_{xx}(\tau)| \leq \Gamma_{xx}(0) = \int_{-\infty}^{+\infty} |x(t)|^2 dt > 0$$

Elle possède donc un encadrement supérieur en τ=0 égal à l'énergie du signal

démonstration

En appliquant l'inégalité de Schwartz sur l'intercorrélation $< x, y_\tau >$, on obtient :

$$| < x, y_\tau > |^2 \leq < x, x > < y_\tau, y_\tau >,$$

que l'on peut écrire :

$$|\Gamma_{xy}(\tau)|^2 \leq \Gamma_{xx}(0)\Gamma_{yy}(0).$$

Dans le cas particulier où $x(t) = y(t)$, on a :

$$|\Gamma_{xx}(\tau)| \leq \Gamma_{xx}(0).$$

Détection d'objets

Corrélation 2D… mesure de similarité pour l'image

Produit de corrélation: version 2D continue

Considérons deux signaux continues **x** et y . Alors le produit de corrélation analogique $(x \circ y)(\tau, \eta)$ est défini par:

$$\Gamma_{xy}(\tau, \eta) = (x \circ y)(\tau, \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x(u, v) y^*(u + \tau, v + \eta) du\, dv$$

Produit de corrélation : version 2D numérique

Considérons un signal échantillonné 2D **x** comprenant N*N échantillons et un signal échantillonné 2D **y** comprenant M*M coefficients, avec M<<N. Alors le produit de corrélation numérique (x o y)(n,m) est défini par:

$$(x \circ y)(n, m) = \frac{1}{M^2} \sum_{k=-E[\frac{M}{2}]}^{E[\frac{M}{2}]} \sum_{l=-E[\frac{M}{2}]}^{E[\frac{M}{2}]} x(n + k, m + l) y(k, l), \quad (n, m) = [0, \dots, N]$$

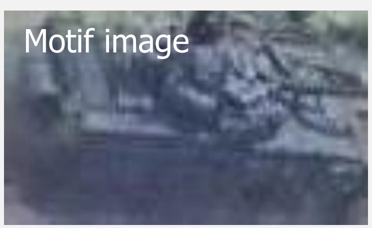
Matlab : *z=xcorr2(x,y)*

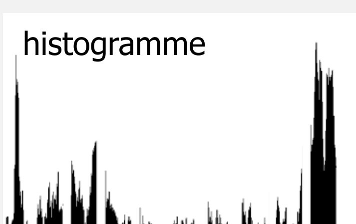
Principe général

Illustration

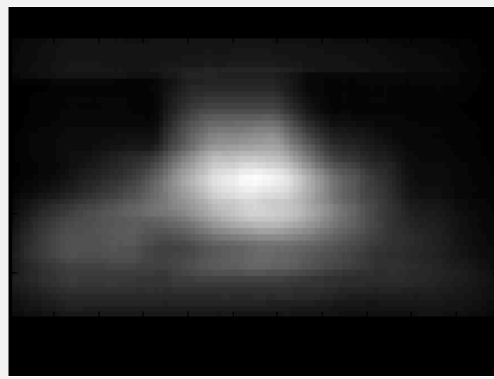
Image **I** comportant l'objet d'intéret

Modèle d'apparence **X**





Vraisemblance sur la position de l'objet



Résultat du suivi

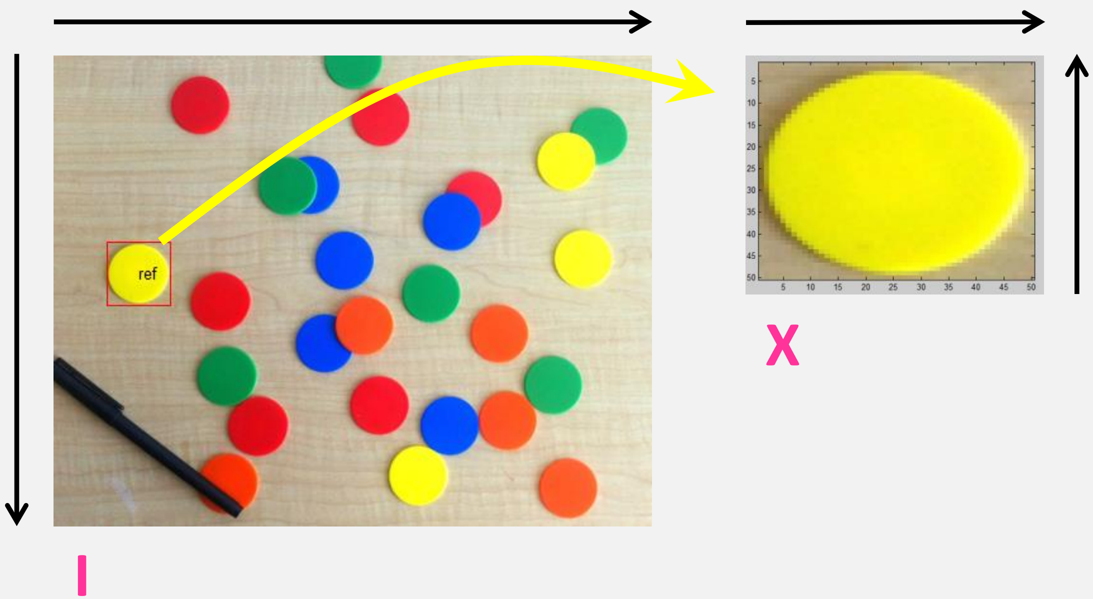
Méthodologie:

(e.g. correlation, SSD,…)

Corrélation

Ingrédients

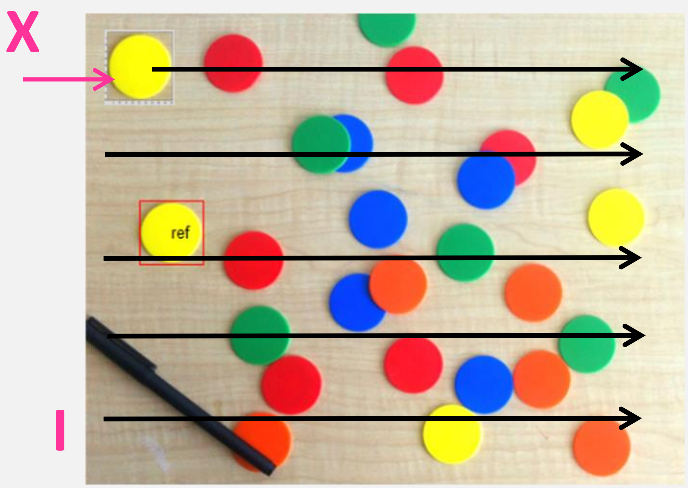
- Soit I une image de taille N.M dans laquelle on cherche à détecter les pièces jaunes.
- Soit X une image « référence » de la pièce jaune de taille T

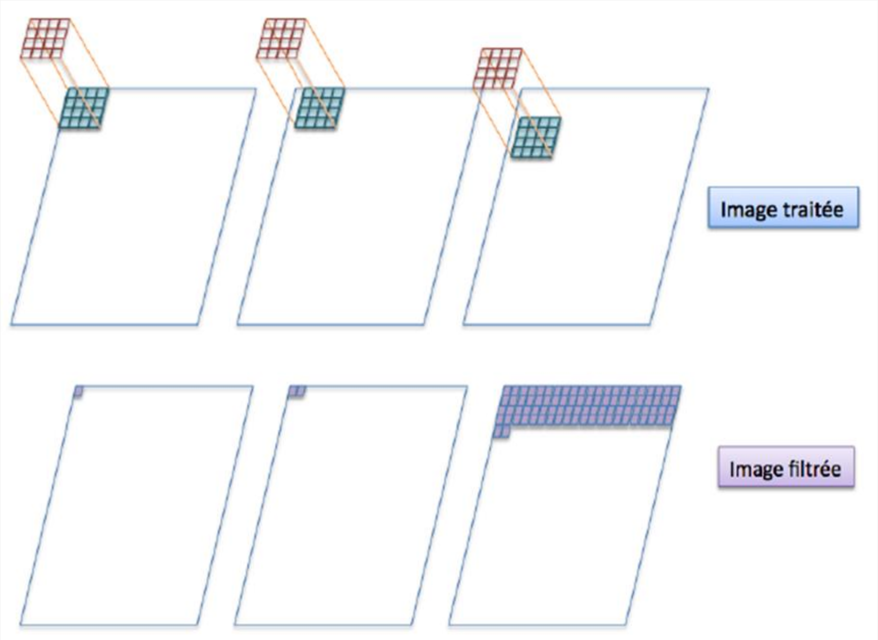


Corrélation

Principe

- On balaye le motif X sur toute l'image I et on calcule le produit de corrélation en chaque pixel de l'image





Corrélation

Un peu de math!

- Image en niveaux de gris: On balaye le motif X sur toute l'image I et on calcule le produit de corrélation en chaque pixel de l'image

$$C(i,j) = \frac{1}{T^2} \sum_{k=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} \sum_{l=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} I(\boldsymbol{i}+k,\boldsymbol{j}+l)X(k,l), \quad (\boldsymbol{i},\boldsymbol{j}) = [0,\dots,N]$$

X

- Image Couleur: On balaye le motif X=(r,g,b) sur toute l'image I= (R,G,B) et on calcule le produit de corrélation en chaque pixel de l'image

$$C(i,j) = \frac{1}{3T^2} \sum_{k=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} \sum_{l=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} (\boldsymbol{R}(\boldsymbol{i}+k,\boldsymbol{j}+l)\boldsymbol{r}(k,l) + \boldsymbol{G}(\boldsymbol{i}+k,\boldsymbol{j}+l)\boldsymbol{g}(k,l) + \boldsymbol{B}(\boldsymbol{i}+k,\boldsymbol{j}+l)\boldsymbol{b}(k,l))$$

Corrélation

Résultat du score de corrélation

- Que remarque t'on?

Image en niveaux de gris (imshow(C,[]))

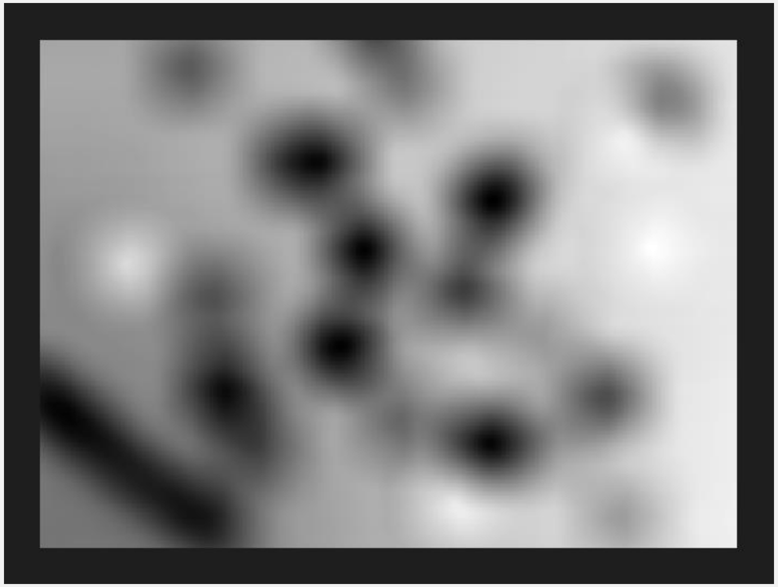
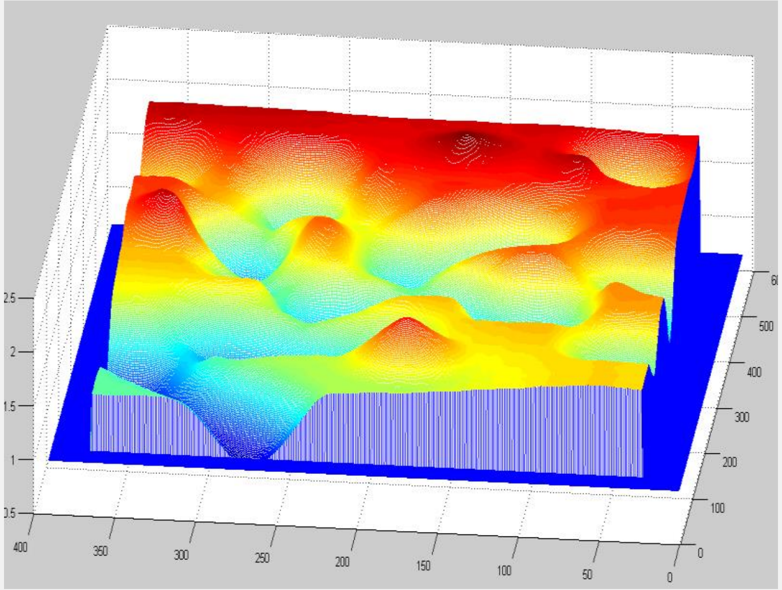


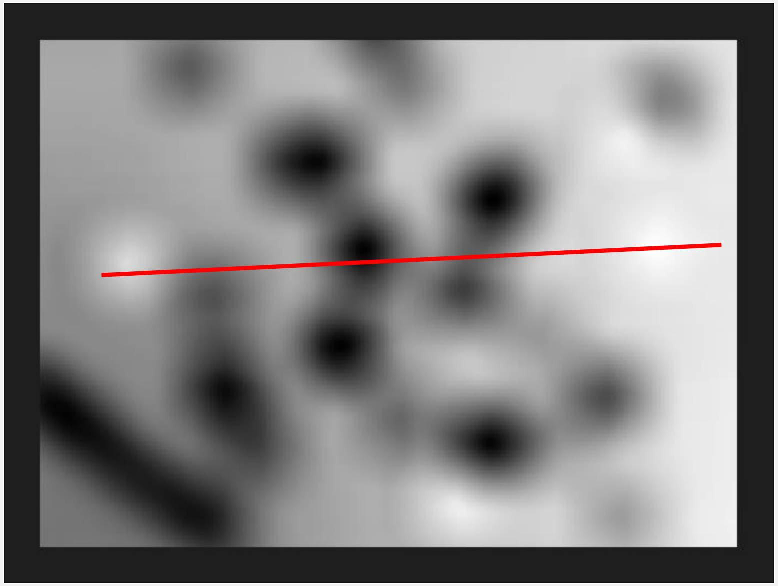
Image 3D (mesh(C))

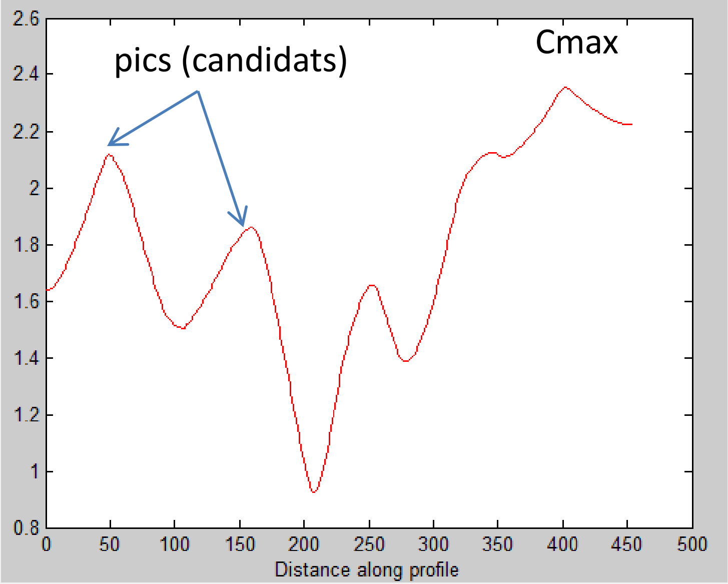


Corrélation

Comment détecter: seuillage

- Rechercher le score max: Cmax=max(C)
- Problème:
  - présence de plusieurs pics car plusieurs instances de la pièces sont dans la scène
  - Pas la même réponse (pas la même amplitude)

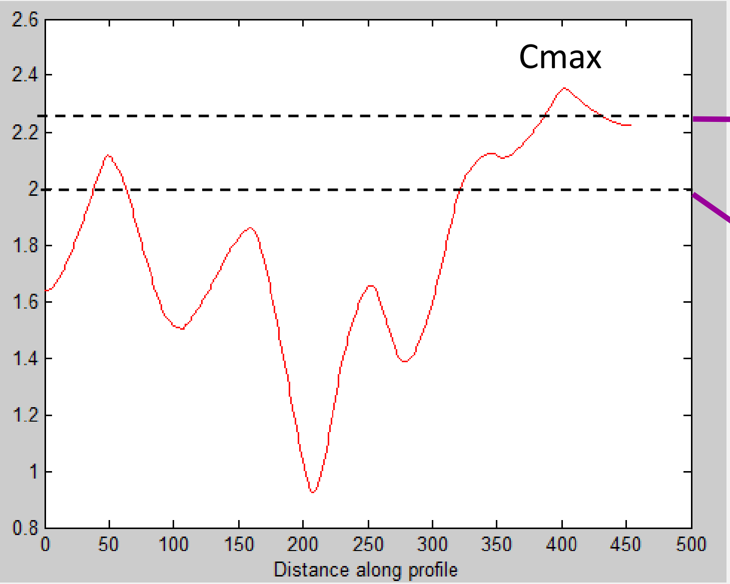


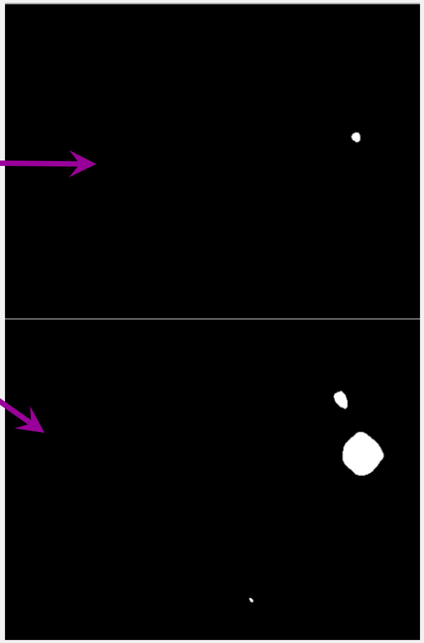


Corrélation

Comment détecter: seuillage

- A quel niveau on peut seuiller?



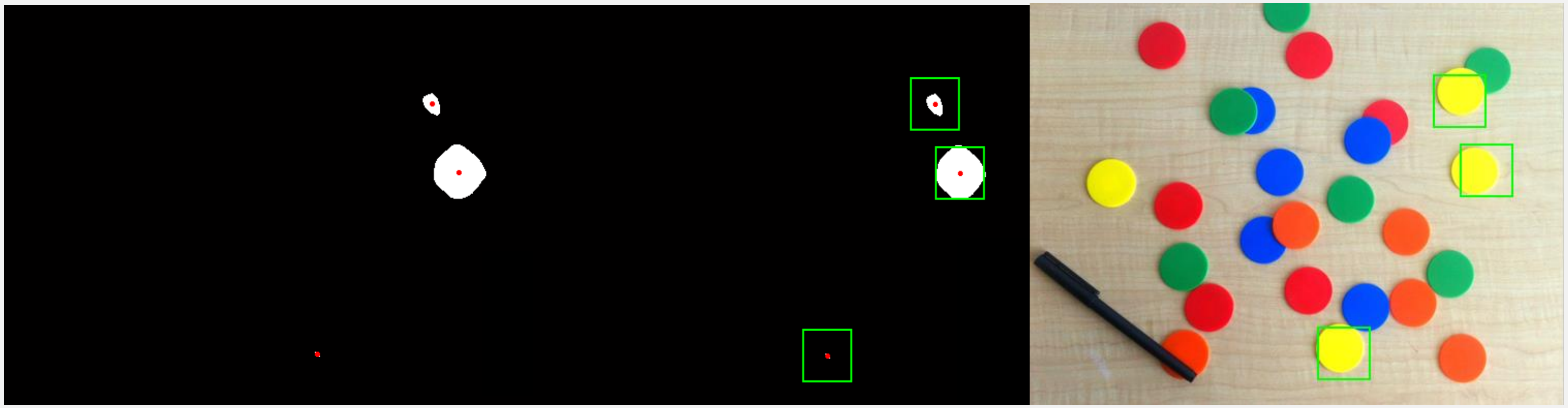


- Pas si simple: réponse multiple - plus le seuil diminue et plus le nombre de régions candidates apparaissent

Corrélation

Comment détecter: seuillage

- Recherche des centres des régions puis affichage des résultats:



- Résultats: toutes les pièces jaunes ne sont pas détectées et apparition d'un décalage dans la détection des pièces. Faut t'il diminuer le seuil encore? Le prochain épisode en TP…

Corrélation

Corrélation: inconvénient

- Appliquons notre motif X de taille 3.3 à un morceau d'image de niveau de gris constant v >=max(a,b,c,d,e,f,g,h,i)

a	b	c
d	e	f
g	h	i

$\circ$

v	v	v
v	v	v
v	v	v

Resultat: $9*v*(a+b+c+d+e+f+g+h+i) > (a+b+c+d+e+f+g+h+i)^2$

- Le score est donc plus grand et donc il peut apparaitre des erreurs de détection

Corrélation

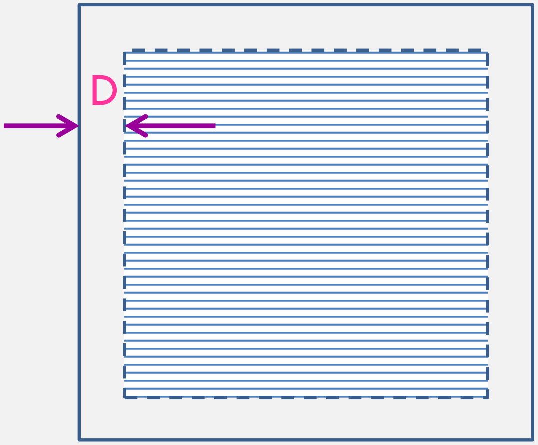
algorithme: **CorrImage**

- Entrée: I (image, N×M), X (référence, T×T)
- Sortie: C (score de corrélation)

- Algorithme:

1. Init de C: C=zeros(N,M);
2. Définition des bords: D=(T-1)/2;   % T taille du bloc référence (impaire)
3. Balayage:

```
For i =1+D : N-D
        For j=1+D : M-D
                A=I(i-D : i+D, j-D : j+D);
                B=A.*X;
                C(i,j)=sum(B( : ));
        End
End

C=C/(T*T);
```



Corrélation

Corrélation normalisée

- Objectif: rendre la mesure de corrélation plus discriminante en normalisant la référence et l'image

- Entrée: I (image, N×M), X (référence, T×T)
- Sortie: C (score de corrélation)

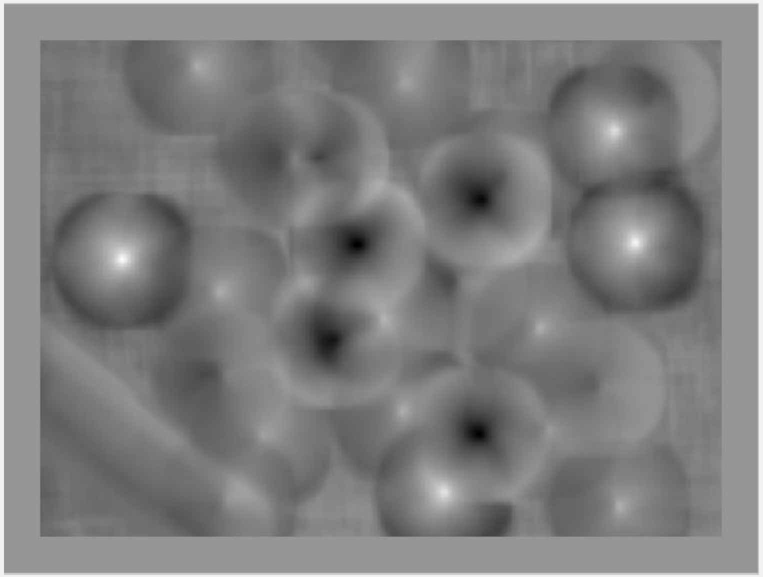
1. Init de C: C=zeros(N,M);
2. Définition des bords: D=(T-1)/2; % T taille du bloc référence (impaire)
3. **Standardisation de X**
4. Balayage:

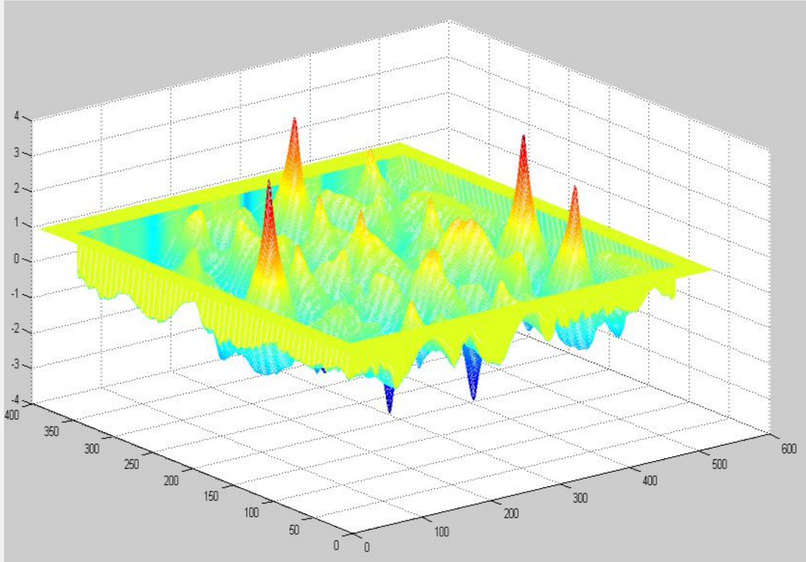
```
For i =1+D : N-D
    For j=1+D : M-D
        A=I(i-D : i+D, j-D : j+D);
        Standardisation de A
        B=A.*X
        C(i,j)=sum(B( : ));
    End
End

C=C/(T*T);
```

Corrélation

Résultat du score de corrélation normalisée

- Que remarque t'on? La corrélation normalisée améliore nettement la détection : les pics blancs sont plus nettes et plus étroits. Ils semblent que les pièces jaunes ont été identifiées mais l'amplitude de réponse n'est pas identique d'une pièce à une autre.. Problème de seuil… (voir TD)





Détection d'objets

Les Mesure de similarité

- La mesure de corrélation
- $\boxed{\text{La somme des différences au carré: SSD}}$
- La mesure de Bhattacharya

La somme des différences aux carrés

Un peu de math!

- Image en niveaux de gris: On balaye le motif X sur toute l’image I et on calcule la différence au carré entre les pixels de l’image I et X

$$SSD(i,j) = \frac{1}{T^2} \sum_{k=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} \sum_{l=-E[\frac{T}{2}]}^{E[\frac{T}{2}]} (\, I(\boldsymbol{i}+k, \boldsymbol{j}+l) - X(k,l)\,)^2, \quad (\boldsymbol{i},\boldsymbol{j}) = [0, \dots, N]$$

X

- Que donne le développement…

- Et pour une image couleur…

- On ne recherche plus le maximum mais le minimum maintenant…

Détection d'objets

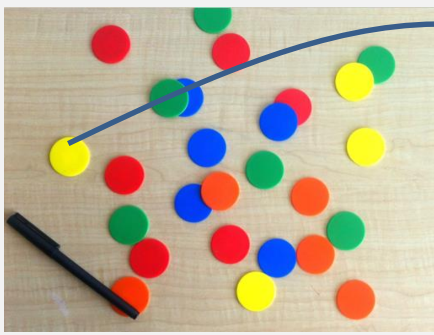
Les mesures de similarité

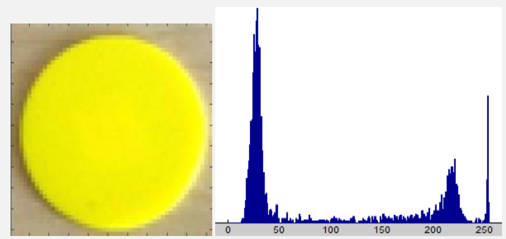
- La mesure de corrélation
- La somme des différences au carré: SSD
- La mesure de Bhattacharya

Détection d'objets

Principe

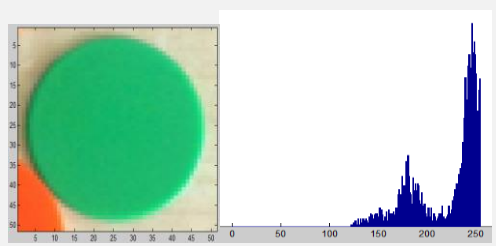
- Mesure de similarité basée sur le calcul de l'histogramme

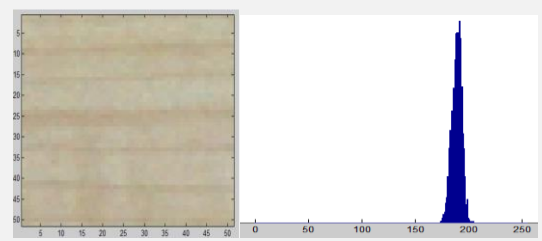


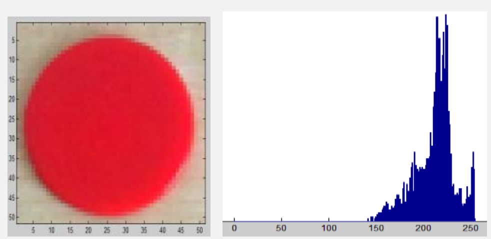


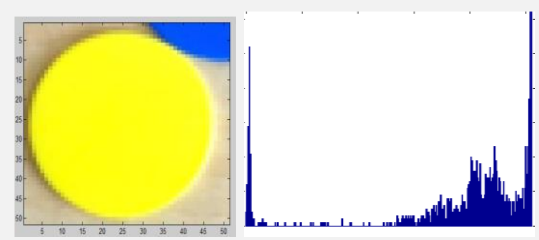
X: référence

Blocs image candidats





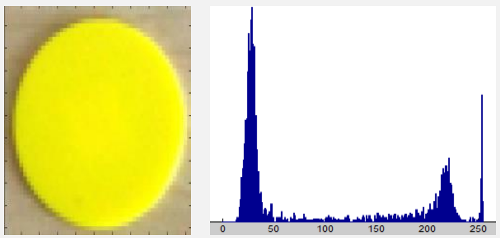




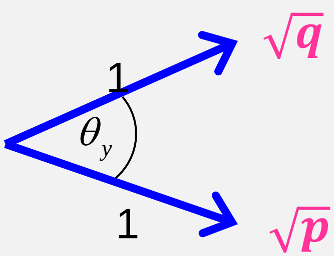
Détection d'objets

Mesure de Bhattacharya

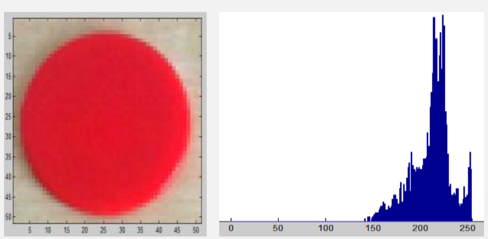
X: référence



$$\boldsymbol{q} = \{q_1, q_2, \dots, q_m\} \qquad \sum_{k=1}^{m} q_k = 1$$



Bloc candidat



$$\boldsymbol{p} = \{p_1, p_2, \dots, p_m\} \qquad \sum_{k=1}^{m} p_k = 1$$

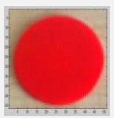
Coefficient de Battacharya

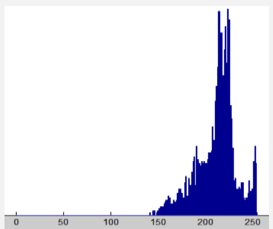
$$\rho = \cos(\theta) = \frac{< \sqrt{p}, \sqrt{q} >}{|\sqrt{p}||\sqrt{q}|} = \sum_{k=1}^{m} \sqrt{p_k q_k}$$

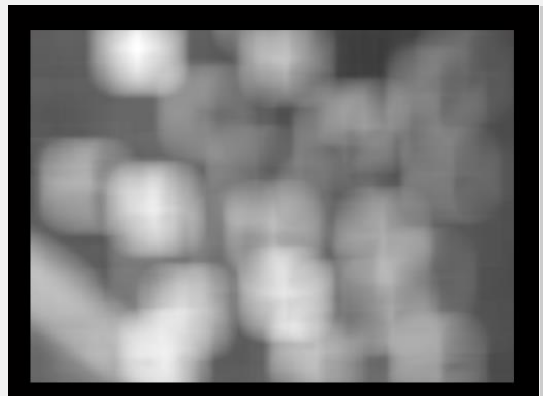
Détection d'objets

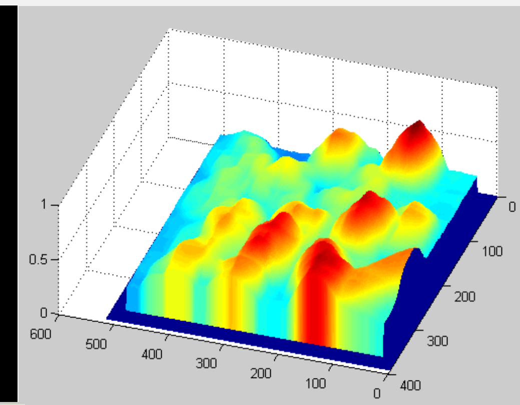
Resultats

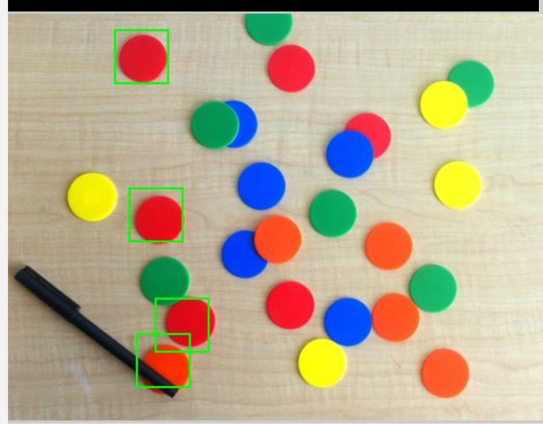
X: référence









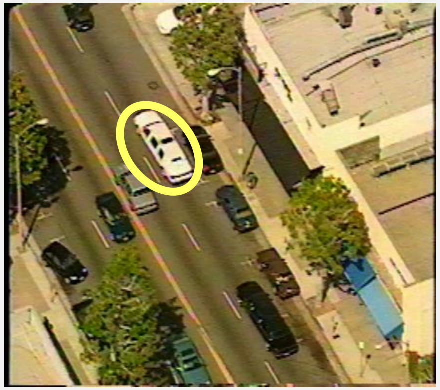


Seuil=0.9

Quelques questions pratiques

Forme de la fenêtre

Forme de fenêtre mal adaptée à la forme de l'objet. Intégration importante de l'arrière plan

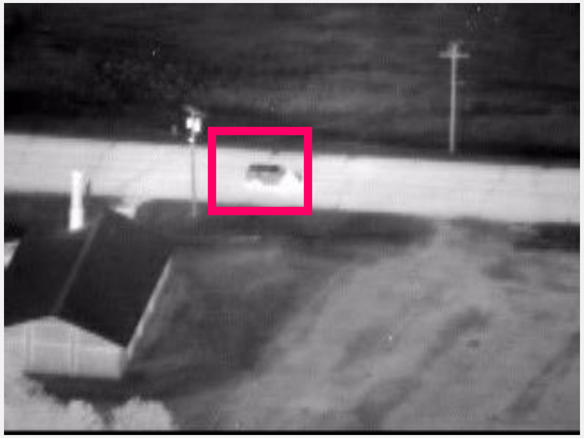


Utilisation d'une fonction de fenêtrage gaussienne pour pondérer plus fortement les pixels appartenant à l'objet que ceux issus de l'arrière plan

Quelques questions pratiques

Le temps de calcul

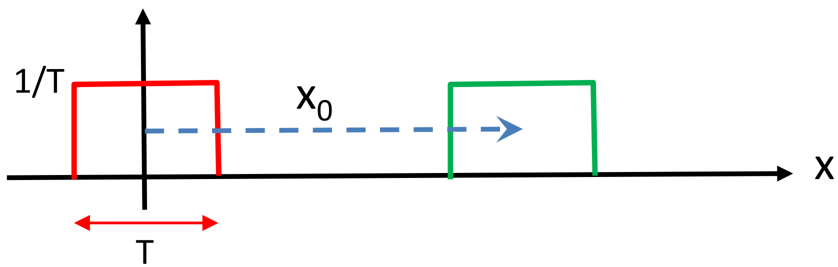
Solution: la zone de recherche doit être limitée. Sa taille doit être fonction de l'amplitude de déplacement. Nécessité d'une connaissance de la dynamique de l'objet et de la caméra (si elle bouge).



Détection d’objets

Détection d’objets par leur mouvement: transformée de Fourier

Illustration en 1D: une fonction porte



$$u(x) = \frac{1}{T}\Pi\left(\frac{x}{T}\right)$$

$$\mathcal{F}\big(u(x)\big) = sin_c(\pi f T)$$

$$v(x) = \frac{1}{T}\Pi\left(\frac{x-x_0}{T}\right) = u(x - x0)$$

$$\boxed{\mathcal{F}[f(x-x_0)] = \exp\left(-2j\pi\nu x_0\right)\, F(\nu)}$$

$$\mathcal{F}\big(v(x)\big) = sin_c(\pi f T)\exp(-j2\pi x_0 f)$$

Soit la fonction $\Pi_T(x)$, définie par,

$$\Pi_T(x) = \frac{1}{T}\Pi\left(\frac{x}{T}\right) = \begin{cases} \frac{1}{T} & \text{si} \quad |x| \le \frac{T}{2} \\ 0 & \text{si} \quad |x| > \frac{T}{2} \end{cases}$$

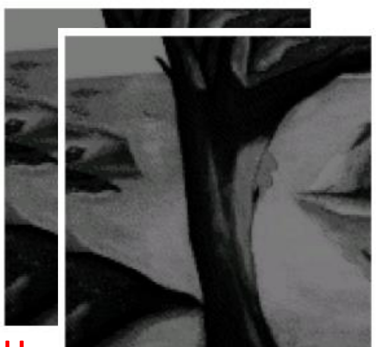
$$\mathcal{F}\left(\frac{1}{T}\Pi\left(\frac{x}{T}\right)\right) = \frac{\sin \pi\nu T}{\pi\nu T}$$

$$\boldsymbol{\mathcal{H}(f) = \frac{\mathcal{F}(v)}{\mathcal{F}(u)} = \exp(-j2\pi x_0 f)}$$

$$\Rightarrow \boldsymbol{h(x) = \mathcal{F}^{-1}(\mathcal{H}(f))?}$$

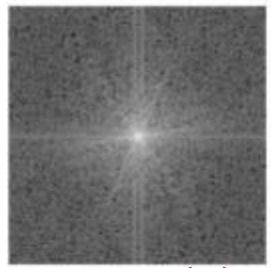
Détection d'objets

Illustration: deux images séparées par une translation tx=7, ty=11

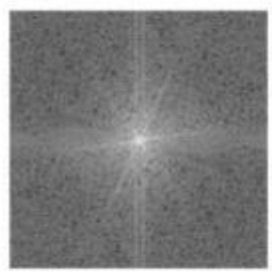


u

v

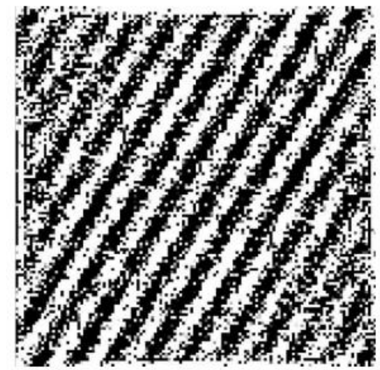


$U = \mathcal{F}(u)$



$V = \mathcal{F}(v)$

$\frac{V}{U}$



$\mathcal{F}^{-1}\left(\frac{V}{U}\right)$