



Laboratoire Bordelais  
de Recherche  
en Informatique

# Indexation des images et de la vidéo

---

Jenny Benois-Pineau  
LABRI UMR 5800 Université Bordeaux  
1/Bordeaux2/IPB-ENSEIRB-Matmeca

06/04/2012



# Indexation des images et de la vidéo

---

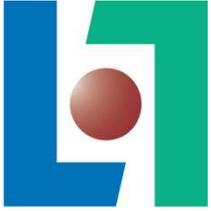
- 1. Description du contenu des images
  - Descripteurs globaux
  - Descripteurs locaux
  - Quantification : Notion de BOF et BOVW
    - Mesures de similarité et distances
    - Prise en compte du contexte spatial : SPMK, GraphWords
  
- 2. Extensions spatio-temporelles
  - Comparaisons des séquences vidéo
  - Mouvement : descripteur particulier
  - Estimation
  - Segmentation en plages du mouvement homogène.



# 1. Description du contenu des images

---

- Objectifs :
- -recherche des images par similarité dans des bases de données
- - reconnaissance des scènes, des objets
- - classification non supervisée /groupement des images similaires dans une BD ou des images une une vidéo
- - classification supervisée
- Formulation générale du problème de recherche par similarité :



- Etant donné  $D \subseteq R^n, X_q \in D, X = \{X_p\}, X_p \in D, p = 1, \dots, N, p \neq q$ , et  $d(Y, Z)$  - métrique sur  $R^n$  ou  $sim(Y, Z)$  - mesure de similarité sur  $R^n$ , trouver
- $X^* \in X$ :

$$X^* = \underset{p=1, \dots, N}{\operatorname{Argmin}} d(X_q, X_p)$$

ou

$$X^* = \underset{p=1, \dots, N}{\operatorname{Argmax}} sim(X_q, X_p)$$

respectivement

Notre problème 1: définir l'espace de descripteurs  
du contenu  $X$   $d$  ou  $sim$



# Evaluation des performances

---

- TP, FP, TN, FN
- Rappel(R) :
- $R = TP / (TP + FN) = TP / (Ngt)$
- Précision(P) :
- $P = TP / (TP + FP) = TP / (Nd)$
- F-mesure normalisée (F) :
- $F = 2 / (1/R + 1/P)$
- *MAP : Mean average precision – mesure tenant compte du rang*
- *Précision à n :*
- *Soit TP à la position n, P(n) – combien de TP sur n premiers/n*
- *Average Precision:  $\frac{\sum_{n=1}^N P(n)\delta(rel,n)}{nbr\ rel\ doc}$*
  
- $MAP = \frac{1}{Q} \sum_{q=1}^Q AP(q)$



Laboratoire Bordelais  
de Recherche  
en Informatique

# Descripteurs globaux

---

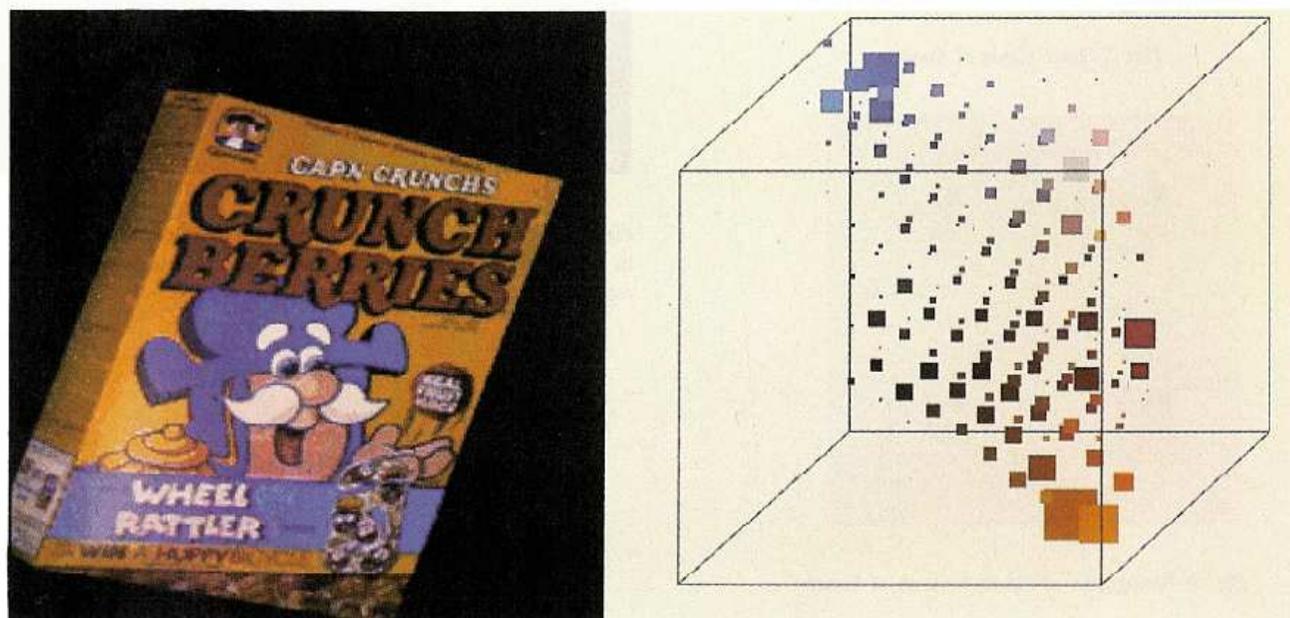
- M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.
- MPEG7: ISO/IEC 15938-3:2002 Information technology -- Multimedia content description interface -- Part 3: Visual
- **Descripteurs globaux du contenu des images en termes de couleur et contrastes couleur, texture, contours..**



Laboratoire Bordelais  
de Recherche  
en Informatique

# Descripteurs globaux : couleur

- Histogramme couleur (S&B)



M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

- Invariance : translation, rotation, et faibles changements de l'angle de vue

06/04/2012



# Mesures de similarité et distances sur des histogrammes(1)

- Intersection des histogrammes

$$\delta(H_q, H_p) = \sum_{k=1}^K \min(H_q(k), H_p(k)) / \text{Card}(I)$$

$$\delta(h_q, h_p) = \sum_{k=1}^K \min(h_q(k), h_p(k))$$

- Coefficient et distance de Battacharya

$$\rho(h_q, h_p) = \sum_{k=1}^K \sqrt{(h_q(k) \cdot h_p(k))}$$

$$D(h_q, h_p) = -\ln(\rho(h_q, h_p))$$



## Mesures de similarité et distances sur les histogrammes(2)

- Distance  $\chi^2$  :

- $\chi^2(H_q, H_p) = \frac{1}{2} \sum \frac{(H_q(k) - H_p(k))^2}{(H_q(k) + H_p(k))}$  Tapez une équation ici.

- Cette distance permet de réduire l'effet des classes de forte population.
- Encore S&B montrent que leur mesure de similarité des histogrammes peut être exprimée via la norme L1:

$$1 - \delta(H_p, H_q) = \frac{1}{2T} \sum_{k=1}^K |H_p(k) - H_q(k)|, T = \text{Card}(I)$$

- Distances entre les histogrammes: normes L1 et L2 dans l'espace  $R^K$
- Plus généralement \*

$$D(H_p, H_q) = \sqrt{(H_p - H_q)^T \Sigma (H_p - H_q)}$$



## Mesures de similarité et distances sur les histogrammes(3)

---

- $\Sigma$  est appelée la matrice de similarité des classes.
- Si  $\Sigma$  est la matrice inverse de covariance,
- La distance est celle de Mahalanobis.
- Dans le cas de
- $\Sigma = E$ , nous avons L2
- D'autres distances sont appliquées : « Earth – mover distance » etc.



## Descripteurs de la couleur en MPEG7 (2)

- **Dominant color descriptor** : l'ensemble des couleurs dominantes dans la région d'intérêt ou dans l'image entière fournit une description compacte facile à indexer.

$$F = \{\{c_i, p_i, v_i\}, s\}, (i = 1, 2, \dots, N)$$

- $c_i$  - ième couleur dominante
- $p_i$  - pourcentage dans l'image
- $v_i$  - variance couleur;
- $s$  - mesure de cohérence spatiale (nombre moyen normalisé des pixels connexes de même couleur dans un voisinage 3x3). = Histogramme de couleurs



## Calcul du DCD – Split-LBG

$$A = \left\{ \left( a^{(1)}, \dots, a^{(n)} \right)_i^T \right\}_{i=1}^{\infty}$$

- Nombre de classes final fixé à priori  $K=2^n$
- Méthode :
  1. Initialisation : tirage aléatoire des centres de  $M < K$  classes  $(M=2) \{c_1, \dots, c_M\} \in A$
  - 2) « Split » de  $D$  :  $c_i \mapsto (c_i + \bar{\epsilon}, c_i - \bar{\epsilon})$
  - 3) LBG
    - a - affectation  $C(a_j) = \text{Arg} \min_k d(a_j, c_k), c_k \in D_i$
    - b - calcul du nouveau centroïde de chaque classe :  $\bar{g}_k = \frac{1}{N_k} \sum_{\bar{a}_l \in C_k} \bar{a}_l$
    - c - calcul de la distorsion  $E = \sum_k \sum_i d(\bar{a}_{ik}, c_k)$
  - si  $E^I < \alpha E^{I-1}$  alors arrêter sinon (3)
  - 4) Arrêt si K classes sinon retour en (2).



# Descripteur Color Layout

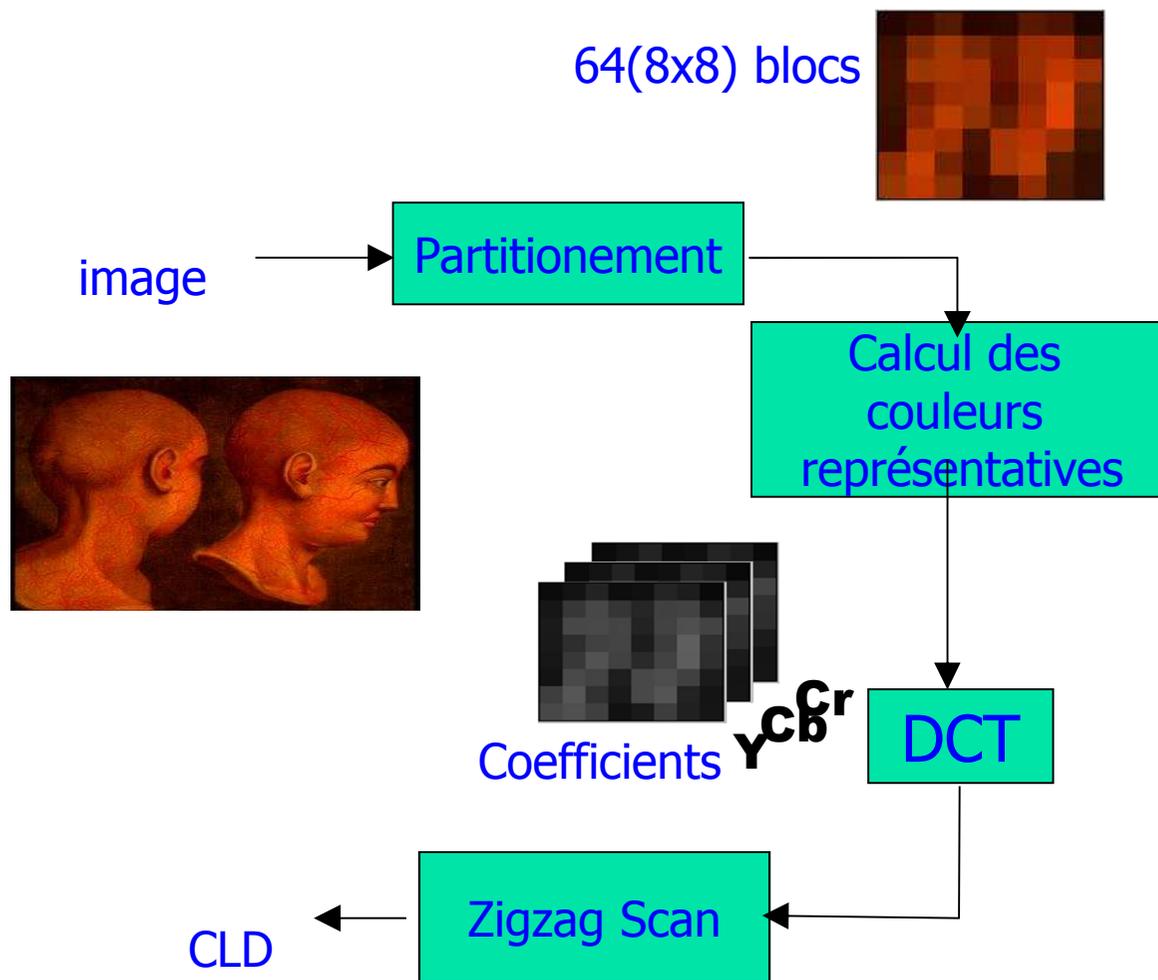
- **Color Layout Descriptor** : représentation compacte de la distribution spatiale des couleurs dans l'image, indépendante de la résolution.
- Calcul :
  - - Partitionner image de taille MxN en 64 blocs (8x8).
  - - Calculer la valeur représentative (moyenne) par bloc
  - Calculer la DCT sur l'imagette des valeurs représentatives

$$CLD = (DY_i, DCr_j, DCb_j), (i = 1, \dots, m, j = 1, \dots, n)$$

- Coefficients DCT de luminance
- Coefficients DCT de chrominance rouge
- Coefficients DCT de chrominance bleu
- $m = 6, n = 3$  par exemple



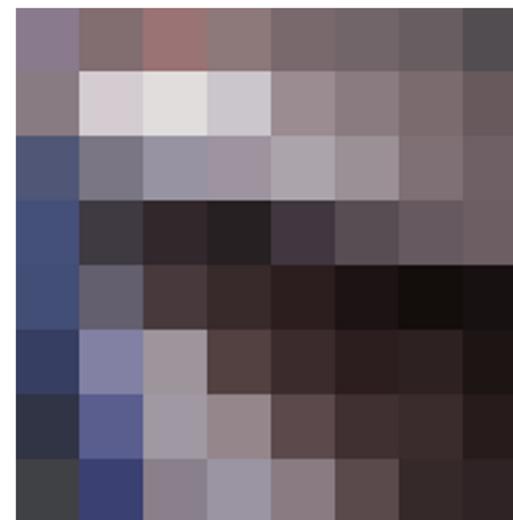
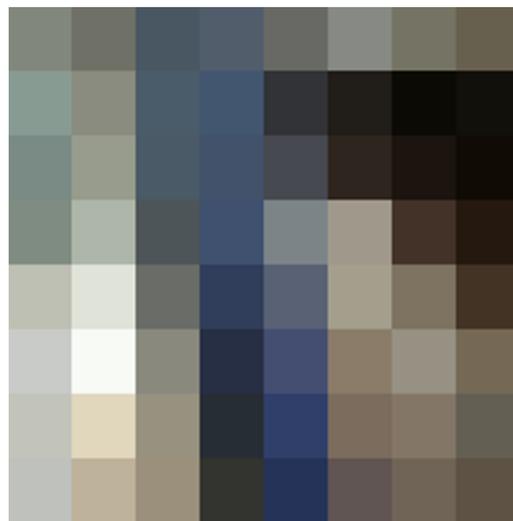
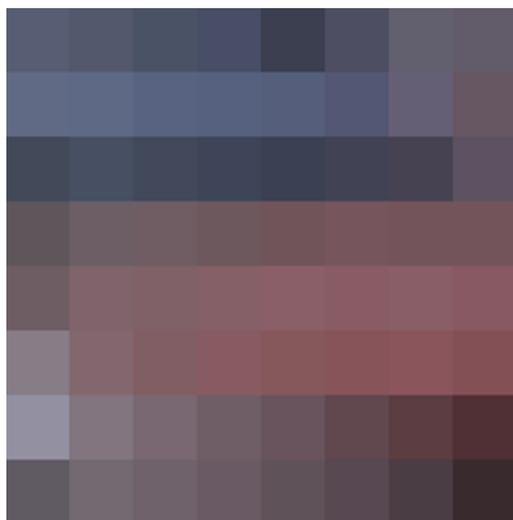
# Calcul du descripteur





Laboratoire Bordelais  
de Recherche  
en Informatique

# Exemples de CLD

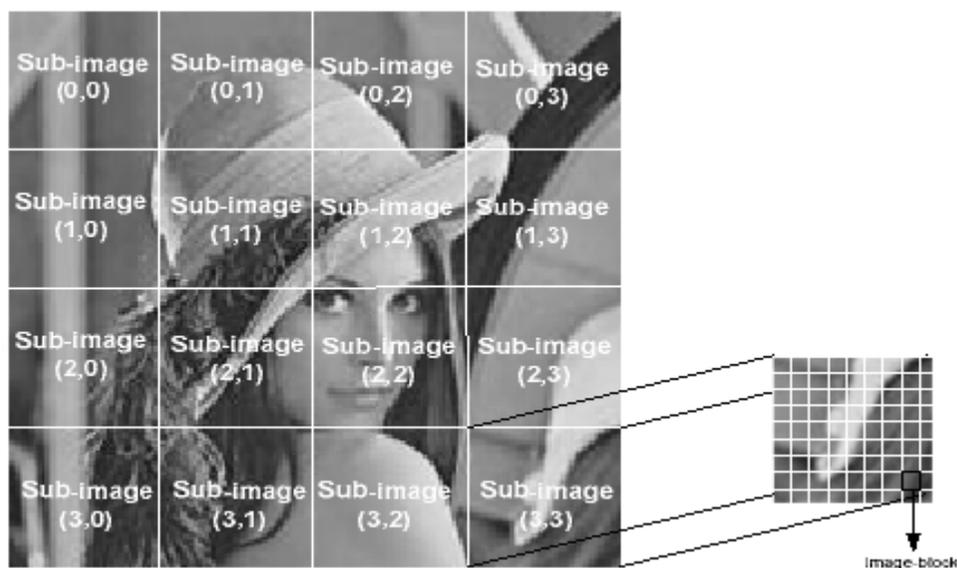


06/04/2012



# Descripteurs de texture/contours

- Descripteur EDH – histogramme de contours



- Permet de caractériser la direction des contours dans des zones de l'image



- Principe de calcul :
- diviser l'image en 16 blocs;
- détecter les contours dans chaque bloc;
- et déterminer leur orientation parmi les 5 :  
H,V,45°,135°,ND (non-directionnal);
- pour chaque bloc calculer l'histogramme normalisé des orientations des contours;
- concaténer les histogrammes dans un vecteur à  $16 \times 5 = 80$  dimensions.



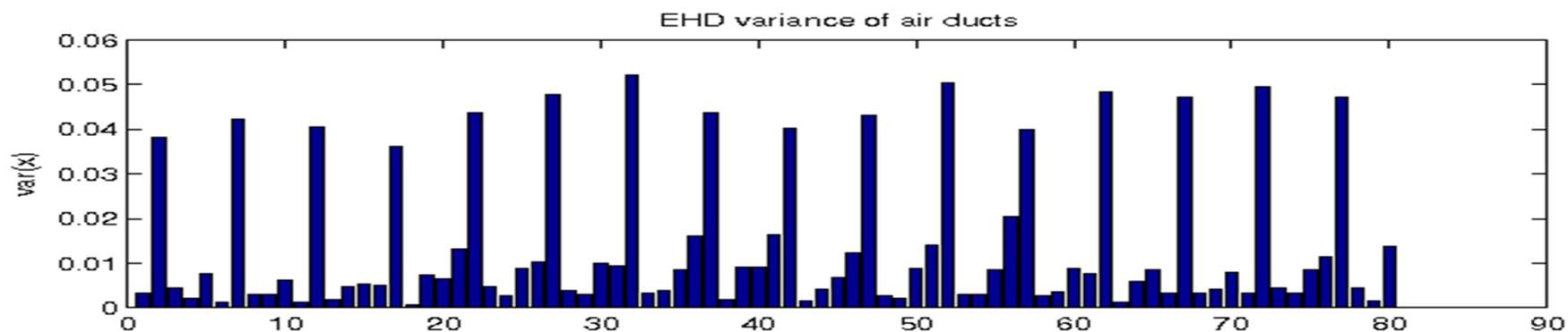
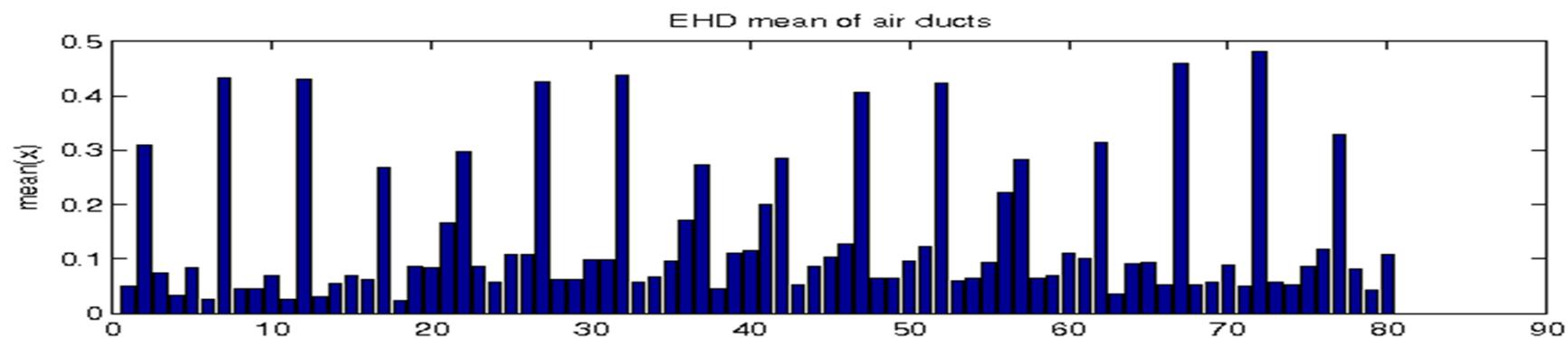
Laborat  
de Recl  
en Info



06/04/2012



- Exemple : les histogrammes moyennes et variance pour un ensemble de 55 exemples





# Descripteurs locaux

- Principe :
  - -a) Détecter des points, des régions « singuliers » dans l'image et décrire la texture autour
  - -b) Décrire la texture en utilisant les descripteurs de (a) en effectuant le découpage régulier
  - « description dense » ou irrégulier « description par régions ».





- SIFT = scale invariant feature transform
- Problème : détecter les points caractéristiques et proposer les descripteurs invariants par rapport aux transformations affines du plan – image et de la liminance.

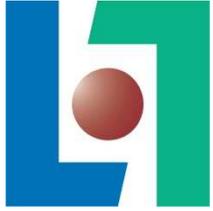
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

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

$\sigma$  -est le paramètre d'échelle

Mokolajczyk (2002) a montré que les points les plus stables sont obtenus comme extremums de  $\sigma^2 \nabla^2 G \circ I$   
Laplacien de Gaussienne d'une image

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



## Descripteurs SIFT (D. Lowe 2004) -2

Laboratoire Bordelais  
de Recherche  
en Informatique

- DOG est une bonne approximation de  $\sigma^2 \nabla^2 G$

- Considérons  $k\sigma = \sigma + \Delta\sigma \Rightarrow \Delta\sigma = (k-1)\sigma$



- Alors  $G(x, y, k\sigma) - G(x, y, \sigma) \approx \frac{\partial G}{\partial \sigma} (k-1)\sigma$  (3)

- $\frac{\partial G}{\partial \sigma} = \frac{1}{2\pi\sigma^5} (x^2 + y^2 - 2\sigma^2) \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$  (4)

- Par ailleurs

- $\nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} = \frac{1}{2\pi\sigma^6} (x^2 + y^2 - 2\sigma^2) \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$  (5)



## Descripteurs SIFT (D. Lowe 2004) -3

---

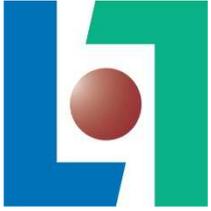
- De (4) et (5) :

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \quad (6)$$

- Finalement de (3) et (6):

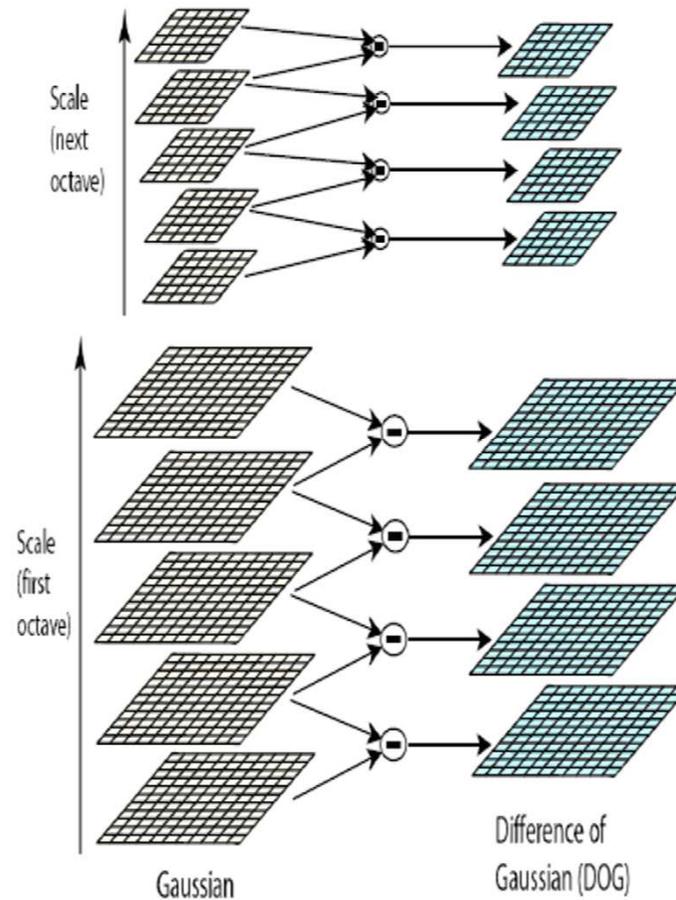
$$G(x, y, k\sigma) - G(x, y, \sigma) \approx \frac{\partial G(x, y, \sigma)}{\partial \sigma} (k-1)\sigma = (k-1)\sigma^2 \nabla^2 G(x, y, \sigma)$$

- DOG déjà comporte la  $\sigma^2$  normalisation requise pour Laplacien inv. à l'échelle. Le facteur (k-1) est constant pour toutes les échelles et n'influence pas la localisation des extremums



## Calcul de la DOG

- Convolver  $I$  avec  $G$  progressivement pour produire  $L$  séparées par  $k$  dans l'espace - échelle
- Une Octave  $\sigma, 2\sigma$  est divisée en  $s$  intervalles,  
 $\sigma, 2^{1/s}\sigma, 2^{2/s}\sigma, \dots, 2\sigma$
- Echelles adjacentes
- DOG:  
$$D(x, y, \sigma_i) = L(x, y, k\sigma_i) - L(x, y, \sigma_i)$$
- Sous-échantillonnage de facteur 2

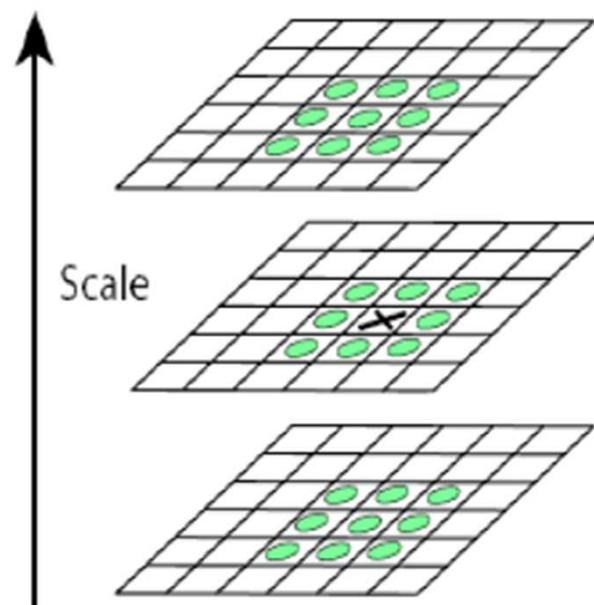




## Détection des extremums locaux

- Un point est retenu si DOG est minimale ou maximale parmi 26 voisins
- Choix du paramètre d'échelle
- valeur  $\sigma = 1,6$   
(Lowe[2004])

L'échelle  $s$  est affectée  
au point





## Filtrage le long des contours (1)

- Un faible pic de DOG (qui correspond au contour) aura une forte courbure dans la direction orthogonale au contour et une faible courbure dans la direction tangentielle.

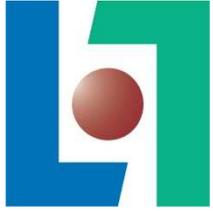
- $$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}$$
 Matrice Hessienne

- Les valeurs propres  $\alpha, \beta$  de H sont proportionnelles aux courbures principales de D au point considéré.

- Soit  $\alpha = r\beta$  avec  $\alpha$  - la valeur propre maximale (  $r \geq 1$  )

- $\alpha, \beta$  sont les racines du polynôme caractéristique  $\det(H - \lambda I)$ , alors

$$\alpha\beta = \text{Det}(H) \quad \alpha + \beta = \text{Tr}(H),$$



## Filtrage le long des contours (2)

---

- Finalement

$$\frac{\text{Tr} (H)^2}{\text{Det} (H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r + 1)^2}{r}$$

- Exclusion du point si

$$\frac{\text{Tr} (H)^2}{\text{Det} (H)} < \frac{(\text{Thr} + 1)^2}{\text{Thr}}$$



Laboratoire Bordelais  
de Recherche  
en Informatique

## Résultats de la détection

Image1



06/04/2012



Laboratoire Bordelais  
de Recherche  
en Informatique

## Résultats de la détection

### Image10

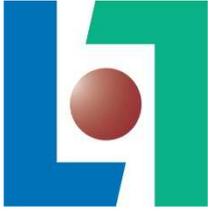


06/04/2012



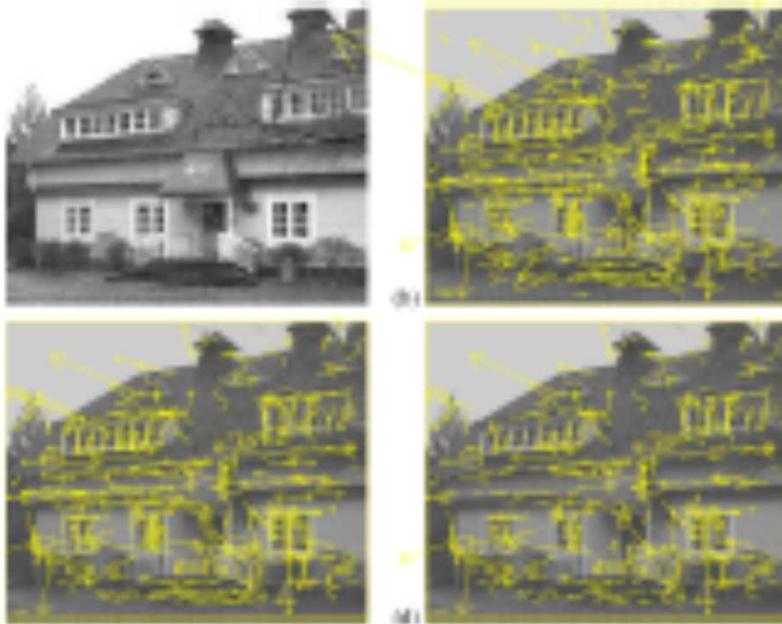
## Affectation de l'orientation

- Objectif : obtenir l'invariance du descripteur par rapport à la rotation en compensant par rapport à l'orientation locale
- For each sample of the image  $L(x, y)$  at a given scale compute :
  - Amplitude du gradient :  $m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$
  - Angle d'orientation :  $\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right)$
- Histogramme d'orientations et calculé dans le voisinage du point caractéristiques (36 orientations :  $360/10$ ).  
$$h(\theta_i) = h(\theta_i) + m(x, y) G(x, y, x_0, y_0, 1.5\sigma_s), i = 1, \dots, 36$$
  - Points multiples : l'histogramme contient plusieurs modes



Laboratoire Bordelais  
de Recherche  
en Informatique

## Affectation de l'orientation(2)

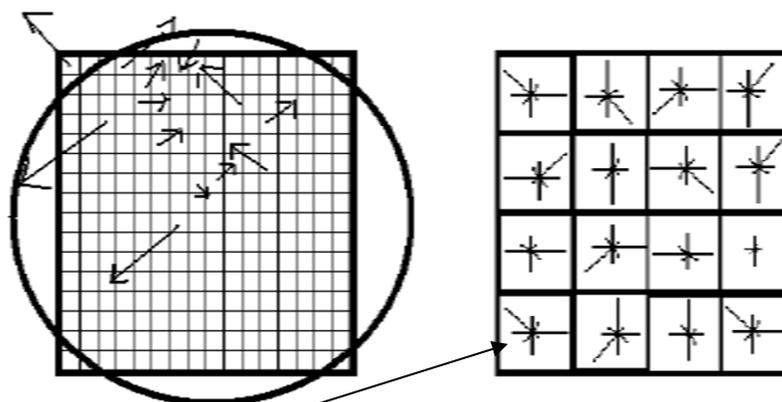


- Exemple de Lowe avec les orientations locales.



## Descripteur local autour du point caractéristique (1)

- Sélectionner une région – carré autour du point caractéristique ( 8x8 or 16x16 pixels);
- Calculer l'amplitude du gradient et l'orientation en chaque point de la région, pondérer l'amplitude avec une Gaussienne (cercle)
- Diviser en blocs de taille 4x4 et calculer l'histogramme d'orientation pour chaque bloc (8 orientations);



Bloc 16x16

Blocs 4x4

- Pondérer chaque contribution dans un bin par l'amplitude lissée du gradient
- Interpoler les histogrammes



## Descripteur local autour du point caractéristique (2)

---

- Le descripteur  $X$  est un vecteur des histogrammes  $4 \times 4$  concaténés  $\dim(X) = 4 \times 4 \times 8 = 128$
- « Tourner » le descripteur  $X$  par rapport à l'orientation du point caractéristique pour obtenir l'invariance par rapport à la rotation
- Normaliser :  $x \leftarrow x / \|x\|$
- - invariance par rapport aux transformations affines de la luminance

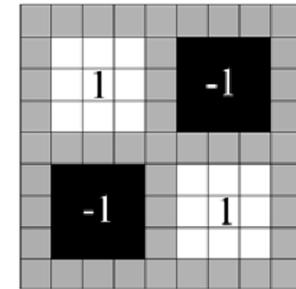
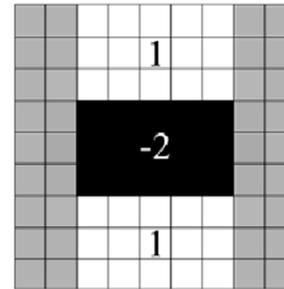
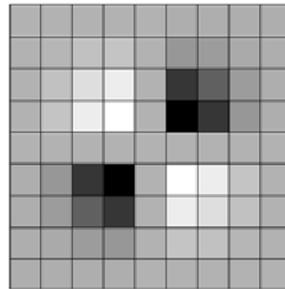
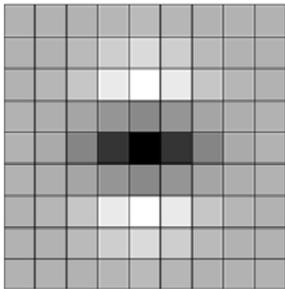


# SURF (speed-up robust features)(1)

- Points caractéristiques : max de det de Hessian - blobs

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

L-convolution de l'image avec la dérivée seconde d'une Gaussienne

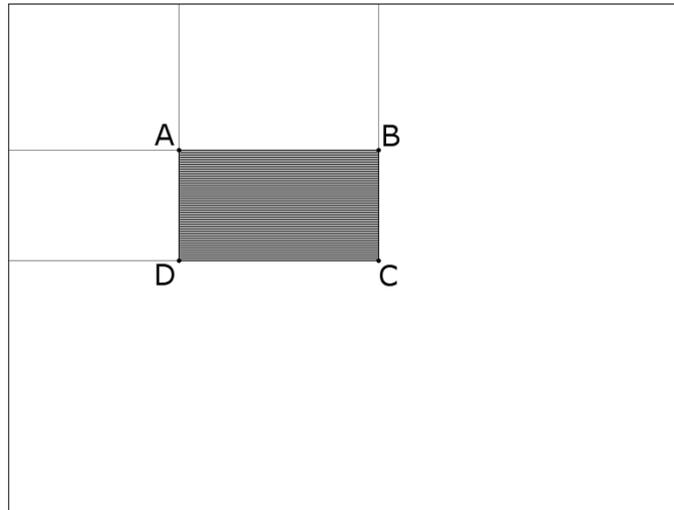




Laboratoire Bordelais  
de Recherche  
en Informatique

# SURF (speed-up robust features)(2) H. Bay et al.

- Image intégrale



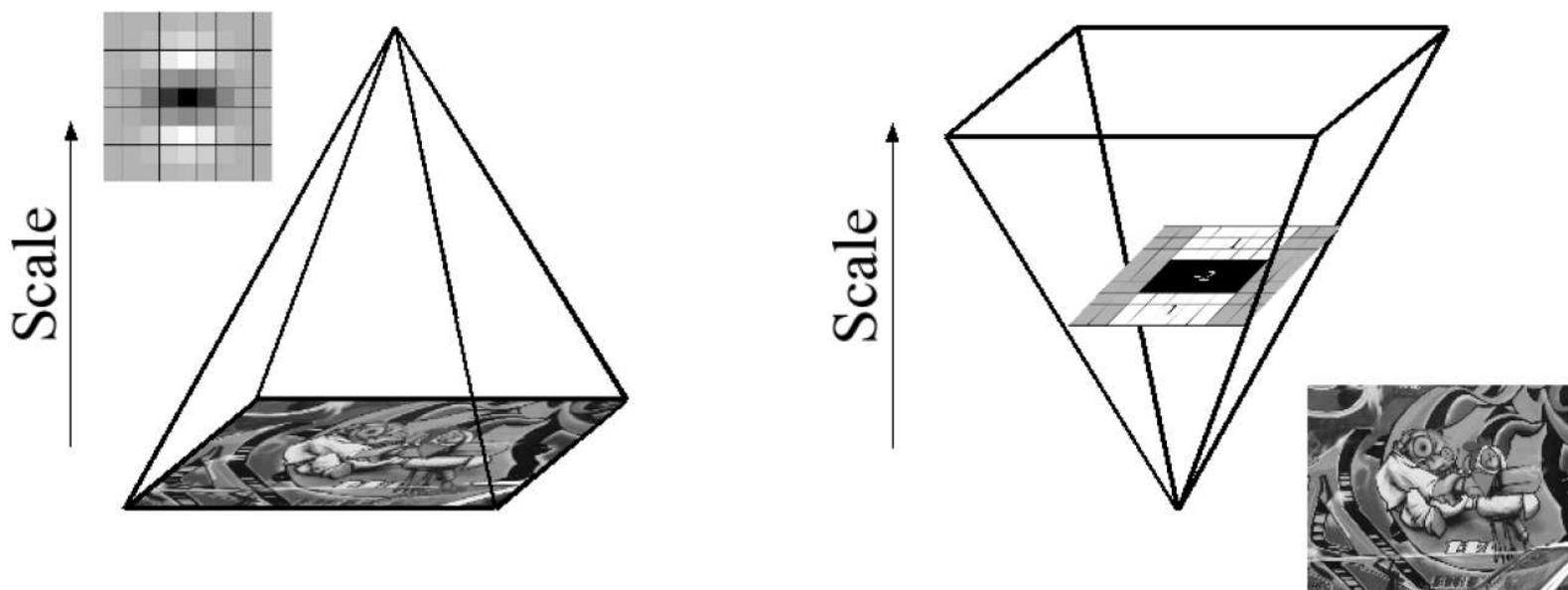
$$C-B-D+A$$

- Calcul rapide de convolution



Laboratoire Bordelais  
de Recherche  
en Informatique

## SURF (speed-up robust features)(3)



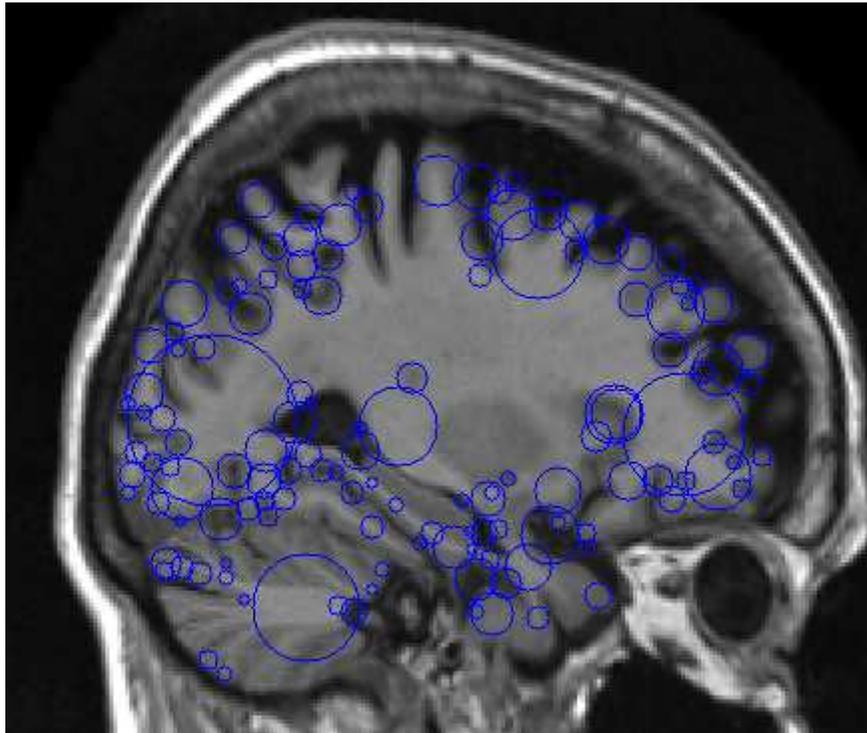
Augmentation de taille de filtre au lieu de  
changement de l'image pour la détection à  
différentes échelles



Laboratoire Bordelais  
de Recherche  
en Informatique

# Exemples de détection

---



06/04/2012



Laboratoire Bordelais  
de Recherche  
en Informatique

# Régions de calcul du descripteur.

Calcul de l'orientation à la base des Ondelettes  
de Haar

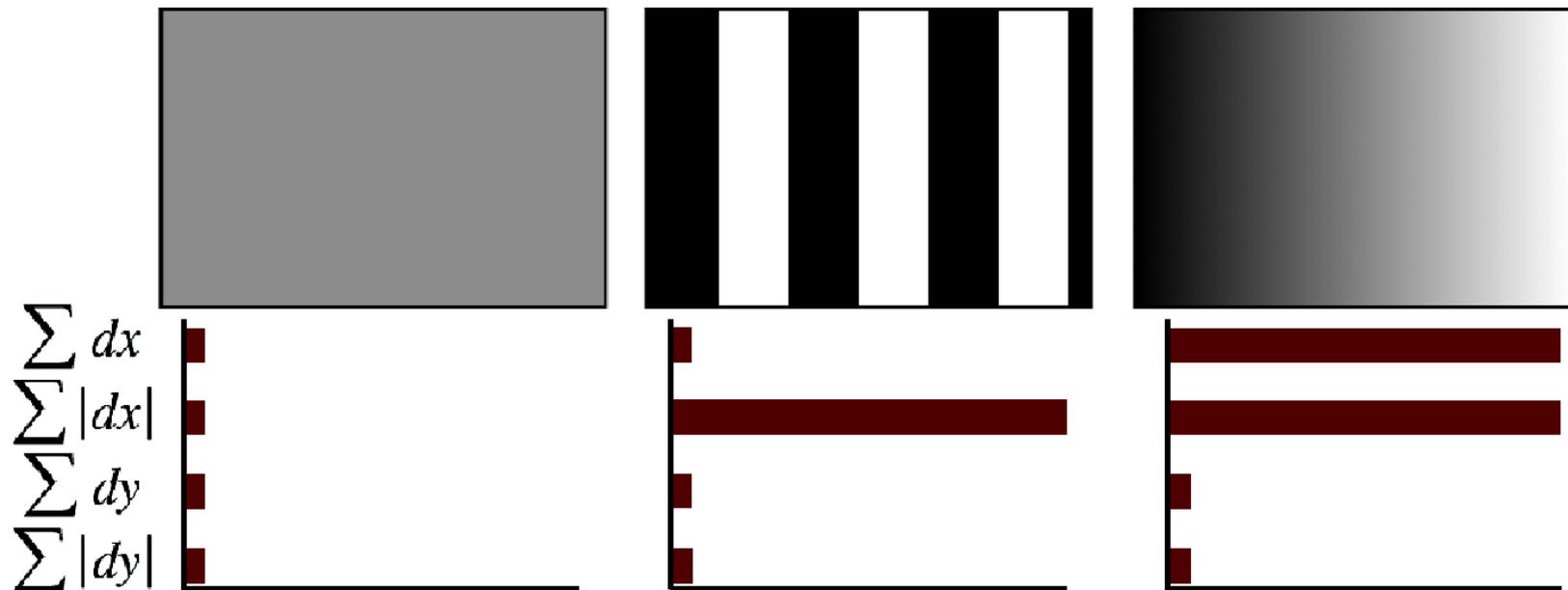


06/04/2012



Laboratoire Bordelais  
de Recherche  
en Informatique

# Descripteur





## Quantification: BOF et BOVW

---

- Notre problème 2 : comment comparer les images représentées par leurs descripteurs locaux
- BOF : représentation d'une image par un ensemble des descripteurs d'origine
- BOVW : quantification vectorielle des descripteurs et représentation d'une image par signature-histogramme des « mots » du dictionnaire.



Laboratoire Bordelais  
de Recherche  
en Informatique

# Approche BagofFeatures (BoF)



D'après Lazebnik (UNC/UIUC)

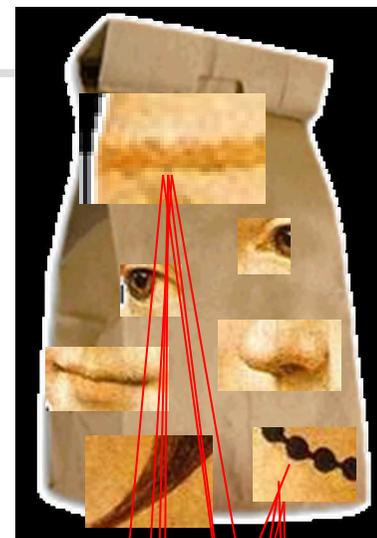


# Comparaison des images

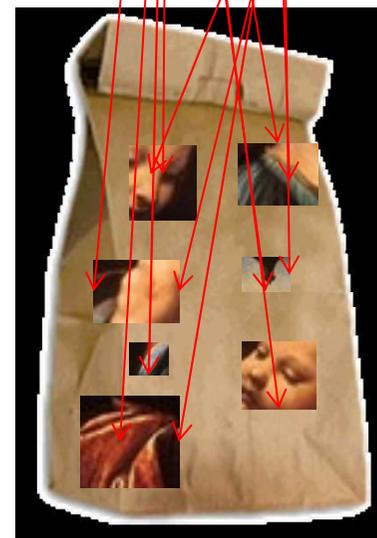
Laboratoire  
de Recherche  
en Informatique



Score(i,j)



$d(F_{k,i}, F_{l,j})$





Laboratoire Bordelais  
de Recherche  
en Informatique

# Recherche des images similaires

---

- Mesure de similarité entre les descripteurs  
(SIFT : distance L1, histogrammes – distance L1, coeff de Battacharya, Ki-2,..., mesures assymetriques)
- **Score** : nombre de descripteurs similaires au sens de la mesure de similarité choisie



## Recherche des images similaires

---

- Soit Image  $I_i$  représentée par le BoF (« sac »)  $B_i$ , composé de  $S$  vecteurs non-ordonnés  $b_{si} \in R^n$ :  $B_j = \{b_{si}\}$
- Soit DB – la base de données des images et  $F$  la base des vecteurs-descripteurs.
- Soit  $I_q$  – l'image requête représentée par le BoF  $B_q = \{b_{pq}\}$  avec  $P$  vecteurs.
- Denotons  $\text{score}_i = \text{score}(I_i, I_q)$  le score entre les images  $I_q$  et  $I_i$ .



## Recherche des images similaires

---

- Ordonner toutes les images de la base par les valeurs décroissantes du score  $i$ .

- Calcul du score :

- $score_i = \sum_{b_{pq} \in B_q} \sum_{b_{si} \in B_i} sim(b_{pq}, b_{si})$



## Stratégie de vote(1)

---

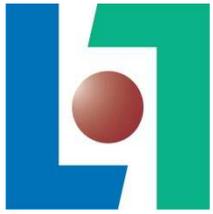
- Dénotons par

$$sim_L = \prod_{L(x,y)}$$

fonction-indicatrice relativement à la «similarité de vote» qui prend la valeur 1 si le prédicat  $L$  est « vrai » et 0 sinon.

- **Le principe de la méthode de vote :**

Chercher dans la base de descripteurs  $F$  les vecteurs les plus proches à l'ensemble de  $P$  vecteurs  $b_{pq}$  de l'image-requête  $I_q$ . Pour chaque vecteur  $b_{si}$  de l'image  $I_i$  proche au sense  $L$  d'un des vecteurs  $b_{pq}$  incrémenter  $score_i$



Laboratoire Bordelais  
de Recherche  
en Informatique

## Stratégie de vote(2)

- Proximité : kNN (kppv), RNN (Rppv)
- kNN : x et y sont similaires si y se trouve parmi les K plus proches voisins de x
- (K est le paramètre de l'algorithme)
- $sim(x, y) = sim_{kNN}(x, y) = \prod_{x \in kNN(y)}$



**K=7** : inconvenients? Non-symétrique



## Stratégie de vote (3)

- RNN (ppv)

$$\text{sim}(x, y) = \text{sim}_{RNN}(x, y) = \prod_{d(x,y) \leq R}$$





## Stratégie de vote

---

- Calcul de « bonnes » correspondances uniquement en considérant toutes les correspondances équivalentes.
- Nécessité d'une similarité plus « sémantique »
- Difficultés de « passage en échelle » - une grande quantité des appariements.



# Approche par Bag of Words BoVW

---

- Approches inspirées par
  - la recherche dans des bases de données textuelles
  - (R. Baeza –Yates, B. Ribeiro-Neto, Modern Information Retrieval, ACM Press, 1999)
  - la quantification visuelle (codage) (Lindo Buso Gray)

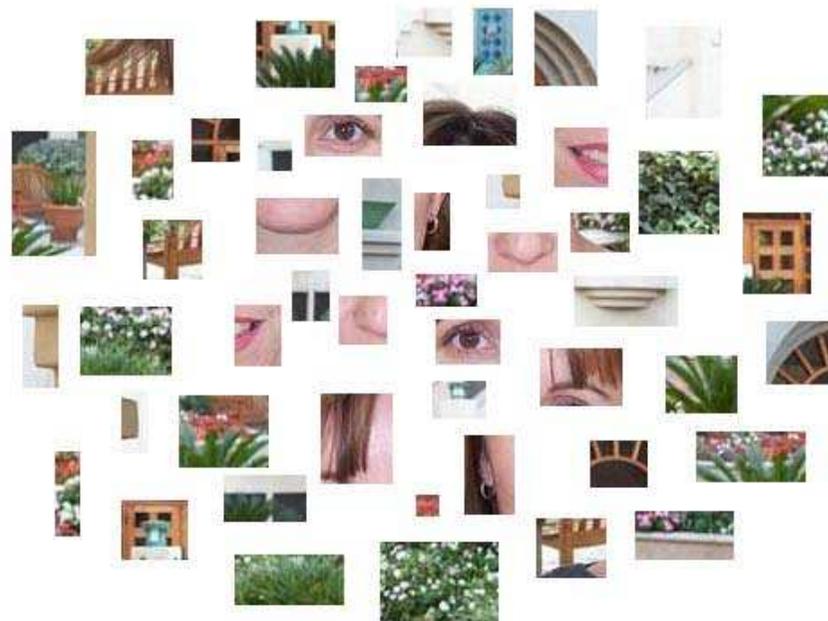
## « Dictionnaires visuels »

J. Sivic and A. Zisserman, "Video google: a text retrieval approach to object matching in videos," ICCV'2003, vol. 2, pp. 1470-1477, 2003.



Laboratoire Bordelais  
de Recherche  
en I

## BoW(1)



- « visage », »herbe », « bâtiment » ????

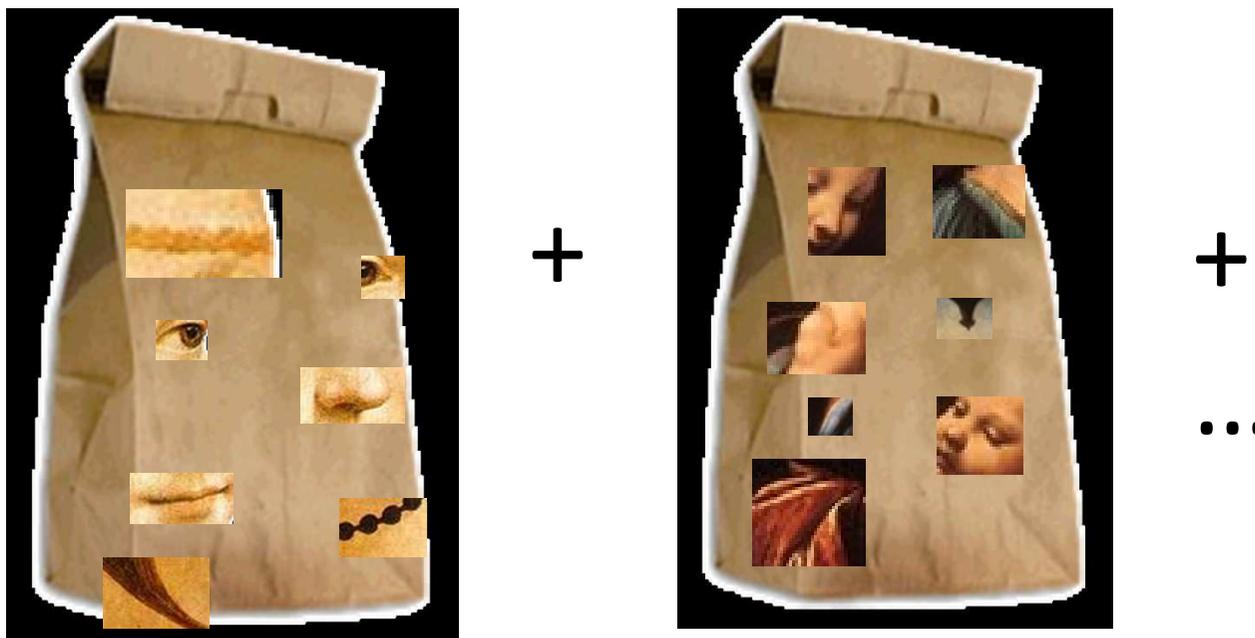
Lazebnik (UNC) conférence sur BoF, proposé par G.  
Szurka en 2004



Laboratoire Bordelais  
de Recherche  
en Informatique

## BoW(2)

- 1. Extraction des descripteurs des éléments
- 2. Construction de l'ensemble des descripteurs





## BoW(3)

---

3. Construire le dictionnaire visuel  $C$  – classification non-supervisée (clustering)

- $Card(C) \ll Card(B_1 \cup B_2 \cup \dots \cup B_n)$  pour  $n$  images dans la base

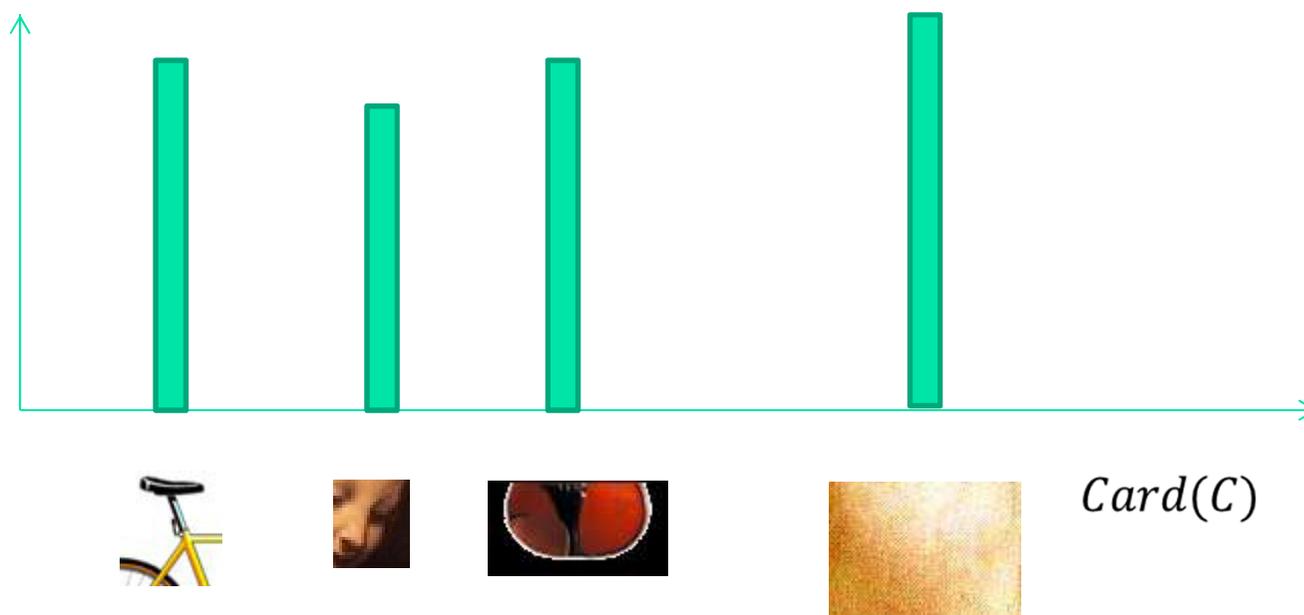
4. Quantifier tous les descripteurs par rapport au dictionnaire :

$$\forall b_{ij} \rightarrow c_{k^*} : d(b_{ij}, c_{k^*}) = \min_k \left( d(b_{ij}, c_k) \right), \\ k = 1, \dots, Card(C)$$



## BoW(4)

- 5. Pour toutes les images  $I_i$  de la base calculer l'histogramme (normalisé) (*BoW*)  $H_i$  des mots  $c_k$  du dictionnaire  $C$



- Le nouvel espace de représentation  $R^{Card(C)}$



Laboratoire Bordelais  
de Recherche  
en Informatique

## BoW(5)

---

- Tailles des dictionnaires :  $Card(C) \approx 10^3$
- Comparaison des histogrammes : kNN, RNN,
- tf-idf

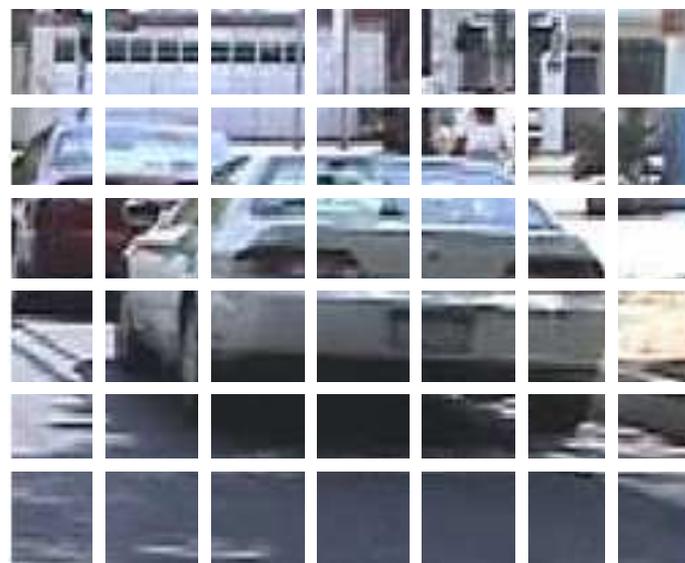


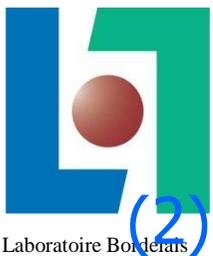
Laboratoire Bordelais  
de Recherche  
en Informatique

# Choix des éléments de l'image et des descripteurs(1)

---

- Vogel & Schiele, 2003
- Fei-Fei & Perona, 2005
- Quenot et Ayache 2005,
- Merialdo, Dumont 2008





Laboratoire Bordelais  
de Recherche  
en Informatique

(2)

## Choix des éléments de l'image et des descripteurs

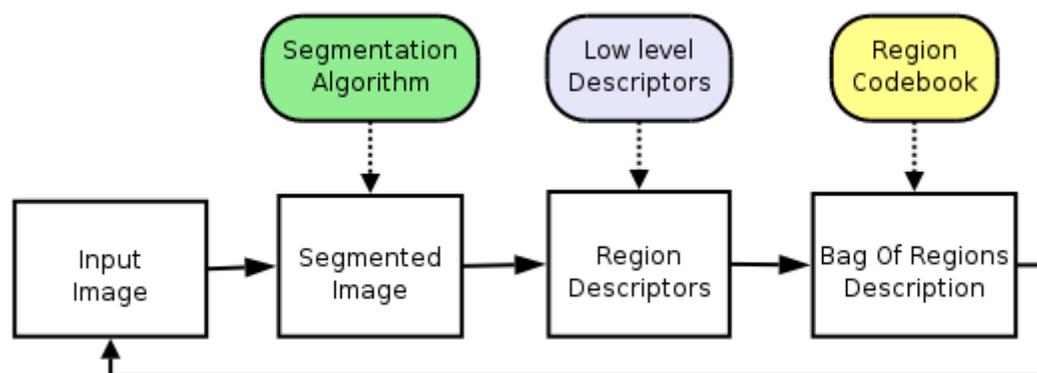
Sivic &  
Zisserman  
2005



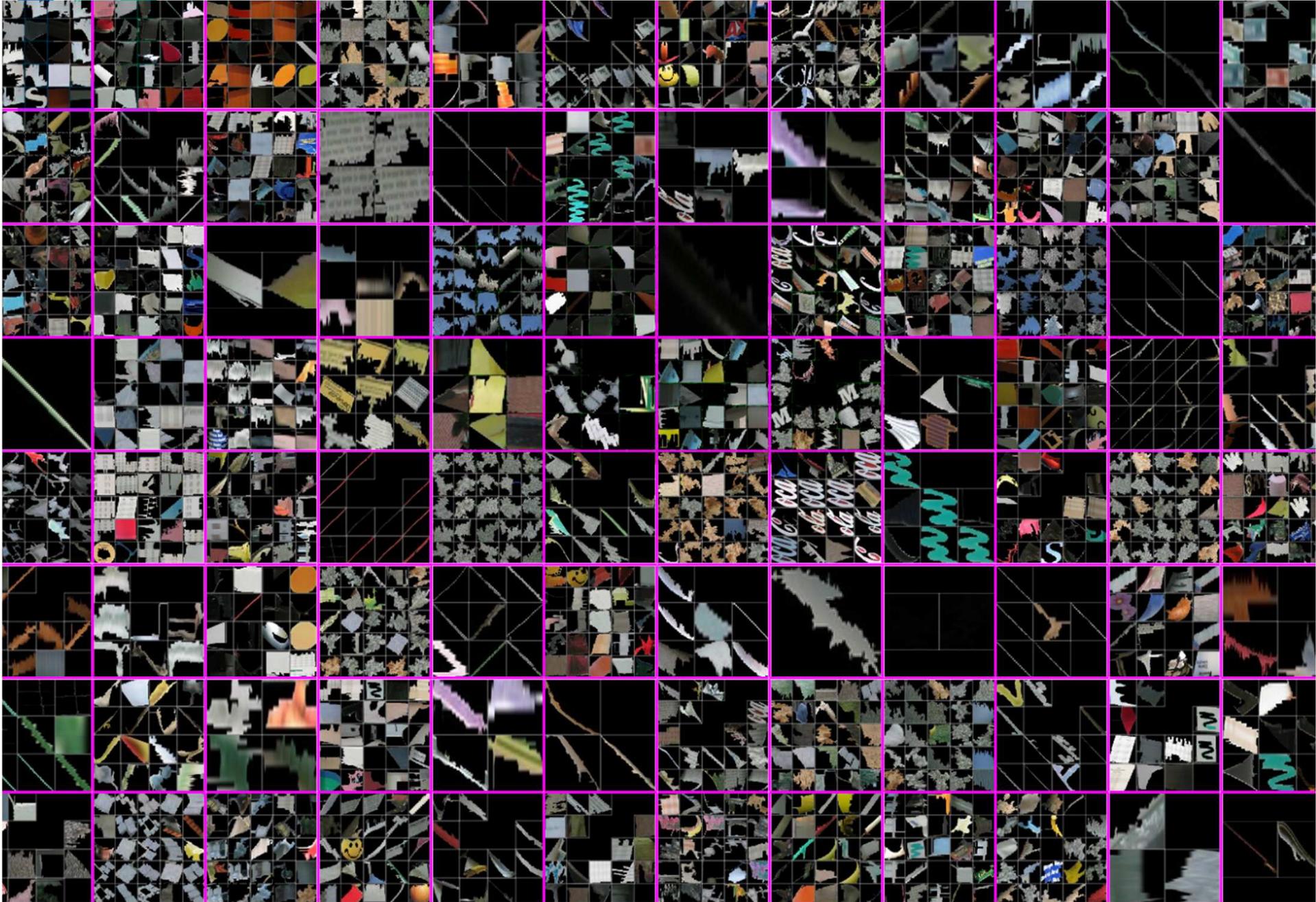


# BORW – Bag of Region-Words

- Principe : les régions sont issues de la fusion des segmentations régulière et irrégulière.



- Descripteurs : histogramme HSV
- Wang DB





## TF-IDF(1)

### • Pondération des BoW

- $v_i = (t_1, t_2, \dots, t_{Card(C)})^T$
- $t_k = h_k \log \frac{N}{N_k}$
- $h_k = \frac{n_{ik}}{n_i}$  - “term frequency”
- $N$  – est le nombre des images dans la base
- $N_k$  – est le nombre des images contenant le mot  $c_k$
- $\log \frac{N}{N_k}$  - est la fréquence “inverse” - IDF



## TF-IDF(2)

---

- L'IDF diminue le poids des mots qui sont rencontrés souvent dans la DB et sont peu discriminants

- Comparaison :

- $sim(v_q, v_i) = \frac{v_q v_i}{\|v_q\|_2 \|v_i\|_2}$  - distance de cos



Laboratoire Bordelais  
de Recherche  
en Informatique

# Méthodes de construction de dictionnaire

- Algorithmes de groupement de données : clustering(CNS)
- Le plus fréquemment utilisé est l'algorithme des  $K$ -

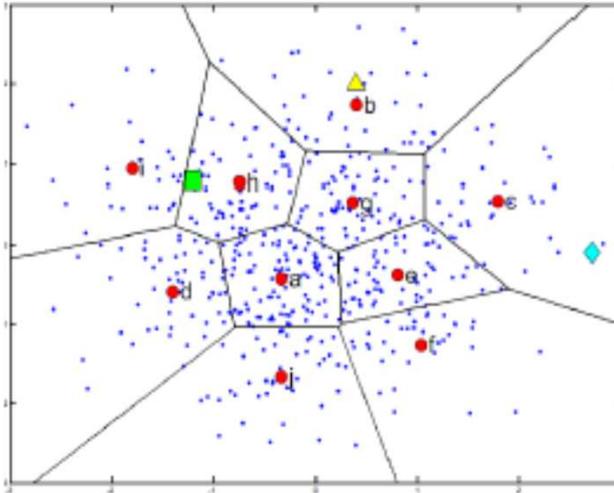


Illustration d'ambiguïté des  
mots visuels

- Problème de passage à l'échelle



# Algorithme de K-moyennes(I)

- J. MacQueen, "Some methods for classification and analysis of multivariate observations", Proc. Of the Fifth Berkley Symposium on Math. Stat. And Prob., pp. 281 – 296, 1967
- **Principe** : CNS avec le nombre des classes connus a priori.
- **Paramètre** : le nombre  $k$  de classes
- **entrée** : un échantillon de  $M$  vecteurs-descripteurs  $x_1, \dots, x_M$ .
- (1) Choisir  $k$  centres initiaux  $c_1, \dots, c_k$
- (2). Pour chacun des  $M$  vecteurs, l'affecter à la classe  $i$  dont le centre  $c_i$  est le plus proche
- (3). Si aucun élément ne change de classe alors arrêt.
- (4). Calculer les nouveaux centres : pour tout  $i$ ,  $c_i$  est la moyenne des éléments de la classe  $i$
- (5). Aller en 2



# Clustering incrémental

---

- 1. Choisir les vecteurs initiaux pour former  $k$  centres des clusters,  $c_k, k = 1, \dots, K$ .
- 2. Sélectionner la donnée suivante  $x$  de l'ensemble de données  $D$
- 3. Calculer la distance de  $x$  à  $K$  centres des clusters.
- 4. Trouver le centre du cluster le plus proche
- $c_{win} = \arg \min_k d(x, c_k)$
- 5. Si  $d(x, c_{win}) > Th$  alors l'échantillon  $x$  devient le centre du nouveau cluster. Sinon bouger le centre  $c_{win}$  dans la direction de  $x$  :
- $c_{newwin} = c_{oldwin} + e^*(x - c_{oldwin}); \quad 0 \leq e < 1$



# Classification hiérarchique agglomérative

## ■ Principes :

- (1) A l'initialisation chaque vecteur descripteur dans l'échantillon de données forme une classe.
- (2) Tant que le nombre de classes est supérieur à  $k$  ( à la limite  $k=1$ )
  - Regrouper les classes les plus proches au sens d'une distance  $d$

Distance entre les classes

lien minimal  $d_{\max}(C_i, C_j) = \max_{x \in C_i, y \in C_j} d(x, y)$

lien maximal  $d_{\min}(C_i, C_j) = \min_{x \in C_i, y \in C_j} d(x, y)$

distance moyenne

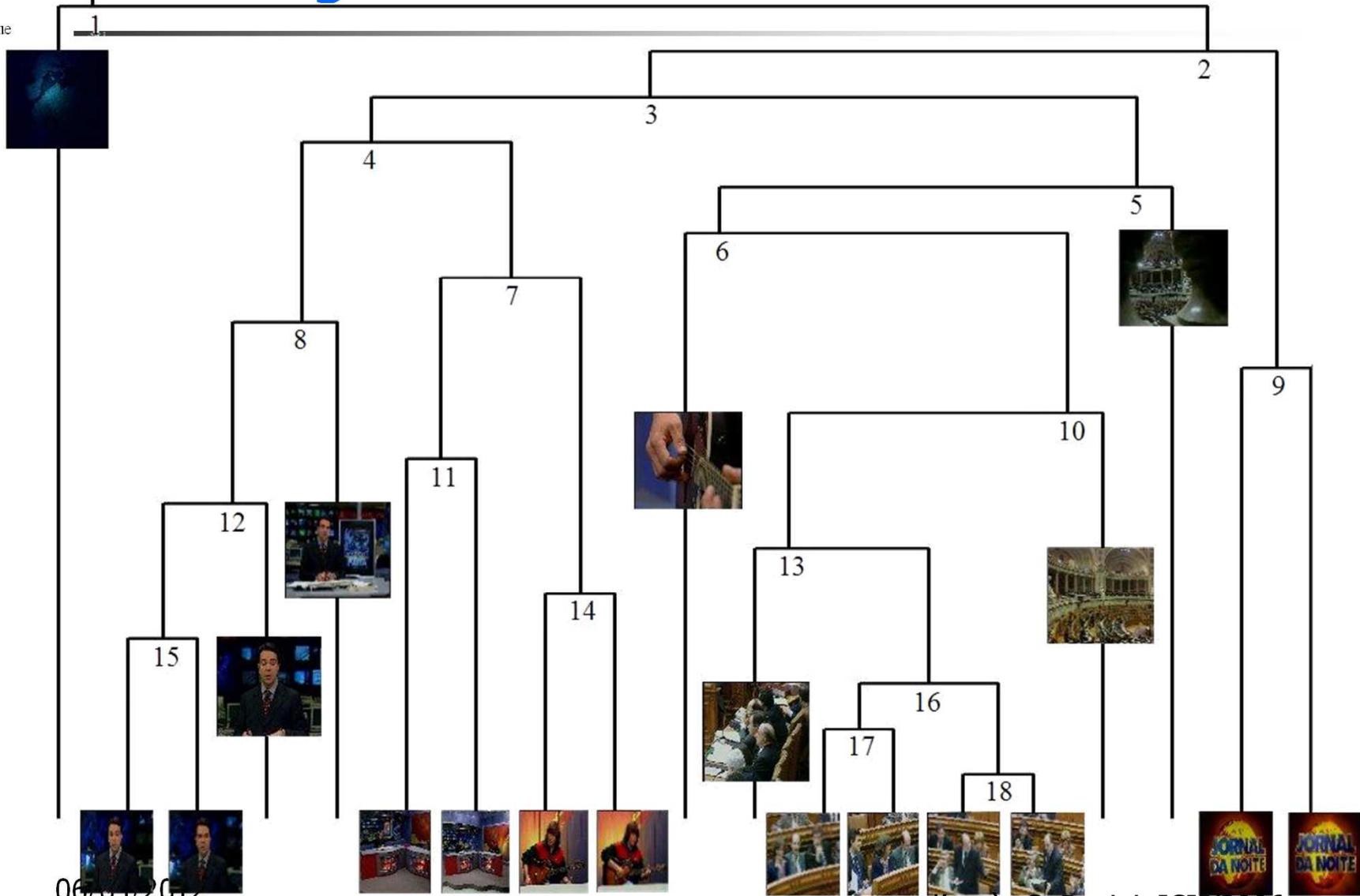
$$d_{\text{moy}}(C_i, C_j) = \frac{1}{n_i \times n_j} \sum_{l=1}^{l=n_i} \sum_{p=1}^{p=n_j} d(x_l, y_p)$$

d'après A. Laurent, LIRMM



Laboratoire Bordelais  
de Recherche  
en Informatique

# Dendrogramme



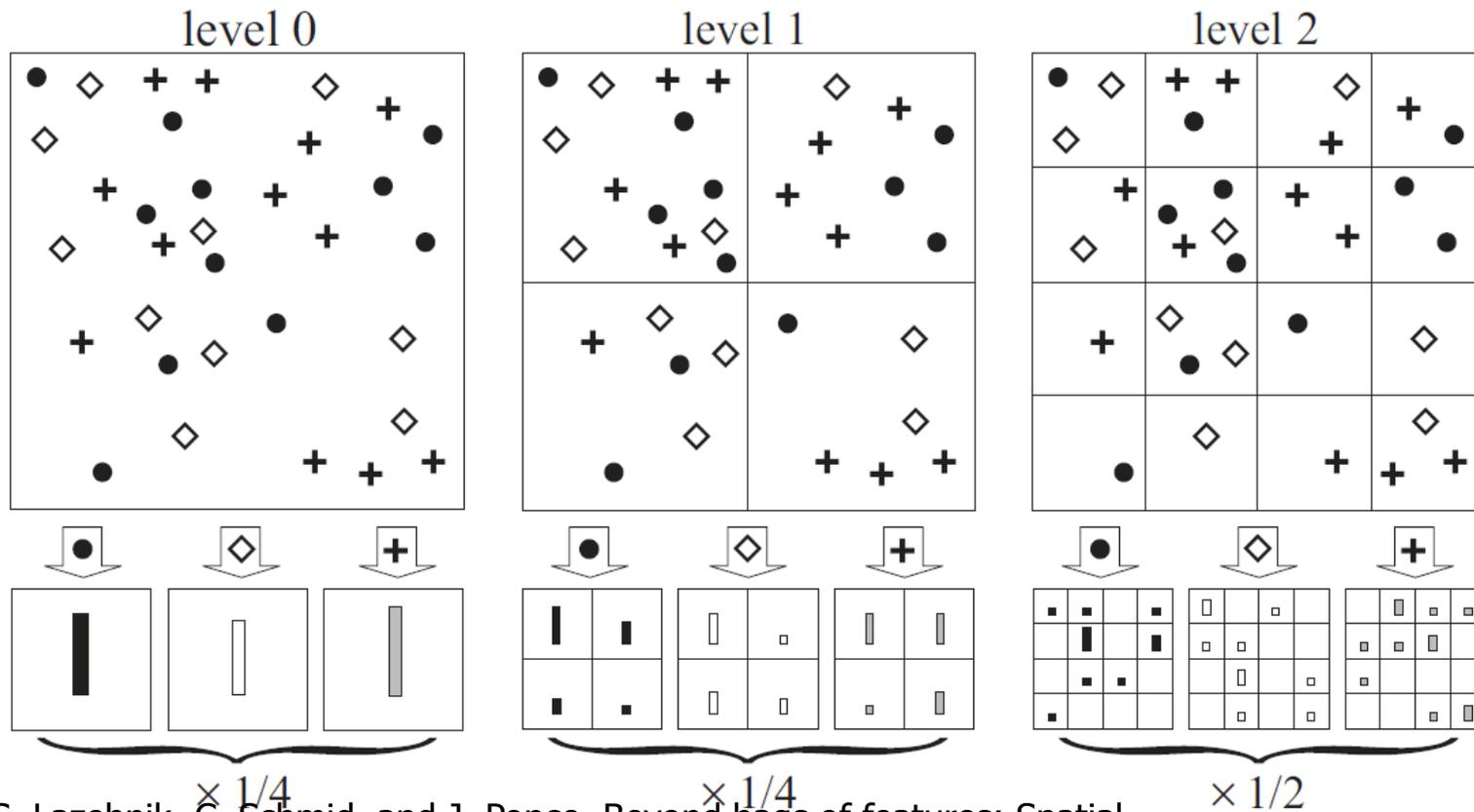
06/01/2012

d'après S. Benini, ICIP'2006



# Prise en compte du contexte spatial: SPMK

- La description éparsée ne tient pas compte de la distribution des « features » dans le plan-image



S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2169–2178. IEEE, 2006.



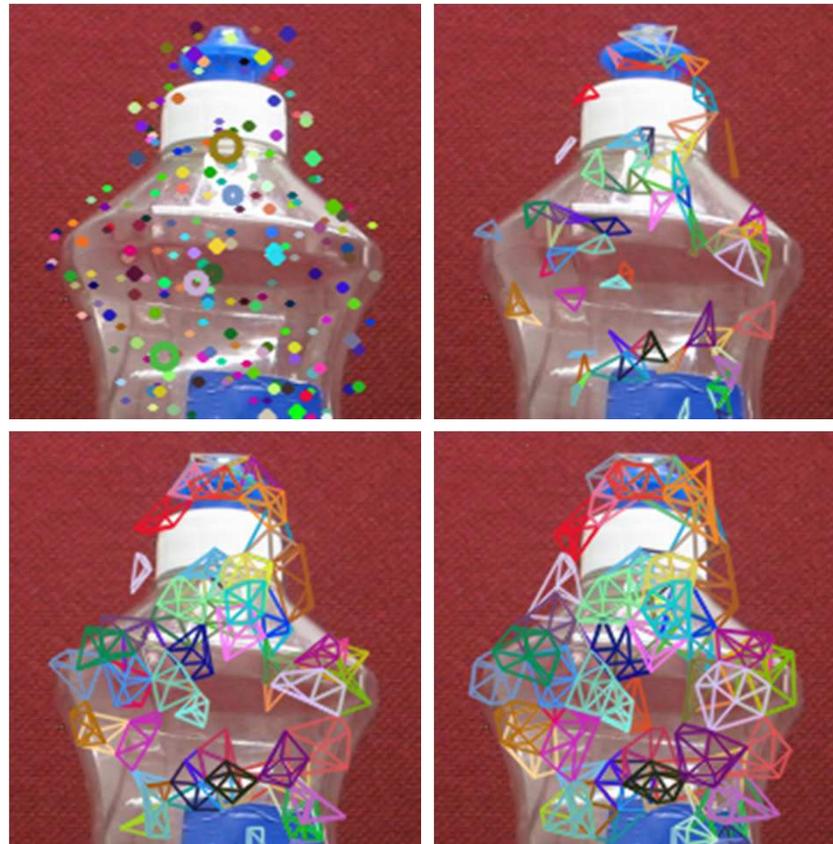
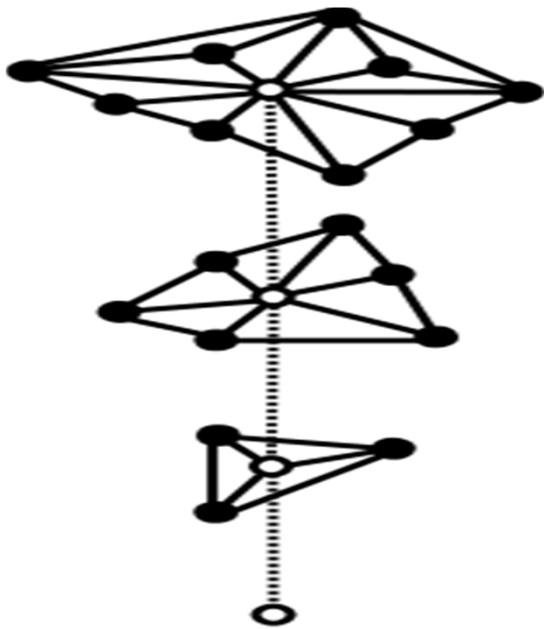
- Similarité : Spatial Pyramid Match Kernel.
- Considérons images  $X$  et  $Y$
- Considérons  $m=1, \dots, M$  classes dans le dictionnaire;
- Considérons  $I$  – la mesure d'intersection des histogrammes,
- Considérons  $l=0, \dots, L$  les partitions de l'image par des grilles emboîtées avec les cellules  $\frac{1}{2^l} N \times \frac{1}{2^l} N$  ( pour simplifier  $H=L=N$ )
- $$k^L(X_m, Y_m) = \frac{1}{2^L} I(H^0(X_m), H^0(Y_m)) + \sum_{l=1}^L \frac{1}{2^{L-l+1}} I(H^l(X_m), H^l(Y_m))$$
- $$K^L(X, Y) = \sum_{m=1}^M k^L(X_m, Y_m)$$
- Ceci revient à concatener les histogrammes pour tous les niveaux et pour toutes les cellules avec la pondération.
- Inconvénient : non-invariance par rapport aux transformations affines du plan –image.



Laboratoire Bordelais  
de Recherche  
en Informatique

# GraphWords

- Construction de graphes par triangulation de Delaunay sur des points SURF





# Problème de clustering des graphes

- (1) Distance:
  - $d(P, Q) = s(P, P) + s(Q, Q) - 2s(P, Q) \in [0, 1]$
  - Nécessité de définir la similarité  $s$
- (2) Méthode de quantification/classification
  - HAG car impossible d'interpoler les graphes – espace non-vectoriel
  - Modèle de classe : graphe médian

$$\text{median} = \operatorname{argmin}_{G \in \mathcal{C}} \sum_{i=1}^m \|c_i - G\|$$



Laboratoire Bordelais  
de Recherche  
en Informatique

# Exemples

Base de données SIVAL (25 objets, 10 environnements, 6 positions par environnements)



(a) DirtyWork-Gloves



(b) Striper-Notebook



(c) Glazed-WoodPot



(d) GoldMedal



(e) DirtyRun-ningShoe

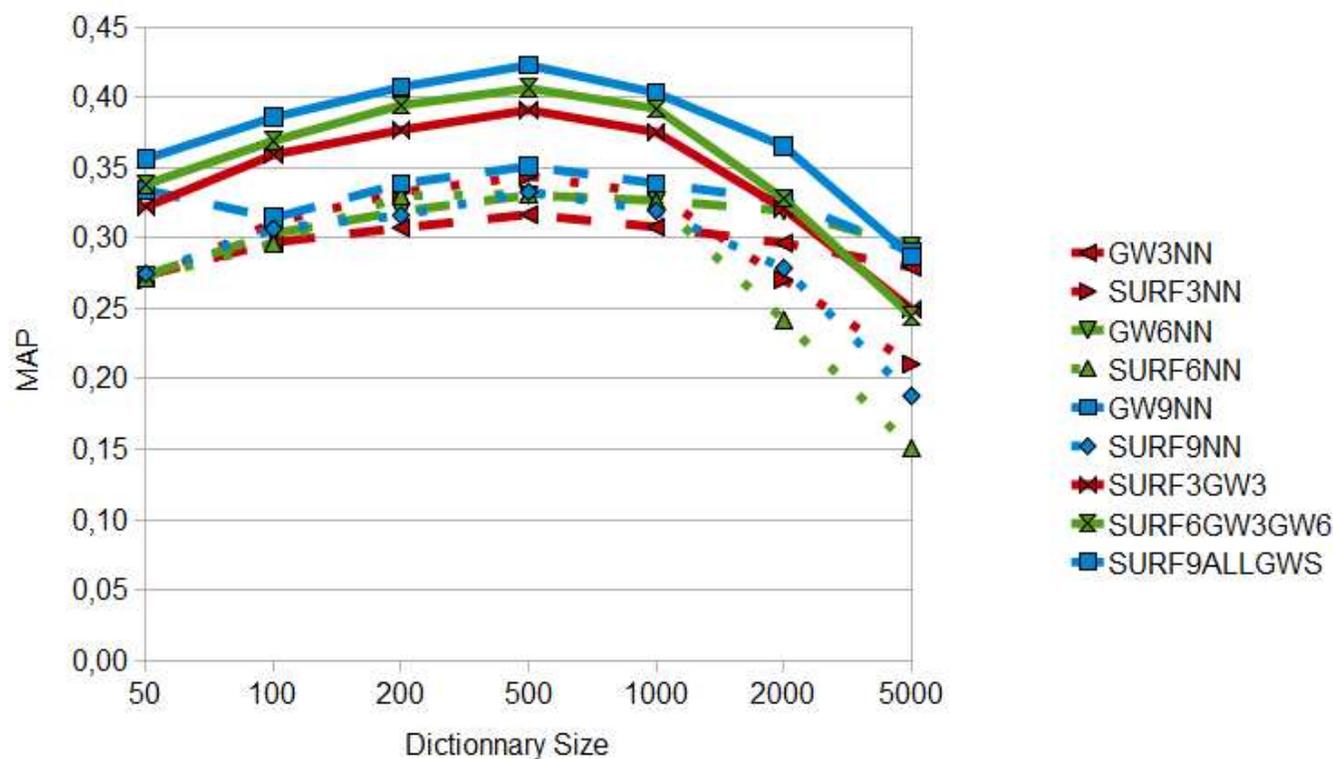


06/04/2012



## Évaluation de la méthode

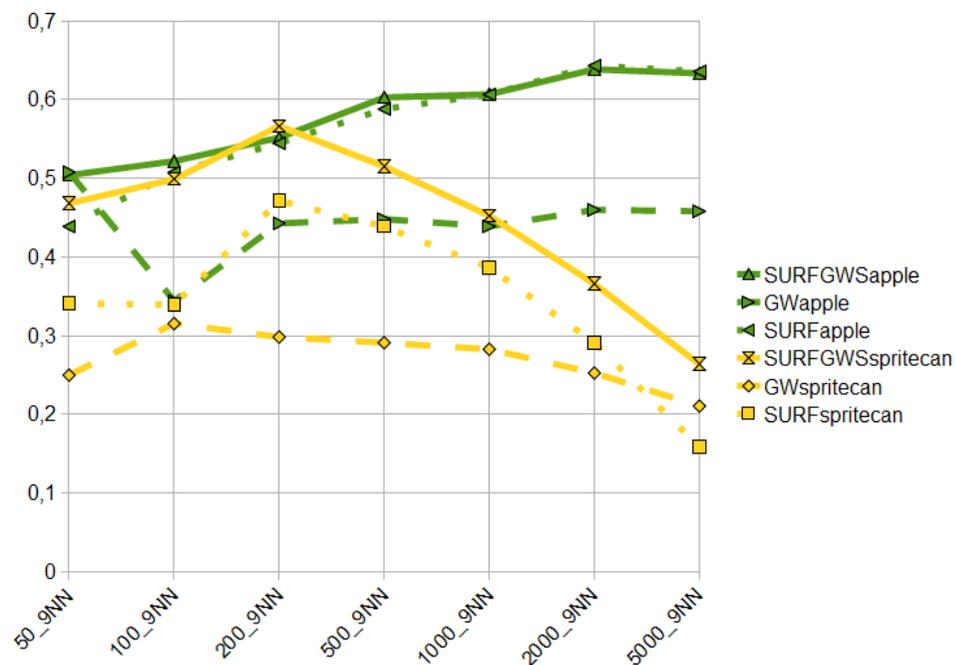
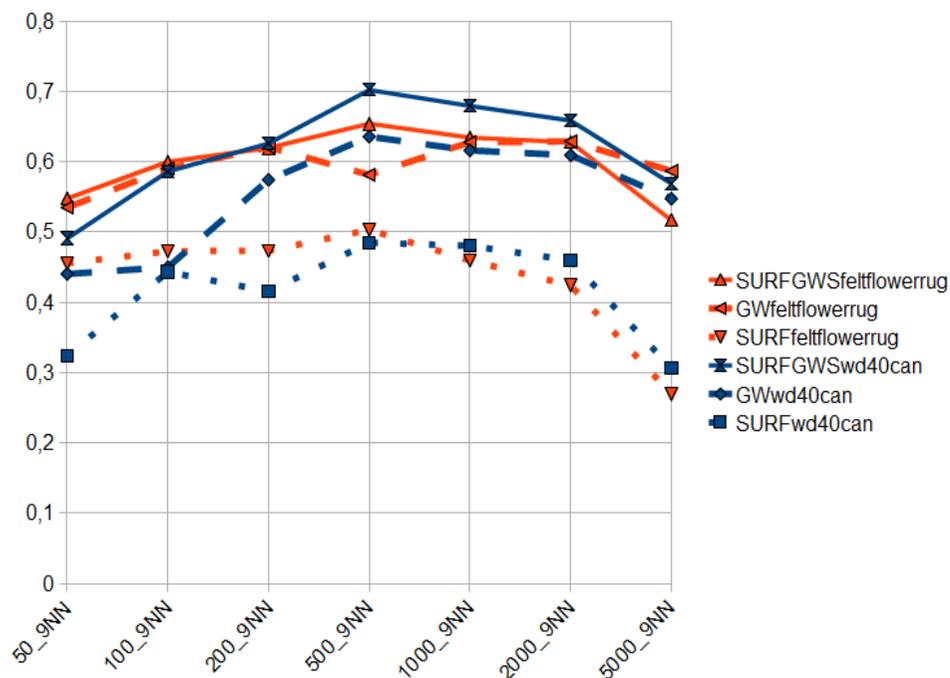
- Base de données SIVAL (25 objets, 10 environnements, 6 positions par environnements)





# Évaluation de la méthode

- Fusion (continu) toujours meilleure que graphes (tirés) ou points SURF isolés (pointillés)
- [Soumis à ACM MM'2011]





Laboratoire Bordelais  
de Recherche  
en Informatique

## 2. Extensions spatio-temporelles Comparaisons des séquences vidéo(1)

- (1) Key-framing



«Chancres» - CERIMES

$$\text{sim}(V_q, V_j) = \prod_{(I_q, I_j)}$$





## Comparaisons des séquences vidéo(2)

---

- Soit  $V_q$  – clip requête et  $V_j \in DB$

- 1. Soit  $I_{qi} \in V_q$  - une image clé,

- $score_j = \max_{I_{jk} \in V_j} (score(I_{qi}, I_{jk}))$

- 2.

$$score_j = \max_{I_{qi} \in V_q} \max_{I_{jk} \in V_j} (score(I_{qi}, I_{jk}))$$



## Mouvement : descripteur particulier

---

- Représentation du mouvement dans le plan - image
- Modèles du mouvement
- Méthodes d'estimation
  - Méthodes directes / Estimation par bloc
  - Méthodes paramétriques/Estimation robuste
- Méthodes de la segmentation basées mouvement
  - Méthodes par comparaison du mouvement



Laboratoire Bordelais  
de Recherche  
en Informatique

# Représentation du mouvement dans le plan - image

- Une séquence vidéo est une image 2D du monde 3D en mouvement
- On ne perçoit le mouvement que grâce au changement de la luminance / couleur



06/04/2012

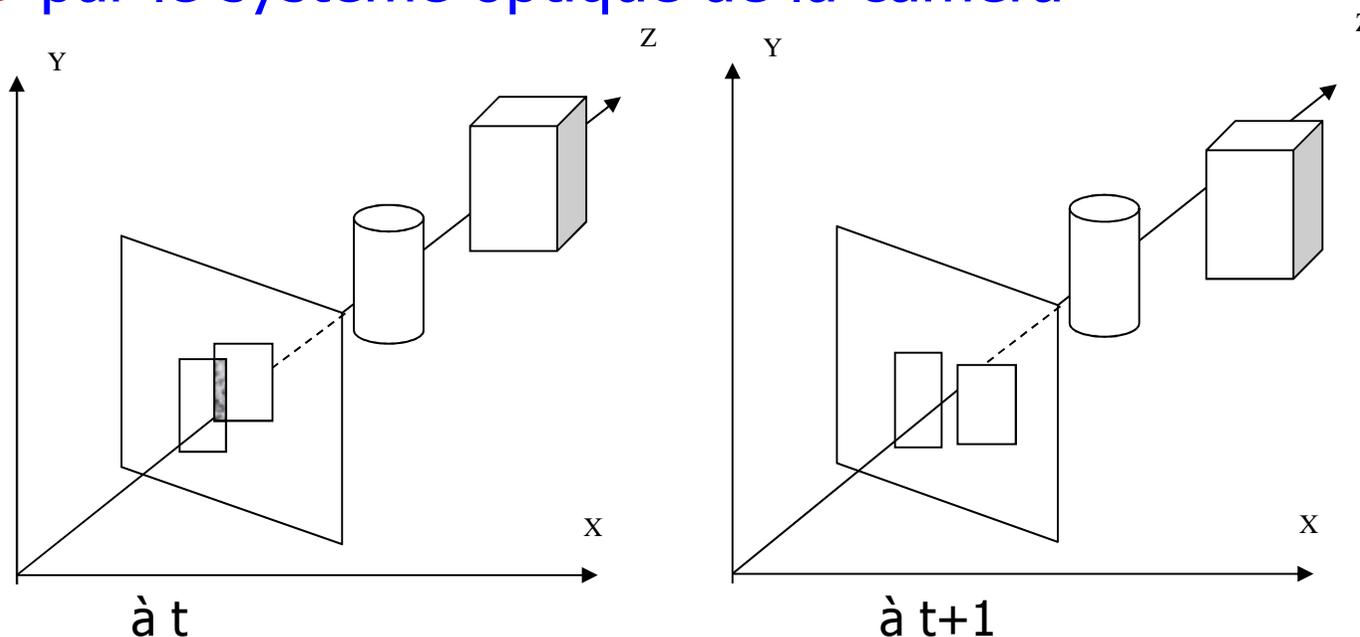
Mouvement Apparent



Laboratoire Bordelais  
de Recherche  
en Informatique

# Mouvement réel 2D vs mouvement apparent

Mouvement réel 2D est la projection du mouvement 3D par le système optique de la caméra



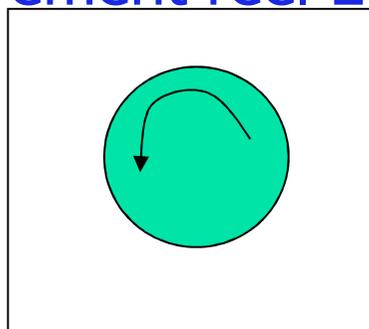
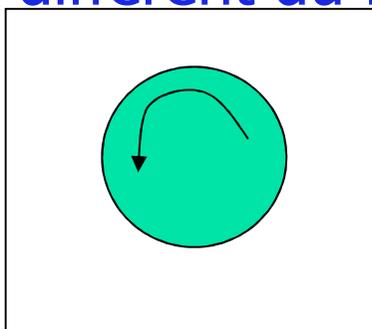
Mouvement apparent \_ "flot optique" est observé dans le plan image 2D grâce au changements de la luminance



Laboratoire Bordelais  
de Recherche  
en Informatique

# Mouvement réel 2D vs mouvement apparent (2)

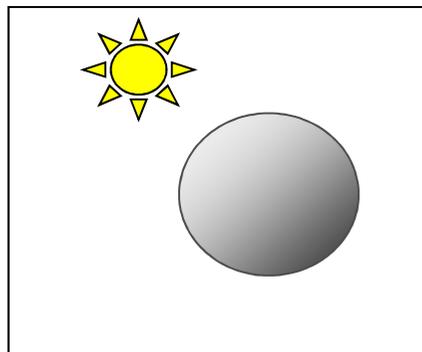
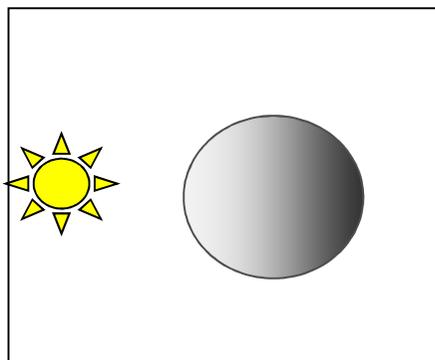
Mouvement apparent est dans le cas général différent du mouvement réel 2D



a) Insuffisance du gradient spatial

MR - oui

MA - non



b) Changements d'illumination extérieure

MR - non

MA - oui

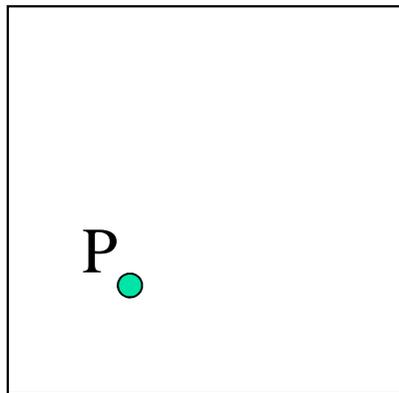
Néanmoins!

Hypothèse: Mouvement Apparent=Mouvement reel 2D

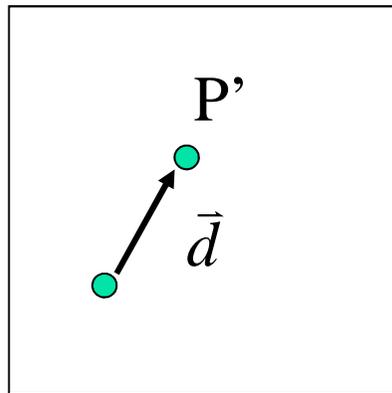
06/04/2012



# Caractérisation locale du mouvement



t



t+1

$$\vec{d} = (dx, dy)^T$$

Vecteur de déplacement  
élémentaire

$$\vec{w} = (dx/dt, dy/dt)^T$$

vecteur vitesse

Premier niveau de caractérisation du mouvement  
consiste à calculer le flot optique  $W = \{\vec{w}\}_\Omega$  ou  $D = \{\vec{d}\}_\Omega$   
"champ de déplacement"



# Modèles de mouvement

En développant  $\vec{d} = (dx, dy)^T$  en série de Taylor  
autour de  $(x_g, y_g)$  jusqu'au 1er ordre

$$\vec{d} = \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \underbrace{\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}}_M \begin{pmatrix} x - x_g \\ y - y_g \end{pmatrix} \quad (9)$$

Ici

$$a_1 = \frac{\partial dx}{\partial x} \quad a_2 = \frac{\partial dx}{\partial y} \quad b_1 = \frac{\partial dy}{\partial x} \quad b_2 = \frac{\partial dy}{\partial y}$$

$\Theta_{aff} = (a_0, b_0, a_1, a_2, b_1, b_2)^T$  **Modèle affine à 6-paramètres**



# Modèles affines

## Exprimant

$$M = \frac{1}{2}(\text{trace}M)I + \frac{1}{2}(M - M^T) + \frac{1}{2}[M + M^T - (\text{trace}M)I]$$

$$\text{trace}M = a_1 + b_2 = \frac{\partial dx}{\partial x} + \frac{\partial dy}{\partial y} = \text{div}(\vec{d})$$

$$\frac{1}{2}(M - M^T) = \frac{1}{2} \underbrace{\begin{pmatrix} \frac{\partial dy}{\partial x} & \frac{\partial dx}{\partial y} \\ \frac{\partial dx}{\partial x} & \frac{\partial dy}{\partial y} \end{pmatrix}}_{\text{rot}_z(\vec{d})} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$\frac{1}{2}[M + M^T - (\text{trace}M)I] = \begin{pmatrix} -b_2 + a_1 & b_1 + a_2 \\ b_1 + a_2 & b_2 - a_1 \end{pmatrix}$$

$$\text{hyp1} = \frac{\partial dx}{\partial x} - \frac{\partial dy}{\partial y} \qquad \text{hyp2} = \frac{\partial dy}{\partial x} + \frac{\partial dx}{\partial y}$$

$$M = \frac{1}{2} \text{div} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \text{rot} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \frac{1}{2} \text{hyp1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2} \text{hyp2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

# Hiérarchie des modèles affines

$$\begin{cases} dx = t_x \\ dy = t_y \end{cases}$$

$$\begin{cases} dx = t_x + k(x - x_g) \\ dy = t_y + k(y - y_g) \end{cases}$$

$$\begin{cases} dx = x_2 - x_1 = t_x + k(x_1 - x_{g1}) - \theta(y_1 - y_{g1}) \\ dy = y_2 - y_1 = t_y + \theta(x_1 - x_{g1}) + k(y_1 - y_{g1}) \end{cases}$$

$$\begin{cases} dx = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} x - x_g \\ y - y_g \end{pmatrix} \end{cases}$$

$$\begin{cases} dx = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} div \cdot (x - x_g) - rot \cdot (y - y_g) + hyp1 \cdot (x - x_g) + hyp2 \cdot (y - y_g) \\ div \cdot (y - y_g) + rot \cdot (x - x_g) - hyp1 \cdot (y - y_g) + hyp2 \cdot (x - x_g) \end{pmatrix} \end{cases}$$

$$\begin{cases} dx = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} x - x_g \\ y - y_g \end{pmatrix} + \begin{pmatrix} a_3 & a_4 \\ b_3 & b_4 \end{pmatrix} \begin{pmatrix} (x - x_g)^2 \\ (y - y_g)^2 \end{pmatrix} + \begin{pmatrix} a_5 \\ b_5 \end{pmatrix} (x - x_g)(y - y_g) \end{cases}$$



Laboratoire Bordelais  
de Recherche  
en Informatique

# Méthodes d'estimation

---

Objectif : mesurer le mouvement apparent

Méthodes :- directes (estimation du flot optique)

- indirectes (paramétriques-estimation du modèle global)

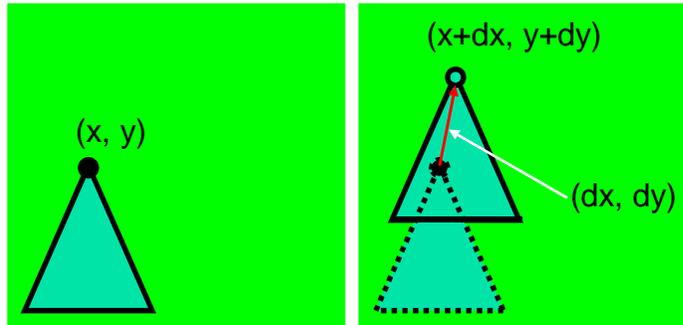
Modes : basé-pixel, basé-bloc, basé-région



# Estimation du mouvement(1)

**Hypothèse principale** : conservation de l'intensité lumineuse d'un point le long du trajectoire

$$DFD(x, y, \vec{d}) = I(x + dx, y + dy, t + dt) - I(x, y, t) = 0 \quad (12)$$



$DFD(x, y, \vec{d})$  n'est jamais nulle à cause du bruit et de changement d'éclairage



## Estimation du mouvement(2)

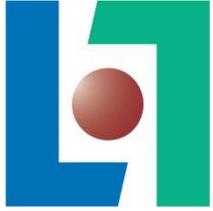
Critères à minimiser: EQM, MAD

$$EQM = \frac{1}{\|\Omega\|} \sum_{(x,y) \in \Omega} DFD^2(x, y, d(x, y)) \longrightarrow \min$$

$$MAD = \frac{1}{\|\Omega\|} \sum_{(x,y) \in \Omega} |DFD(x, y, d(x, y))| \longrightarrow \min$$

$$D^*(x, y) = \{\vec{d}(x, y)\} = \underset{\Omega}{\text{Arg min}} EQM(D) \quad \text{Estimation directe}$$

$$\vec{d}(x, y, \Theta) \rightarrow \Theta^* = \underset{\Omega}{\text{Arg min}} EQM(\Theta) \quad \text{Estimation paramétrique}$$



## Estimation du mouvement(3)

Développant  $I(x + dx, y + dy, t + dt)$  en série de Taylor  
autour de  $(x, y, t)$

et supposant la linéarité de  $I(x, y, t)$  on a

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t}$$

$$dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t} = 0 \quad \rightarrow$$

$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$  l'équation de contrainte  
du mouvement apparent (ECMA)



# Estimation du mouvement(4)

## Sous forme vectorielle

Comme  $\vec{w} = (\underbrace{dx/dt}_u, \underbrace{dy/dt}_v)^T$  alors

$$\vec{\nabla} I \cdot \vec{w} = -I_t \quad \text{ECMA}$$

Décomposons  $\vec{w} = \vec{w}_{\parallel} + \vec{w}_{\perp}$ ,  $\vec{w}_{\perp}$  est parallèle au gradient local  
 $\vec{\nabla} I \cdot (\underbrace{\vec{w}_{\parallel}}_{\text{est orthogonale}} + \vec{w}_{\perp}) = -I_t$

Estimation du mouvement est un problème mal posé.  
Uniquement le flot optique normal  $\vec{w}_{\perp}$  est observable



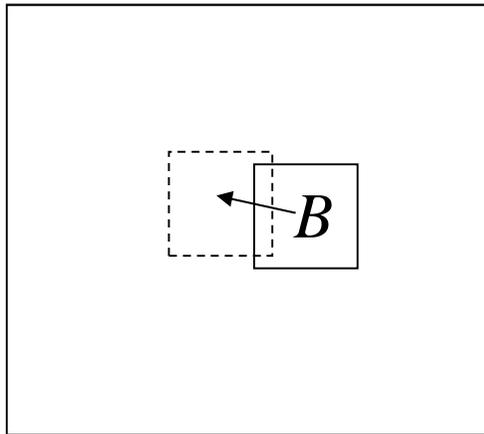
# Estimation par bloc

**Objectif :** obtenir le champ de déplacement  
éparse

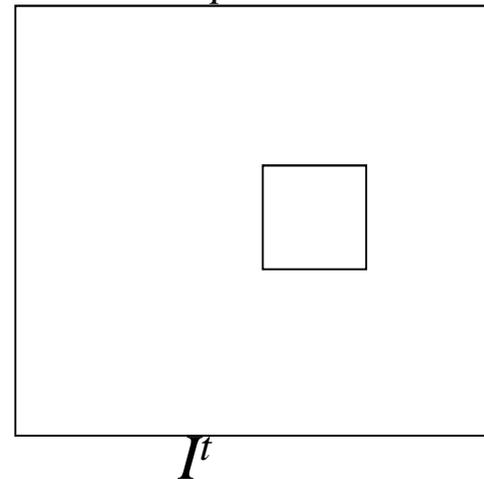
Le FO est supposé constant à l'intérieur d'un bloc

Le critère à minimiser :

$$\min_{\vec{p} \in B} |I(\vec{p}, t) - I(\vec{p} + \vec{d}, t - dt)|$$



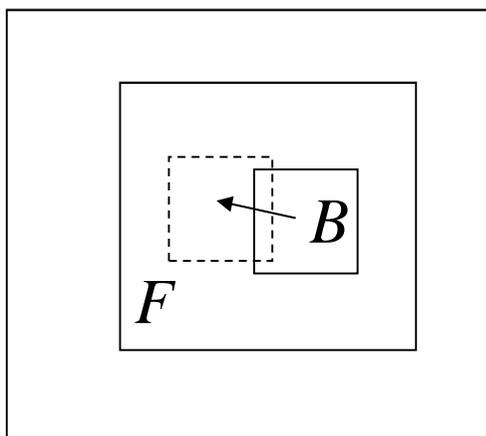
ou 
$$\min_{\vec{d} \in F} \sum_{\vec{p} \in B} (I(\vec{p}, t) - I(\vec{p} + \vec{d}, t - dt))^2$$



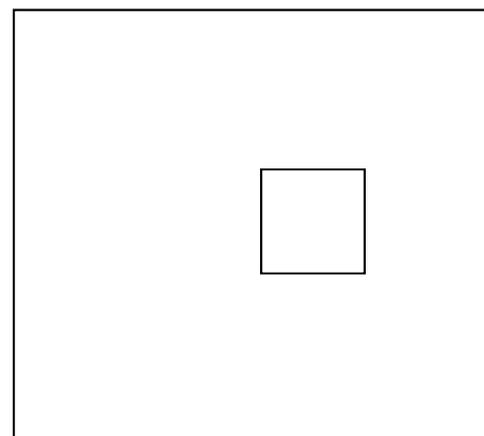


# Méthode de recherche exhaustive

## Estimation « au pixel près »



$I^{t-dt}$

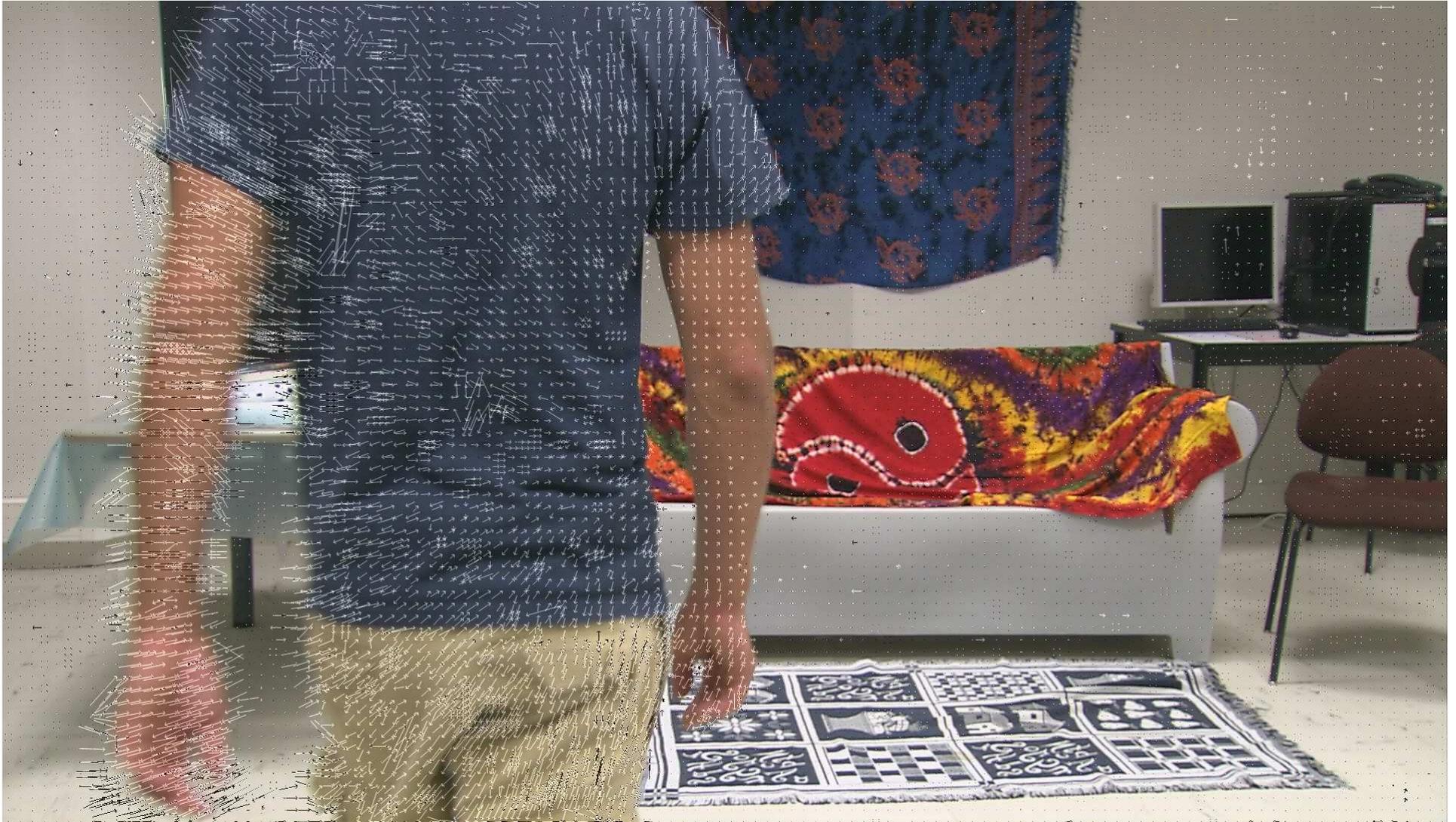


$I^t$

$$\hat{\vec{d}} = \underset{\vec{d} \in F}{\text{Argmin}} \sum_{\vec{p} \in B} |I(\vec{p}, t) - I(\vec{p} + \vec{d}, t - dt)|$$

L'inconvénient majeur : coût opératoire

Les estimateurs basés-blocs sont utilisés pour tous les standards du codage vidéo  
cf. UE « Codage Vidéo »



06/04/2012



# Méthodes de Flot Optique

## Méthode de Horn et Schunk

$$E(u, v) = \int_G (I_x u + I_y v + I_t)^2 + \alpha^2 \left( \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial v}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right) dx dy \rightarrow \min$$

$$\hat{u}^{(i+1)} = \bar{u}^{(i)} - \frac{\partial I}{\partial x} \frac{\vec{\nabla} I \cdot \bar{w}^{(i)} + \frac{\partial I}{\partial t}}{\alpha^2 + \|\vec{\nabla} I\|^2} \quad \hat{v}^{(i+1)} = \bar{v}^{(i)} - \frac{\partial I}{\partial y} \frac{\vec{\nabla} I \cdot \bar{w}^{(i)} + \frac{\partial I}{\partial t}}{\alpha^2 + \|\vec{\nabla} I\|^2}$$

## Méthodes de type "descente"

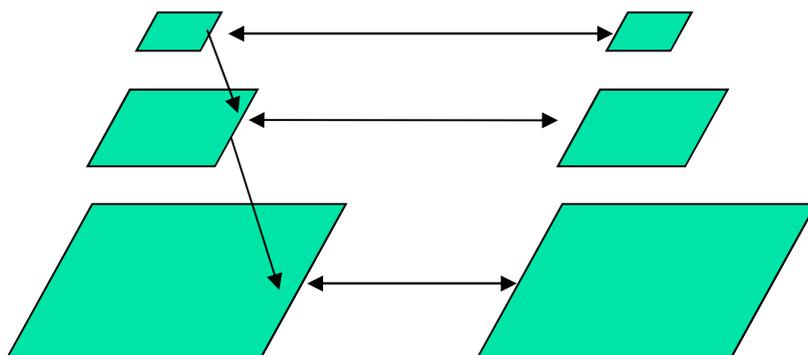
### Méthode of Cafforio and Rocca

$$\vec{d}^{(i+1)} = \vec{d}^{(i)} - \frac{\varepsilon}{\|\nabla I(\vec{p}^{(i)} + \vec{d}^{(i)}, t)\|^2 + \alpha^2} \cdot DFD(\vec{p}^{(i)}, \vec{d}^{(i)}) \cdot \nabla I(\vec{p}^{(i)} + \vec{d}^{(i)}, t)$$



# Approches multi-résolution-multi-échelle

1) Construction des pyramides Gaussiennes pour  $I_t, I_{t+1}$   
 $I_t^0 = I_t$      $I_t^l = (g * I_t^{l-1}) \downarrow$



$$g(k,l) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{k^2+l^2}{2\sigma^2}\right)$$

2) Estimation des paramètres de mouvement commençant par le niveau le plus élevé  $\Theta^L$

3) Propagation

$$t_{x,y}^{l-1} = \rho \cdot t_{x,y}^l \quad \rho \text{ -le facteur de sous-échantillonnage}$$

$$\text{div}, \text{rot}^{l-1} = \text{div}, \text{rot}^l$$



## Estimation robuste (1)

---

- - On suppose le mouvement conforme au modèle paramétrique  $\Theta$
- On considère
  - - les mesures observées  $Y_i$
  - - les mesures conformes au modèle  $M(\Theta, (x, y)_i)$
  - - les résidus  $r_i = Y_i - M(\Theta, (x, y)_i)$
- Principe minimiser un critère d'erreur de façon que les valeurs aberrantes de résidus  $r_i$  ne perturbent l'estimation



## Estimation robuste (2)

- Soit  $P_\sigma(\cdot)$  la loi discrète de distribution d'erreur dépendant du paramètre  $\theta$
- La vraisemblance du paramètre  $\theta$  est définie

$$L(\theta|P_\sigma) = \prod_r P_\sigma(r) \quad \text{Pour toutes les valeurs disponibles de } r$$

- Le max-vraisemblance est trouvé en résolvant

$$\hat{\theta} = \arg \max_{\theta} L(\theta|P_\sigma)$$

- Ceci est équivalent à minimiser

$$\hat{\theta} = \arg \min_{\theta} \sum_r -\log[P_\sigma]$$



## Estimation robuste (3)

$\rho(r, \sigma) = -\log[P_\sigma(r)]$  est appelé “estimateur” (\*)

Supposons que  $r$  suit  $N(0, \sigma^2)$  (Estimateur gaussien). (\*) est l'estimateur aux moindres carrés

$$\rho_G(r, \sigma) = \frac{r^2}{2\sigma^2}$$
$$\sum \frac{r^2}{2\sigma^2} \rightarrow \min$$

Estimateur de Lorentz  $\rho_L(r, \sigma) = \log\left(1 + \frac{r^2}{2\sigma^2}\right)$

Estimateur de Geman-McClur  $\rho_{GM}(r, \sigma) = \frac{r^2 / \sigma^2}{1 + r^2 / \sigma^2}$

D. Hasler, L. Sbaiaz, S. Susstrunk, M. Vetterli, « Outlier Modeling in Image Matching »,  
IEEE TRans on PAMI, v. 25, n3, march 2003



## Estimation robuste (5)

$$\sum_i \rho(r_i) = \sum_i \frac{1}{2} w_i r_i^2, \quad r_i = Z_i - Z(\Theta, (x, y))_i \longrightarrow \min$$

$$\frac{\partial}{\partial \theta_j} (\sum_i \rho(r_i)) = \frac{\partial}{\partial \theta_j} (\sum_i \frac{1}{2} w_i r_i^2), \quad r_i = Y_i - M(\Theta, (x, y))_i$$

$$\sum_i \psi(r_i) \frac{\partial r_i}{\partial \theta_j} = \sum_i w_i r_i \frac{\partial r_i}{\partial \theta_j} = 0 \quad \text{ici} \quad \psi(r_i) = \dot{\rho}(r_i)$$

$$w_i = \frac{\psi(r_i)}{r_i}$$



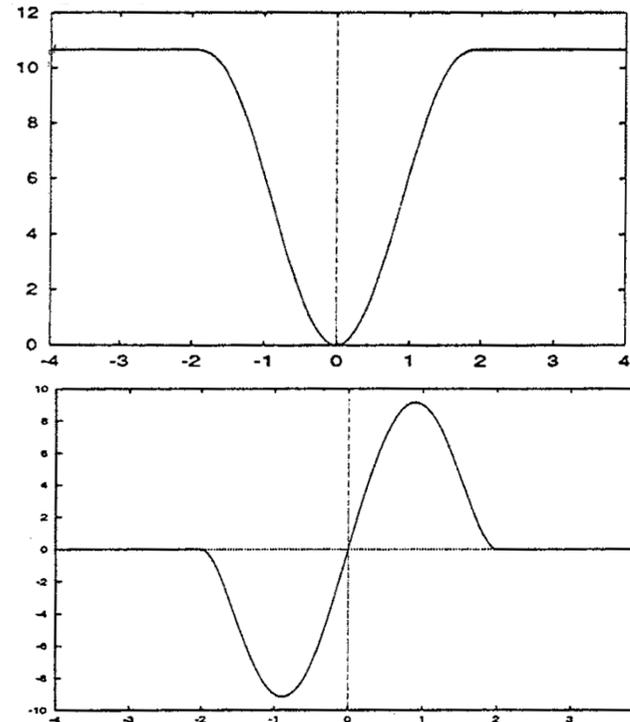
## Estimation robuste (3)

### Estimateur de Tuckey

$$\rho(r, C) = \begin{cases} \frac{r^6}{6} - \frac{2C^2 r^4}{4} + \frac{C^4 r^2}{2}, & \text{if } |r| < C \\ \frac{C^6}{6}, & \text{otherwise} \end{cases}$$

### La dérivée

$$\psi(r, C) = \begin{cases} r(r^2 - C^2)^2, & \text{if } |r| < C \\ 0, & \text{otherwise} \end{cases}$$



$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_i \rho(r_i, C), \text{ } r_i \text{--le résidu}$$

$$\hat{\theta} = (H^T W H)^{-1} H^T W Z$$



Laboratoire Bordelais  
de Recherche  
en Informatique



© LaBRI

© Warner Bros.

06/04/2012



# Segmentation en plages homogènes

- Problème : qualifier le mouvement de la caméra et segmenter en micro-plans
- 1. Reformulation du modèle du mouvement affine complet du 1<sup>er</sup> ordre

$$\begin{cases} dx(x,y)=a_1+a_2x+a_3y \\ dy(x,y)=a_4+a_5x+a_6y \end{cases}$$

$$\Theta = (a_1, a_4, div, rot, hyp_1, hyp_2)^T$$

$$div = \frac{1}{2}(a_2 + a_6) \quad rot = \frac{1}{2}(a_5 - a_3) \quad hyp_1 = \frac{1}{2}(a_2 - a_6) \quad hyp_2 = \frac{1}{2}(a_3 + a_5)$$

- 2. Estimation du modèle simplifié (4 paramètres)

$$PTZ \quad \Theta = (a_1, a_4, div, rot)^T \quad div = a_2 = a_6, rot = -a_3 = a_5$$



# Tests des hypothèses statistiques sur la signification des paramètres du mouvement

---

$H_0$  : le paramètre considéré  $a_j$  est significatif

$H_1$  :  $a_j = 0$ .

En supposant les distributions gaussiennes  
des erreurs d'estimation

$$\frac{Nd}{2} \left( \ln \left( \sum_{i \in Sd} r_{i0}^2 \right) - \ln \left( \sum_{i \in Sd} r_{i1}^2 \right) \right) \underset{H_1}{\overset{H_0}{\leq}} \lambda_A$$

$r_{ij}$  - erreurs résiduels par  
rapport au modèle



Laboratoire Bordelais  
de Recherche  
en Informatique

Deformation	Translation	Movement de la caméra
(0,0,0,0)	(0,0)	Caméra statique
(0,0,0,0)	$\neq(0,0)$	Panoramique ou translation latérale le fond est parallèle au plan focal
(div,0,0,0)		Zoom ou travelling in/out
(0,rot,0,0)		Rotation relative to OA
(div,rot,0,0)		Combinaison des deux précédents
(div,rot,hyp1,hyp2)		Translation avec le fond non-parallèle au plan focal

D'après P. Bouthemy, M. Gelgon

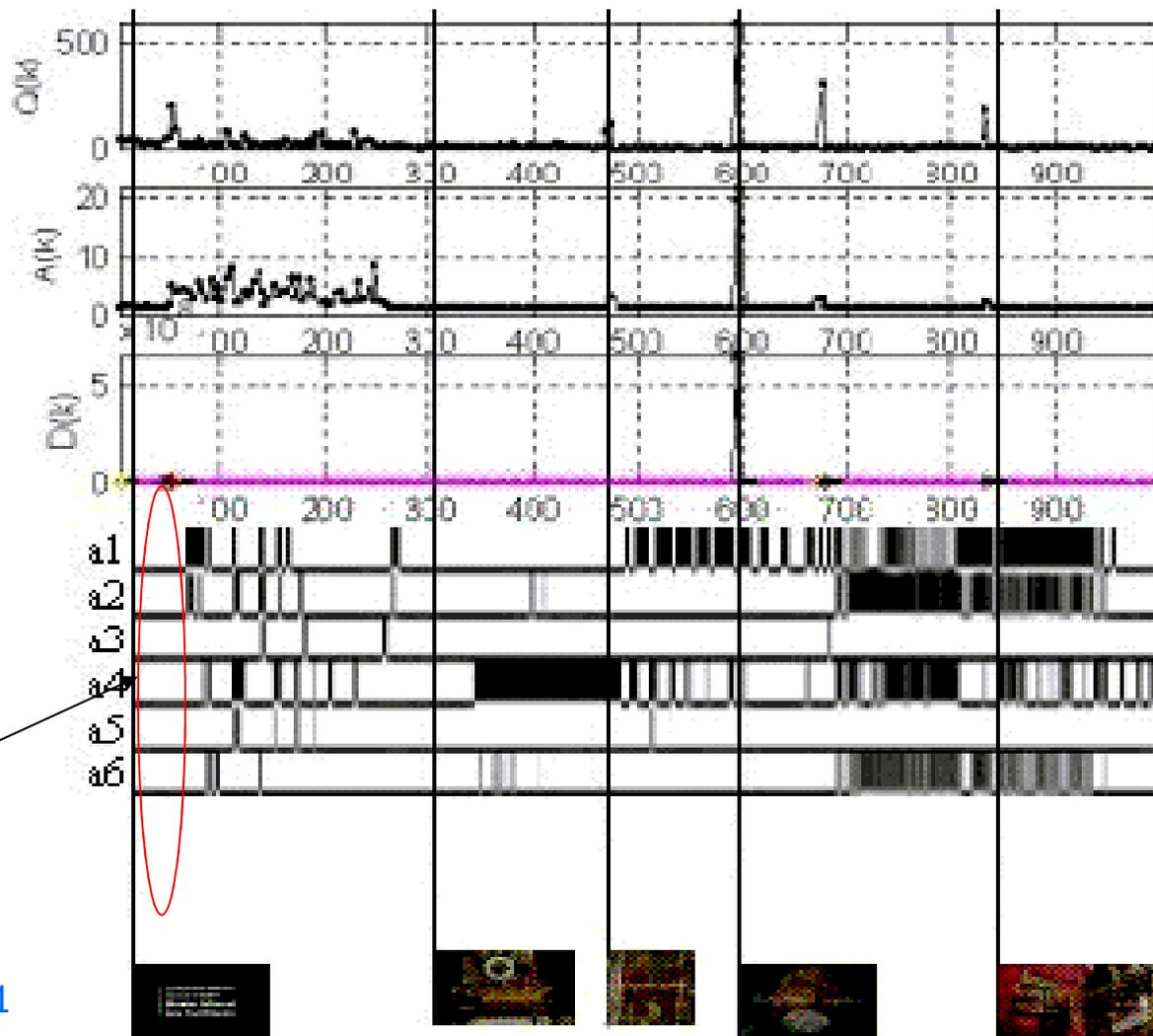
06/04/2012



Laboratoire Bordelais  
de Recherche  
en Informatique

# La mise en œuvre sur FO MPEG2

Caméra  
statique



06/04/201



Laboratoire Bordelais  
de Recherche  
en Informatique

---

# Questions?

06/04/2012