



---

# Systemes de representation multi-echelles pour l'indexation et la restauration d'archives medievals couleur

**Julien DOMBRE**

Laboratoire IRCOM-SIC, UMR-CNRS 6615.

En partenariat avec le Centre d'Études Supérieures des Civilisations Médiévales (CESCM).



# Contexte de l'étude

- ▶ Projet transverse entre deux laboratoires de l'université de Poitiers :
  - ▶ SIC - STIC
  - ▶ CESCUM - SHS
- ▶ Fond iconographique consacré à l'art roman en France et à l'étranger



Restauration



Étude stylistique

1. Introduction
2. Décrire les images pour indexer
3. Segmentation multi-échelles
4. Description des régions
5. Utilisation de la structure pour l'indexation
6. Conclusion / Perspectives

# Présentation du problème

- ▶ **Indexation de bases d'images** : retrouver automatiquement des images d'un contenu bien précis dans une base de données.

≠ classification

- ▶ **Méthode de recherche au CESC** = moteurs de recherche Web

- ▶ Recherche de motifs textuels dans la description :

- ▶ Rechercher les pages contenant "loisirs et Poitiers"
- ▶ Trouver les images contenant un "Christ"

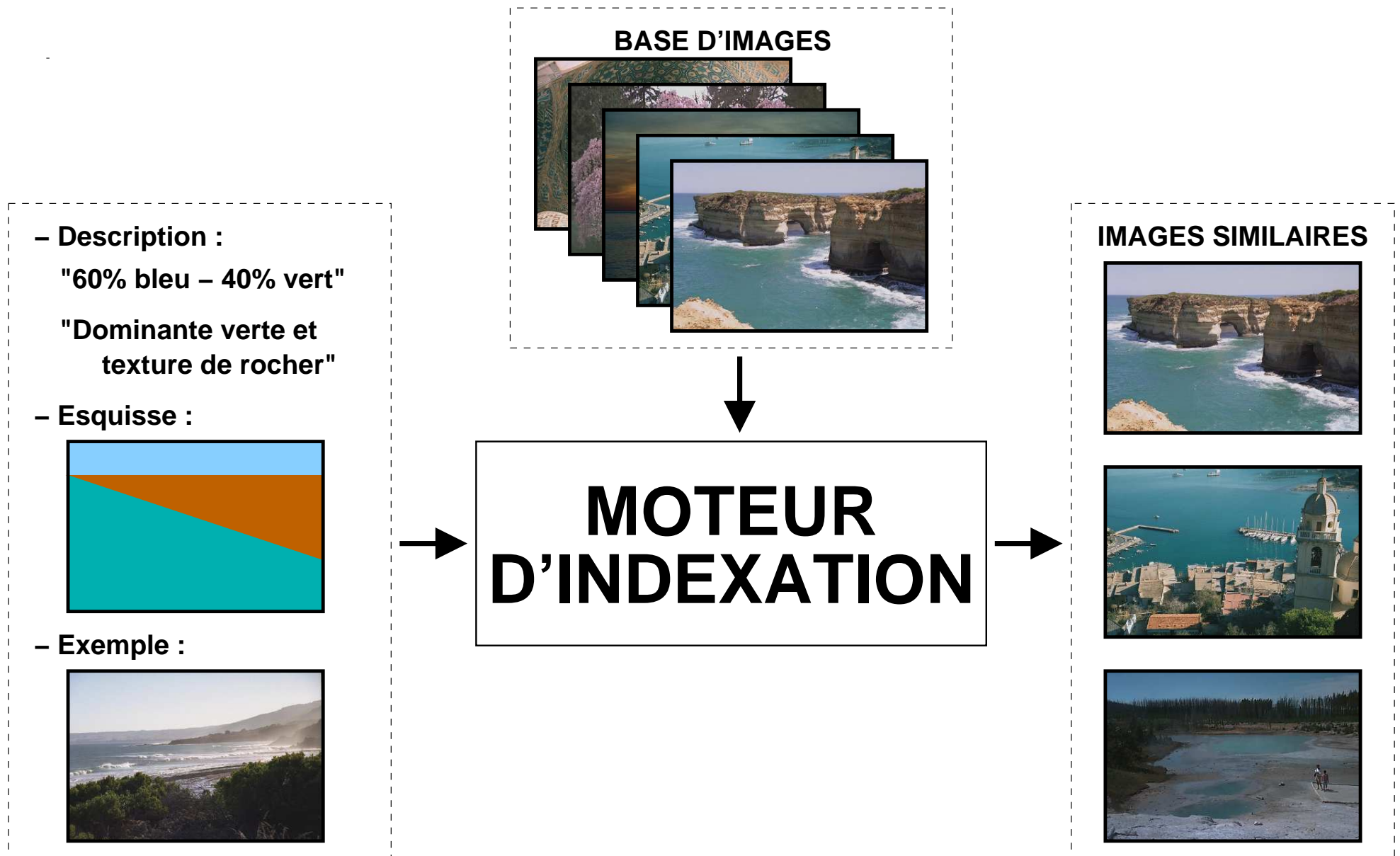
- ▶ Limitations de l'indexation textuelle :

- ▶ Langue utilisée
- ▶ Vocabulaire utilisé (expert/néophyte)
- ▶ Niveau de description

- ▶ **Utilisation du contenu** des images pour les caractériser

Indexation d'images par le contenu (*Content Based Image Retrieval*)

# Introduction / Contexte



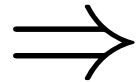
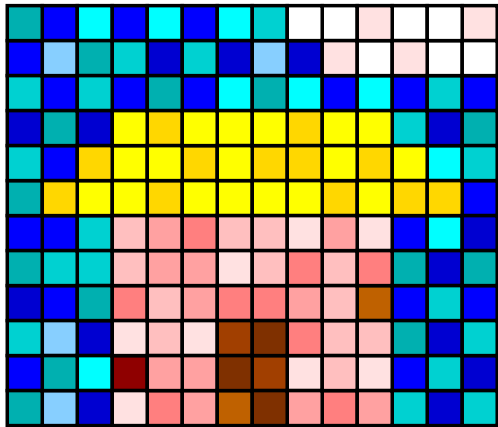
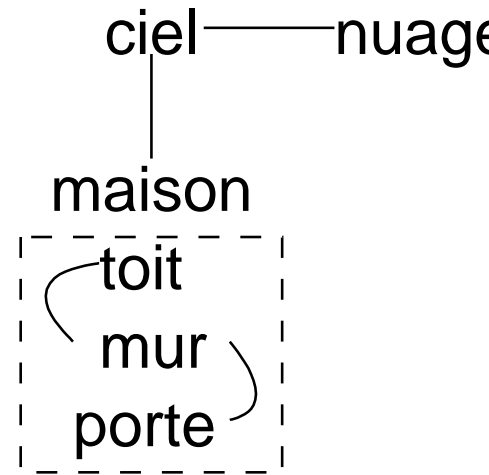
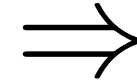
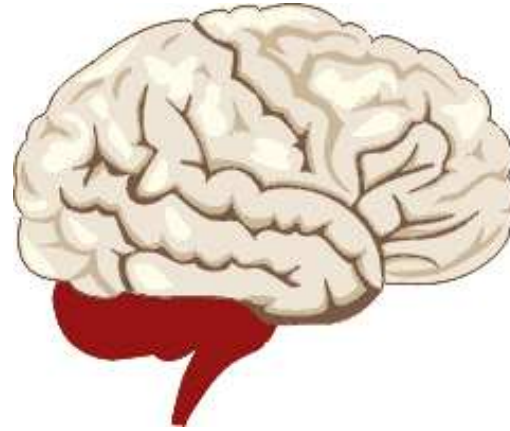
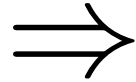
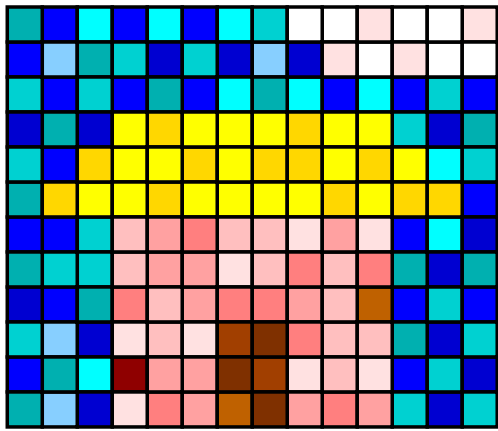


# Plan

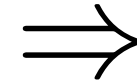
---

1. Introduction
2. Décrire les images pour indexer
3. Segmentation multi-échelles
4. Description des régions
5. Utilisation de la structure pour l'indexation
6. Conclusion / Perspectives

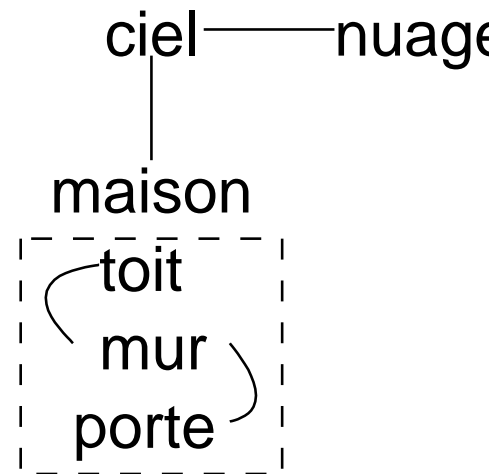
# La vision comme support de l'analyse



??



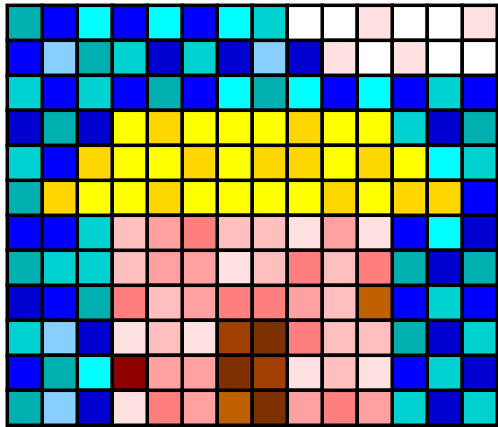
??



► Où s'arrêter pour caractériser / décrire le plus finement l'image



# L'image brute - la matrice



## ► Acquisition :

- Capture du flux lumineux filtré pour l'ensemble des photosites du dispositif d'acquisition
- En couleur : 3 composantes ( $RGB, L^* a^* b^* \dots$ )

## ► Représentation de l'image :

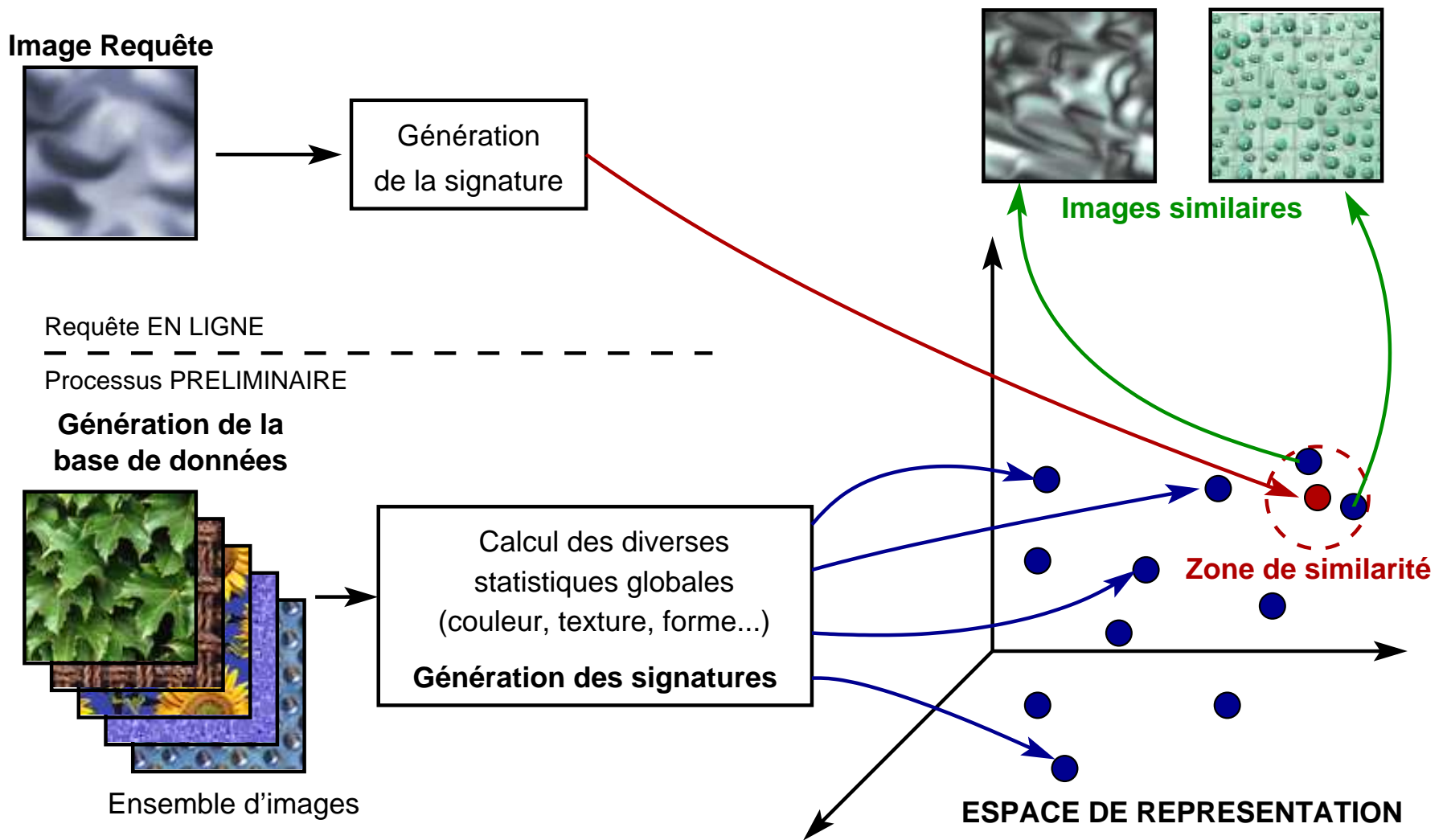
- Toute l'information
- Non structurée

## ► Calcul de statistiques globales aux images → Signature :

- Couleur : histogrammes couleur, *Color Sets*...
- Texture : matrices de cooccurrences, filtres de GABOR...
- ...

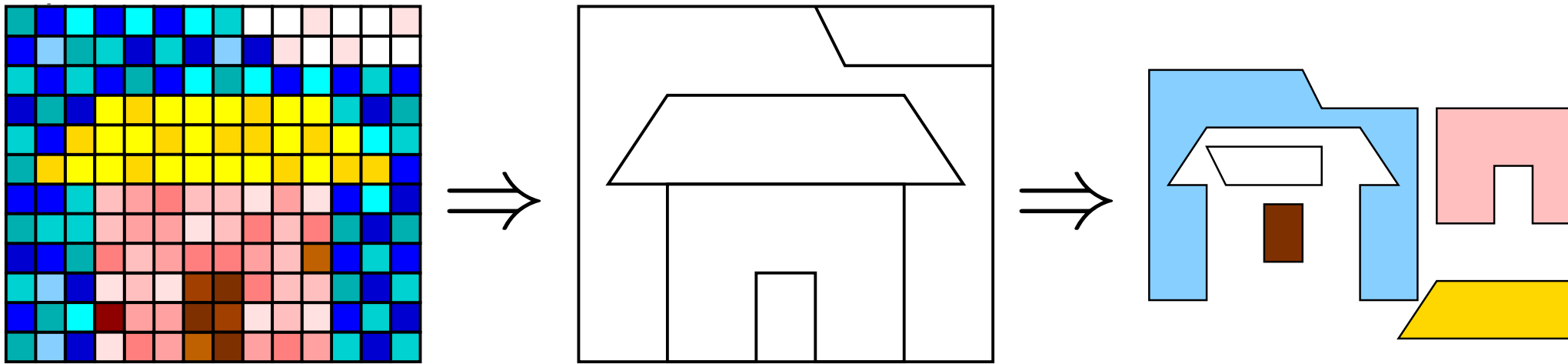


# L'image brute - la matrice



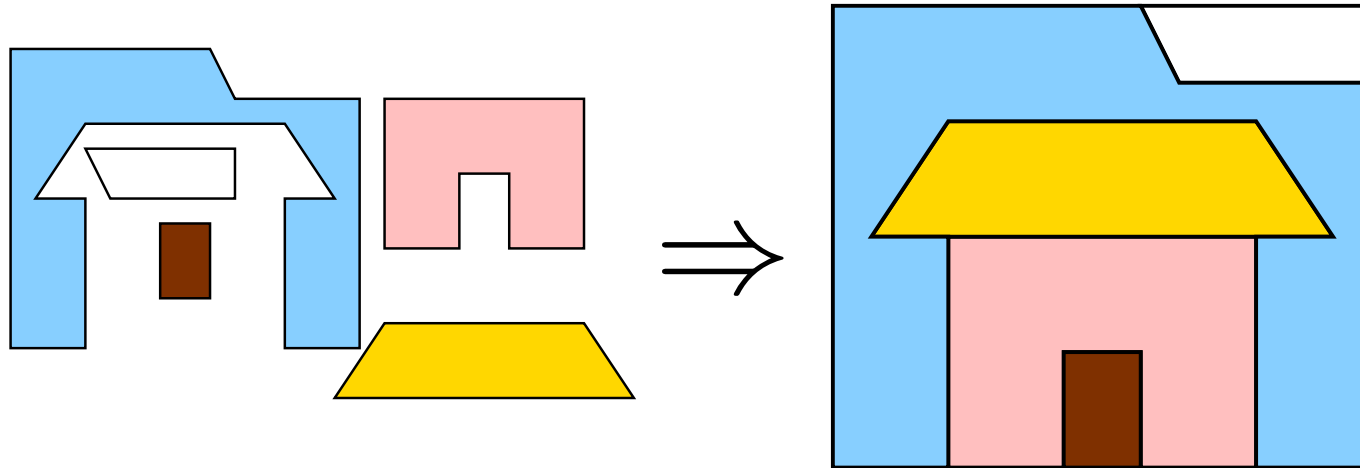
- ▶ Recherche d'objets très difficile : caractérisation noyée au sein de celle de l'image

# Extraction d'éléments simples



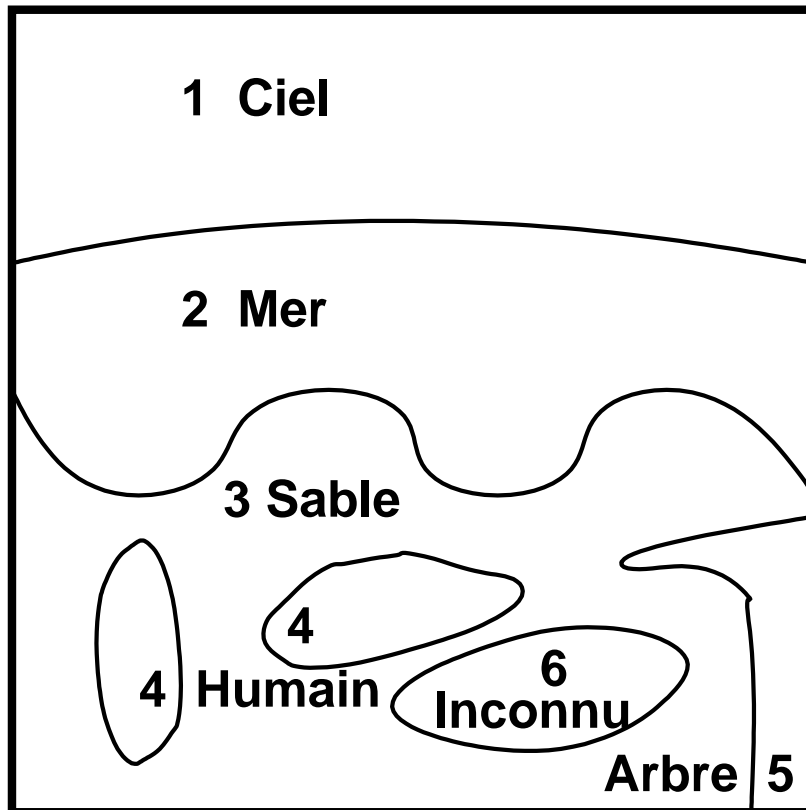
- ▶ Extraction de primitives de l'image (lignes, contours, régions...)
- ▶ Caractérisation des primitives
- ▶ Recherche de primitives et non d'images [MA99,CARSON02]
- ▶ Description insuffisante :
  - ▶ pour trouver un arrangement d'objets donné
  - ▶ pour gommer les erreurs d'extraction

# Arrangement des primitives



- ▶ Image = Ensemble organisé de régions caractérisées
- ▶ Organisation des éléments → indexation spatiale :
  - ▶ Différentes variétés de *2D-Strings* [PETRAKIS93, CHANG91, LEE90]
  - ▶ Graphes d'adjacences de régions labéllisés ou attribués [PETRAKIS97]

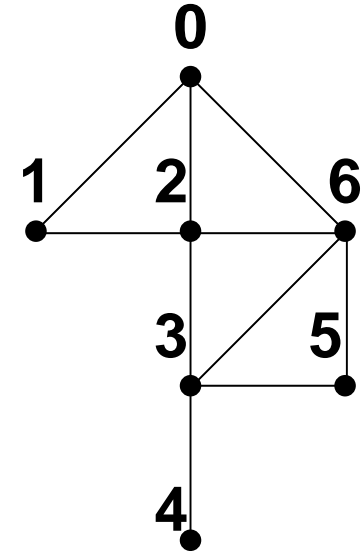
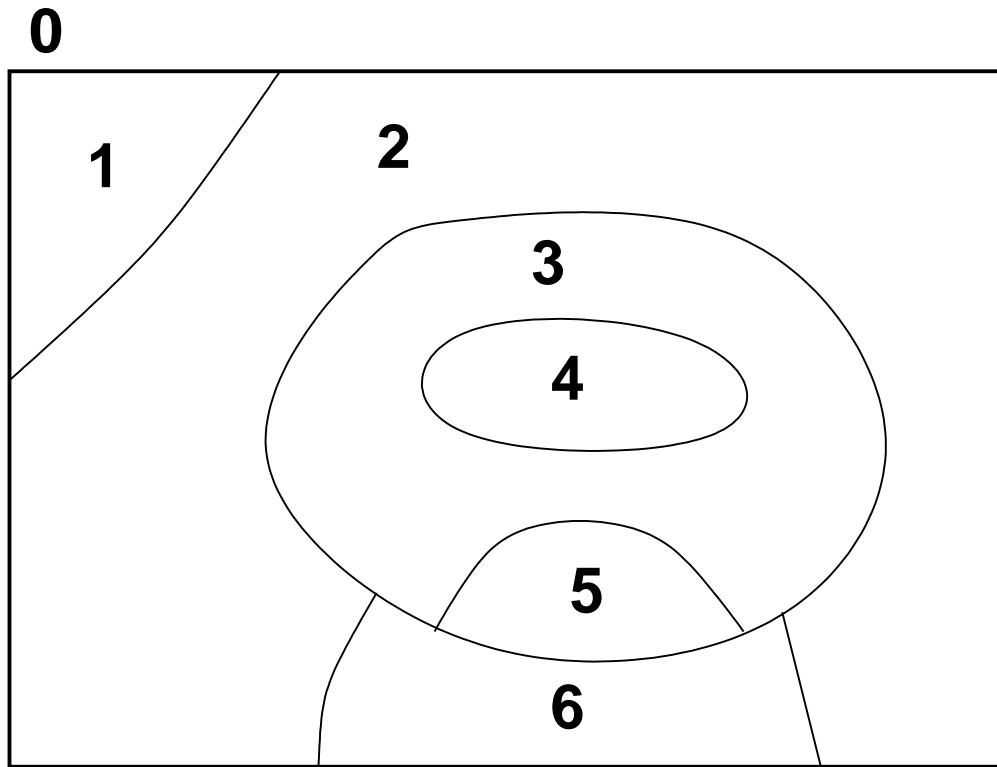
# 2D-Strings



Ordre	2-D string
Horizontal	(4 4 2 3 1 6 5)
Vertical	(1 2 4 3 4 6 5)
Aire	(4 4 6 5 1 2 3)
Périmètre	(4 4 5 6 1 3 2)

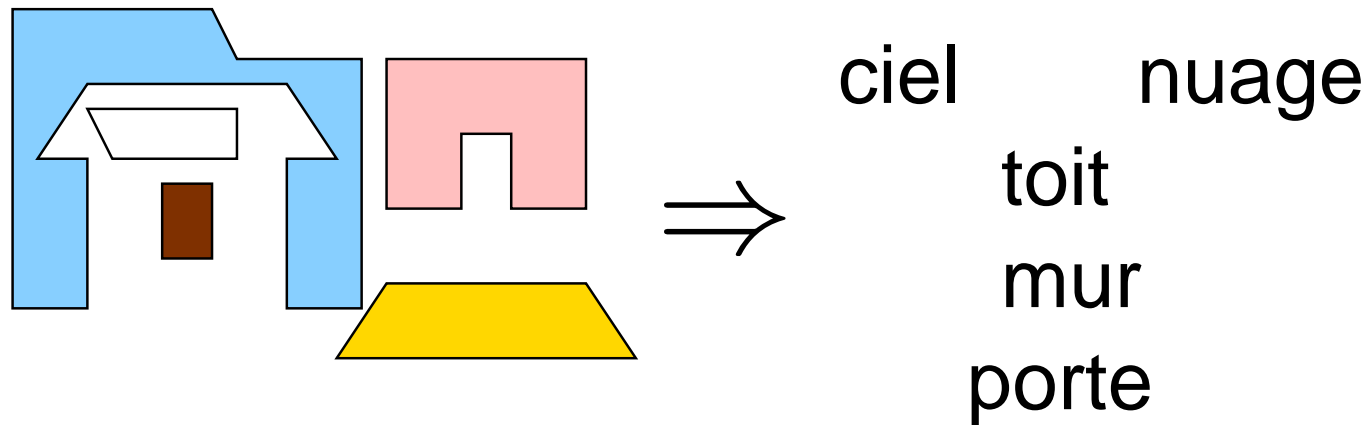
- ▶ Nécessite l'identification des régions
- ▶ Positionnement relatif des régions (centres de gravité)

# Graphe d'adjacences de régions

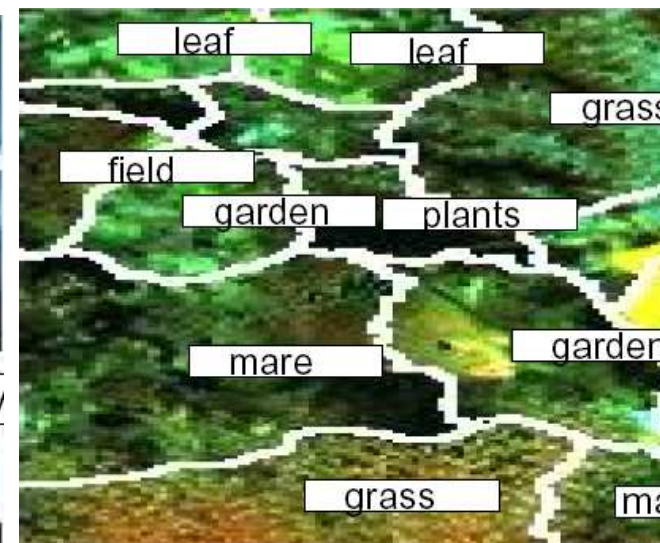


- ▶ Problème de la vision d'un objet à plusieurs échelles
- ▶ Pas de notion de composition

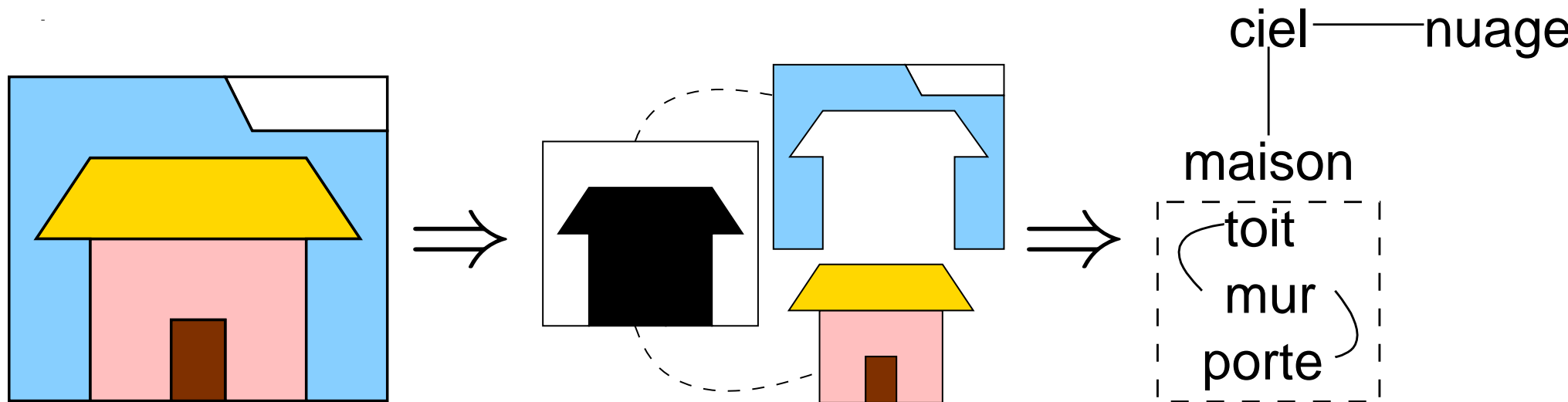
# Remontée jusqu'aux concepts



- ▶ Difficulté de remonter jusqu'à une description sémantique complète
  - ▶ Tentative par KOBUS BARNARD [BARNARD01]
  - ▶ Vocabulaire limité et concepts très généraux (chat  $\neq$  tigre)



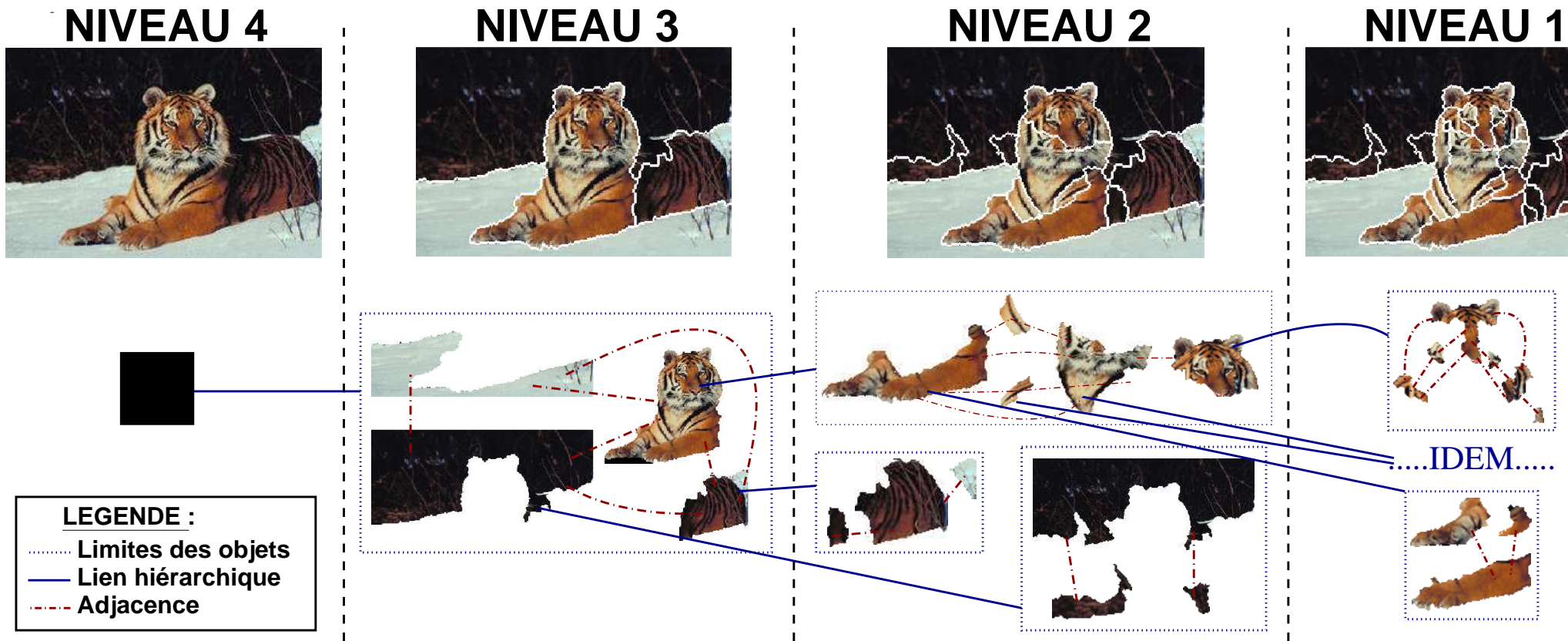
# Arrangement d'objets composés



- ▶ Image = Ensemble d'objets composés en relation
- ▶ Vision d'un objet à différentes échelles
- ▶ Nécessité d'un algorithme de segmentation multi-échelles



# Le graphe pyramidal



- ▶ Caractérisation des régions et des relations au sein de la structure
- ▶ Segmentation multi-echelles orientée objet indispensable



# Plan

---

1. Introduction
2. Décrire les images pour indexer
3. **Segmentation multi-échelles**
4. Description des régions
5. Utilisation de la structure pour l'indexation
6. Conclusion / Perspectives



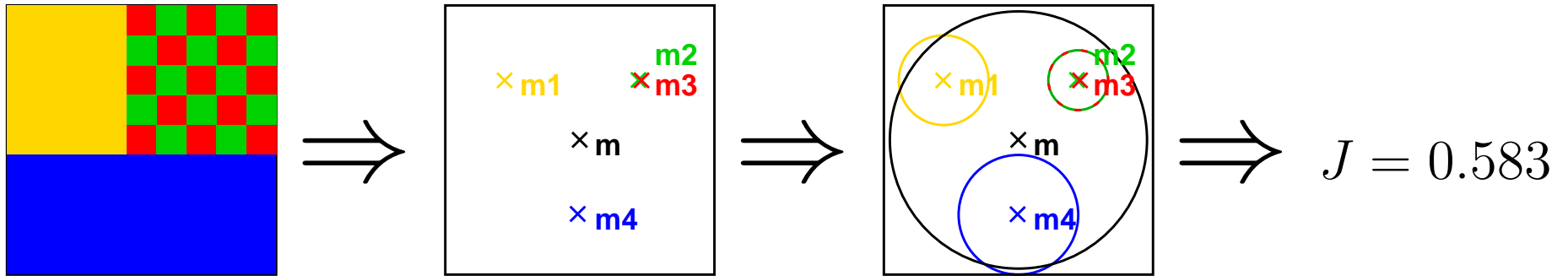
# Méthodes existantes

- ▶ Contraintes :
  - ▶ Multi-échelles
  - ▶ Orientée objet
- ▶ Domaine du critère : couleur, texture, couleur/texture
  - ▶ Seuillage d'histogrammes
  - ▶ Méthodes de classification (*k-means*, *ISODATA*)
  - ▶ ...

→ Difficulté d'utilisation de la nature spatiale de l'image
- ▶ Domaine de l'image
  - ▶ Détection contours
  - ▶ Accroissement de régions
  - ▶ *Split-and-Merge*
  - ▶ ...

→ Méthodes multi-échelles et orientées objets ?

# Critère d'homogénéité



► Critère de rupture de modèle textural défini par DENG [DENG99] :

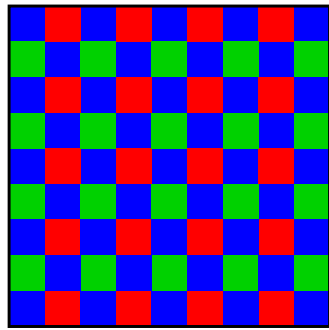
- $Z$  : ensemble de  $N$  points de centre de gravité  $m$
- $Z$  divisé en  $C$  classes :  $Z_i, \forall i \in [1 \dots C]$  de centre de gravité  $m_i$
- Dispersion totale :  $D_t = \sum_{z \in Z} \| z - m \|^2$

- Dispersion relative :  $D_r = \sum_{i=1}^C D_i = \sum_{i=1}^C \sum_{z \in Z_i} \| z - m_i \|^2$

$$J_{Deng} = \frac{D_t - D_r}{D_r} \rightarrow \boxed{J_{proposé} = \frac{D_t - D_r}{D_t}}$$

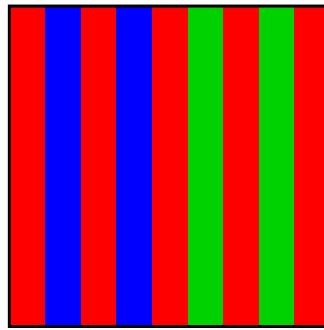
- Si les centres sont confondus  $\Rightarrow D_r = D_t \Rightarrow J = 0$

# Critère d'homogénéité



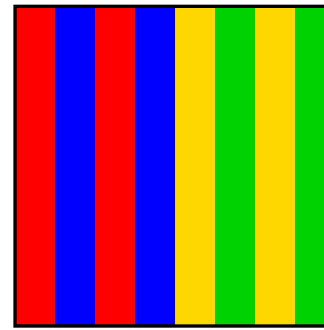
$$D_t = D_r$$

$$J = 0$$



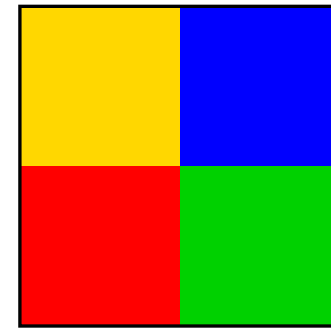
$$D_t > D_r$$

$$J = 0.1333$$



$$D_t \gg D_r$$

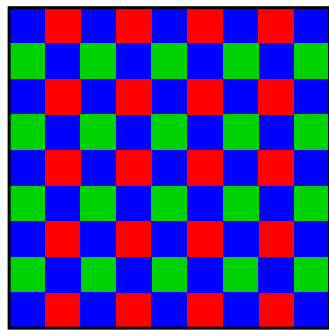
$$J = 0.4048$$



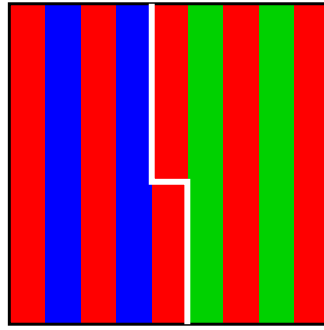
$$D_t \gg D_r$$

$$J = 0.7619$$

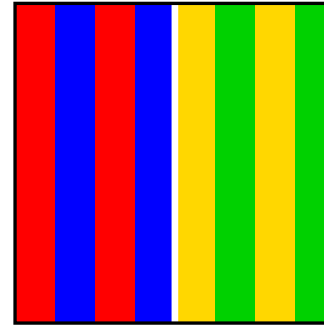
- Restriction du calcul de  $J$  aux régions définies :



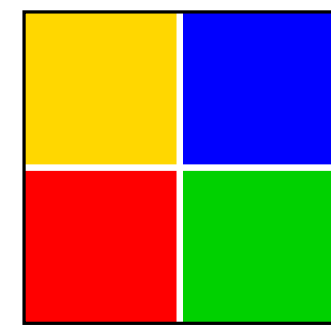
$$\bar{J} = 0$$



$$\bar{J} = 0.0107$$



$$\bar{J} = 0.0385$$



$$\bar{J} = 0$$

- Critère d'évaluation :

$$\bar{J} = \frac{1}{N} \sum_k M_k J_k$$

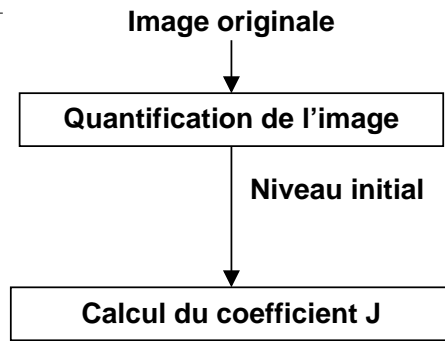
$M_k$  : nombre de pixels de la  $k^e$  région

$J_k$  : valeur de  $J$  de la  $k^e$  région

$N$  : nombre de pixels du voisinage

- Quantification pour obtenir des couleurs visuellement différentes

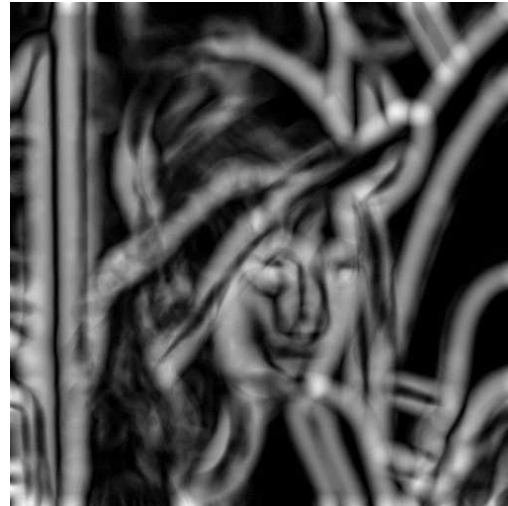
# Segmentation multi-échelles - Cartes de $J$



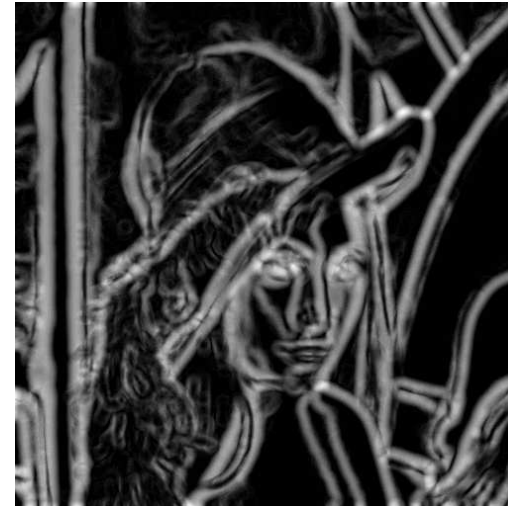
- ▶ Calcul de  $J$  pour tous les pixels de l'image avec un voisinage donné
- ▶ Influence de la taille du voisinage sur les ruptures détectées



originale



diamètre 33

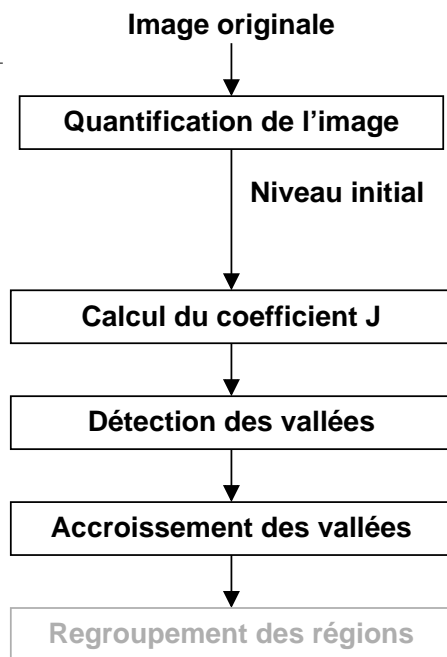


diamètre 17

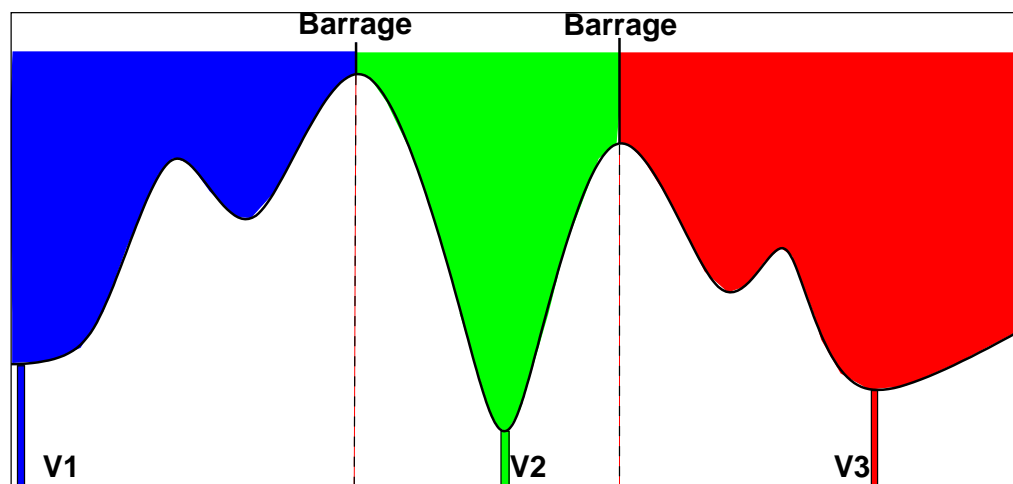


diamètre 9

# Détection / Accroissement des vallées

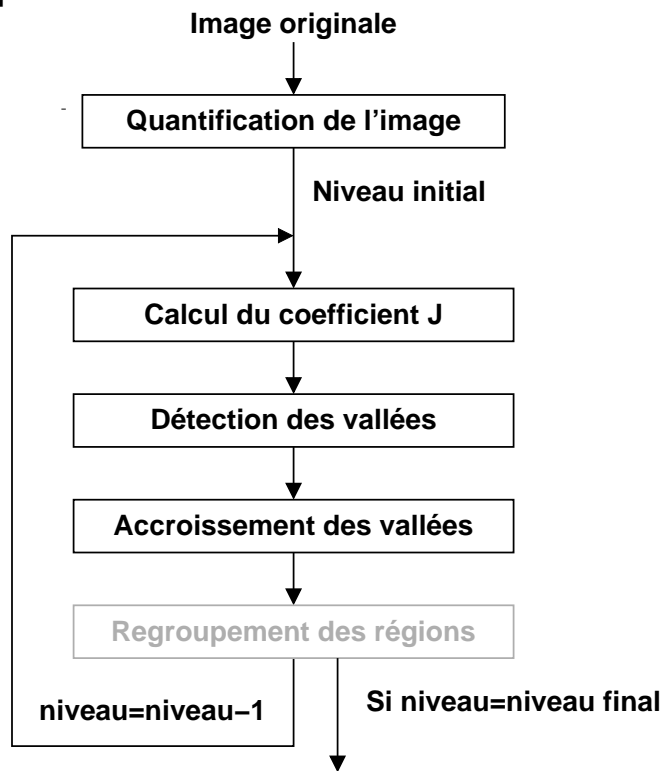


- ▶ Détection des marqueurs par seuillage adaptatif
- ▶ Accroissement par ligne de partage des eaux





# Extension multi-échelles



- ▶ Détection des vallées par seuillage adaptatif au sein des régions déjà générées :
  - ▶ Raffinement progressif de la segmentation
  - ▶ Localisation de la détection des vallées
- ▶ Non restriction lors de l'accroissement des vallées - processus identique



# Repositionnement des échelles

- Cohérence → Repositionnement : confiance sur l'échelle la plus fine

Niveaux Initiaux

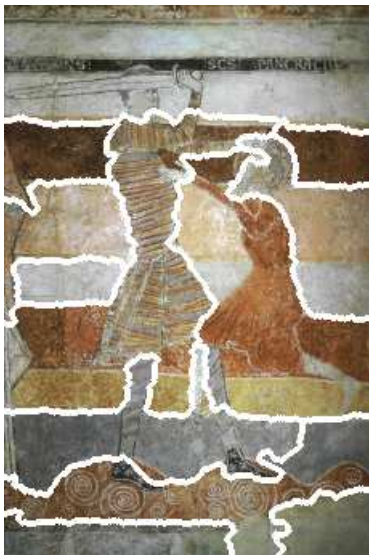


Niveaux Repositionnés

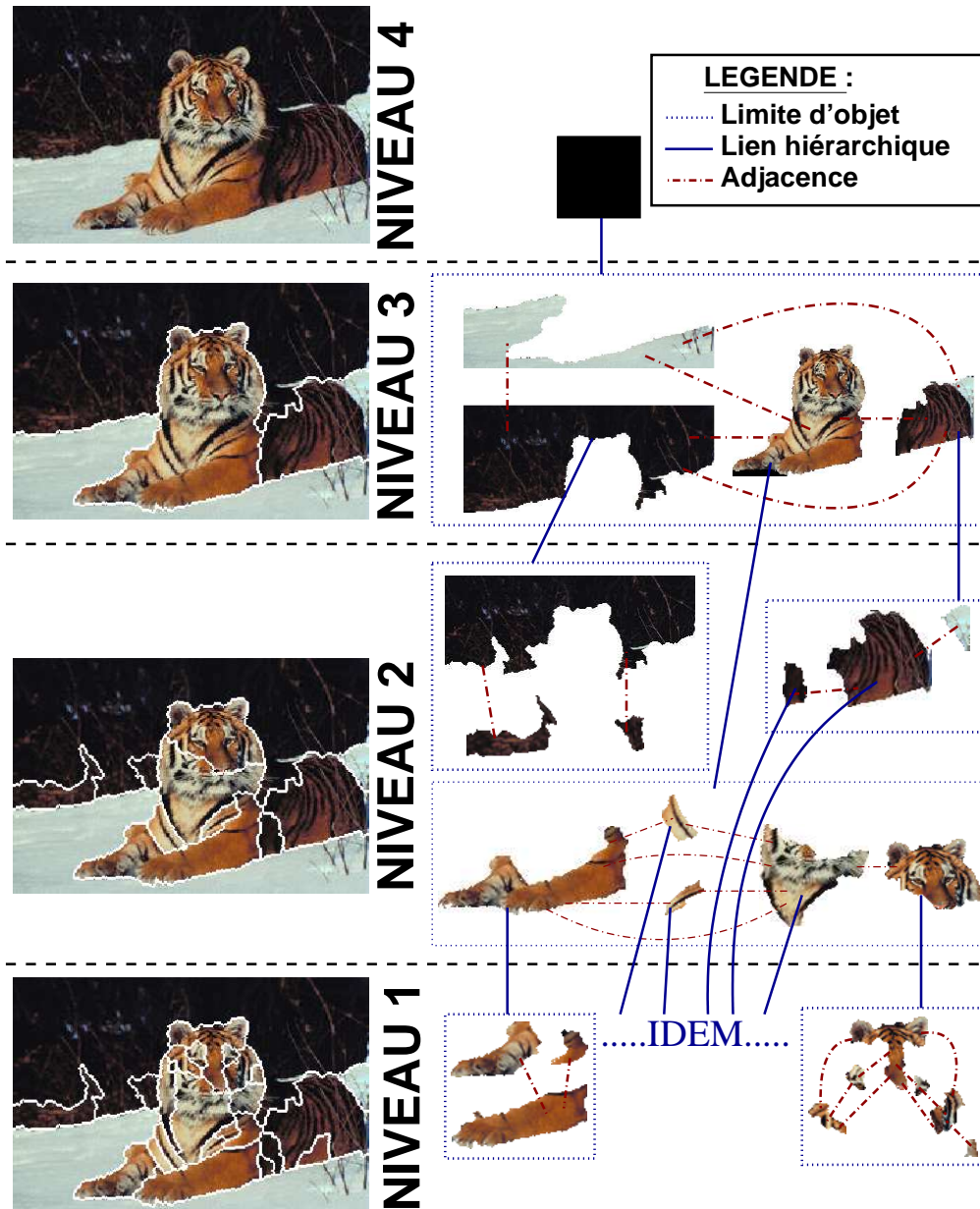




# Quelques résultats



# Construction du graphe pyramidal





# Plan

---

1. Introduction
2. Décrire les images pour indexer
3. Segmentation multi-échelles
4. **Description des régions**
5. Utilisation de la structure pour l'indexation
6. Conclusion / Perspectives

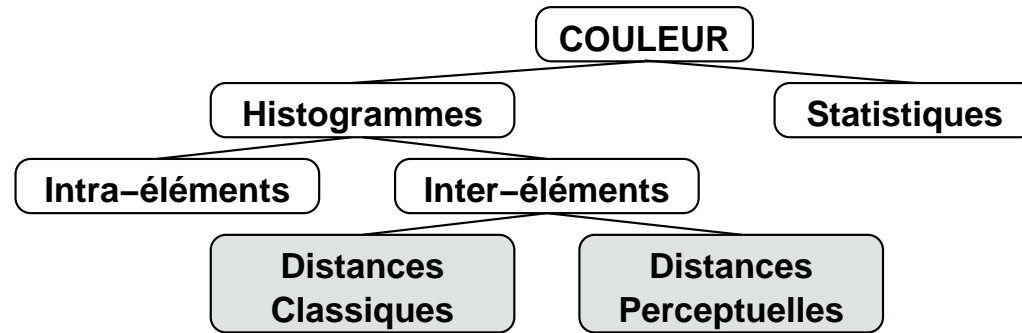
# Description des régions

- ▶ Trois critères différents :
  - ▶ couleur
  - ▶ texture
  - ▶ forme
- ▶ Critères plus ou moins discriminants suivant les applications :
  - ▶ CESCМ : texture  $\simeq$  bruit causé par le vieillissement
  - ▶ Analyse de tissu : couleur et texture
  - ▶ Tri de pièces : suivant la nature des pièces

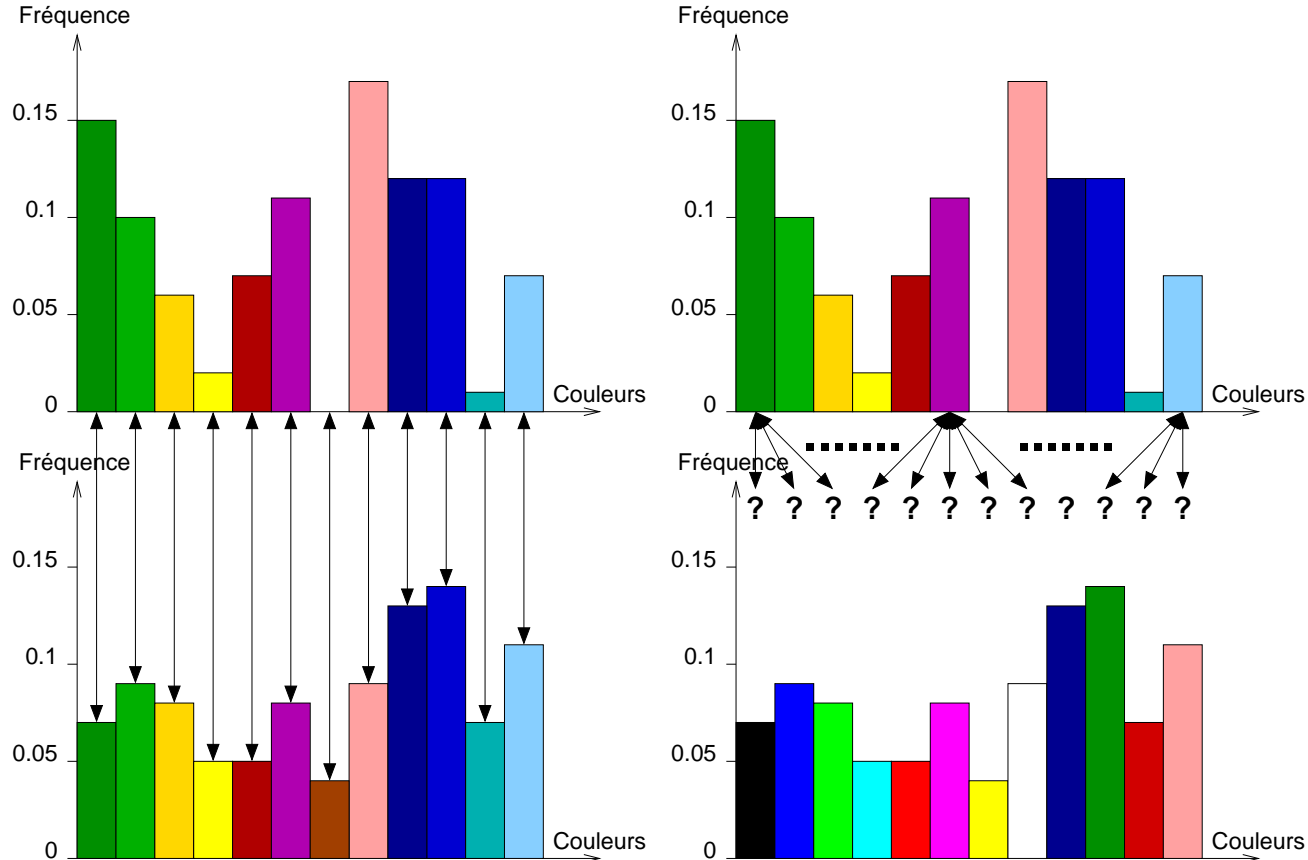


- ▶ Additivité des statistiques → Réduction des calculs
- ▶ Statistiques indépendantes de l'échelle

# Description Couleur

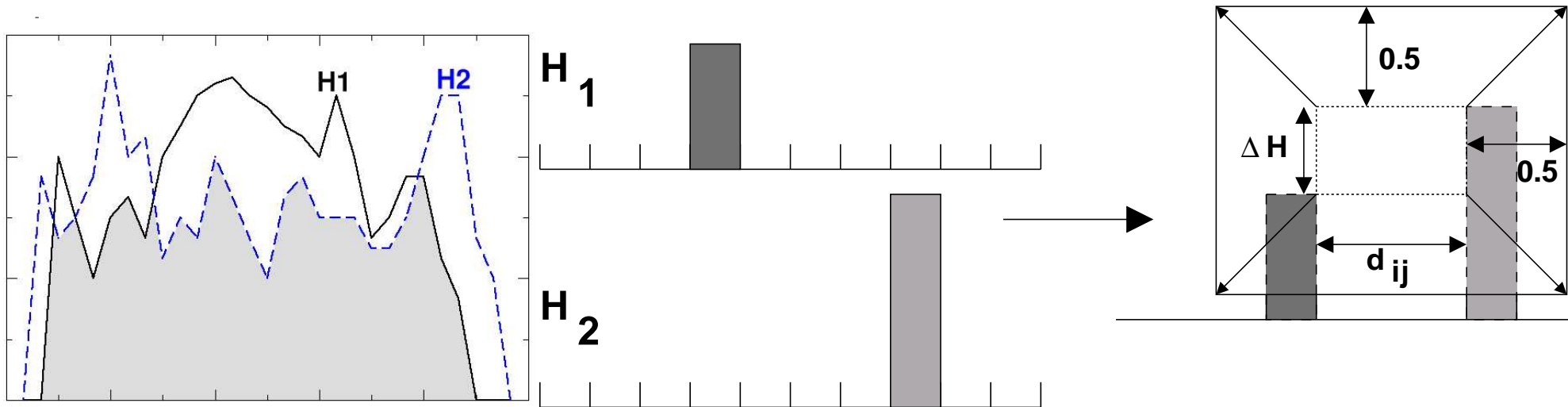


► Histogramme : Fréquences d'apparition des différentes couleurs





# Intersection d'histogrammes

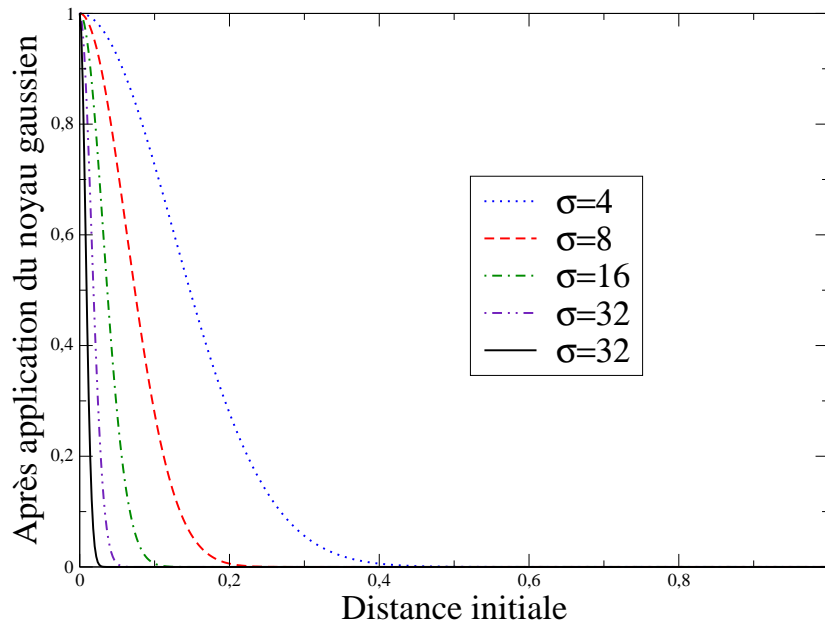


- ▶ SWAIN et BALLARD [SWAIN91] : inutilisable si palettes différentes
- ▶ Évolution pour des supports colorimétriques quelconques :
  - ▶ 2 histogrammes normalisés  $H_1$  et  $H_2$  → Comparaison des fréquences
  - ▶  $d_{ij}$  : distance couleur normalisée → Comparaison colorimétrique

$$s_{12} = \frac{\sum_{i \in \zeta_1} \left[ \min_{j \in \zeta_2} \{ (k_1 + |H_1(i) - H_2(j)|) * (k_2 + d_{ij}) - k_1 k_2 \} \right]}{(1 + k_1 + k_2)} \rightarrow S = \frac{s_{12} + s_{21}}{2}$$

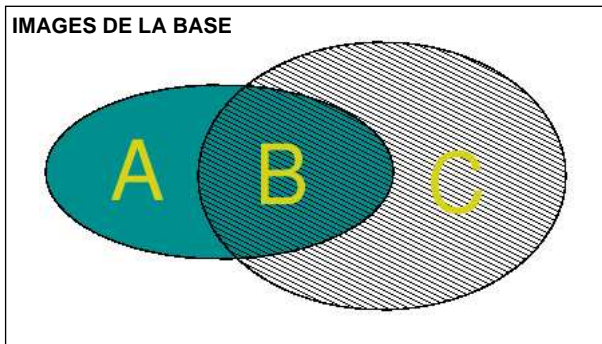
# Autres propositions

- ▶ *Earth Mover's Distance* : métrique de référence déterminant le coût minimum du transport des éléments d'une distribution à l'autre
- ▶ *EMD* très performante mais très lente ( $\mathcal{O}(n^2)$ )
- ▶ Proposition d'un algorithme en  $\mathcal{O}(n)$  :
  - ▶ optimisation par transport prioritaire des couleurs proches
    - ▶ Théoriquement : perte de la notion de métrique
    - ▶ Pratiquement : résultats équivalents entre les deux méthodes
- ▶ Quantification de la distance entre 2 couleurs différentes acceptable

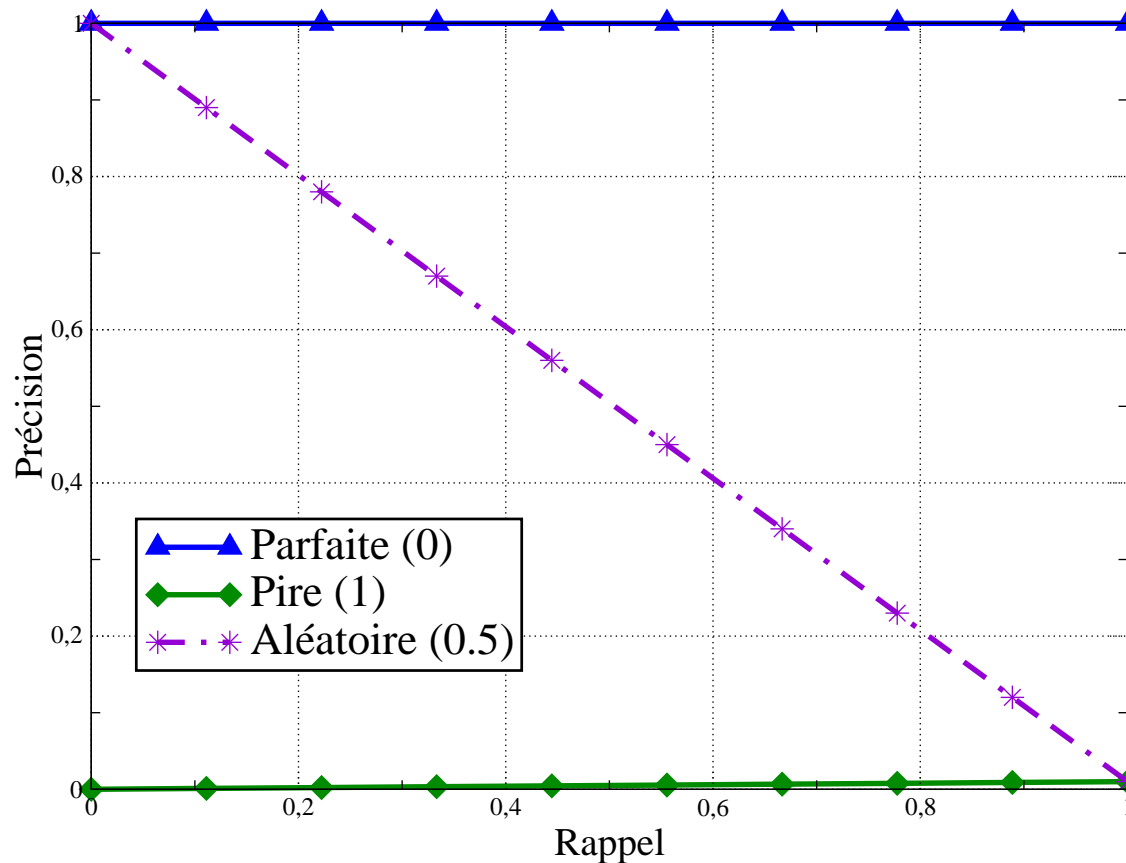


- ▶ rouge-bleu > vert-bleu ?
- ▶ 2 couleurs sont comparables jusqu'à un certain seuil
- ▶  $p(i, j) = A * e^{\frac{-d(i, j)}{2\sigma^2}}$
- ▶ Réglage de  $\sigma$

# Résultats - Présentation

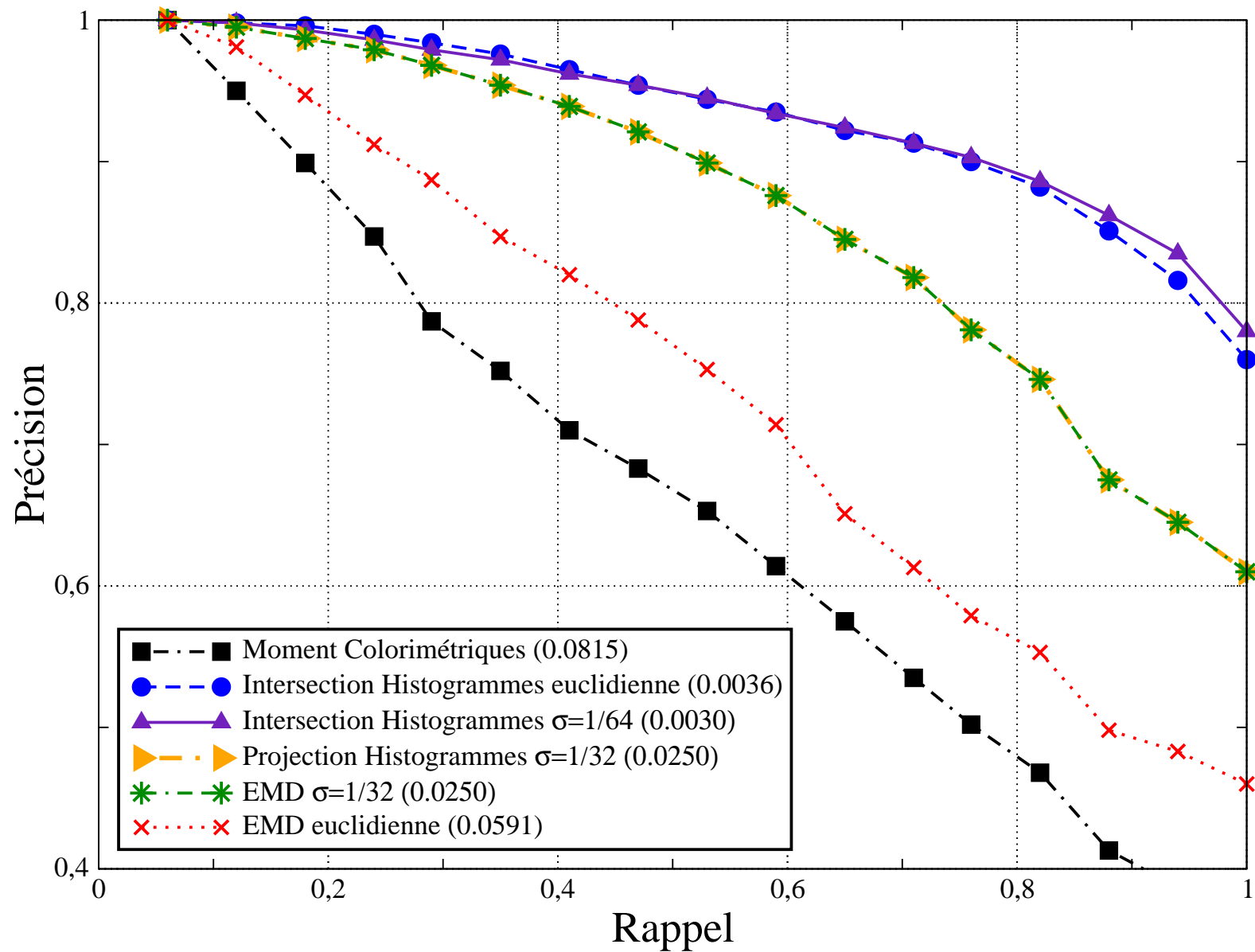
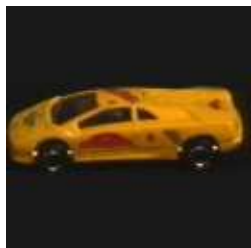


- ▶ Précision =  $\frac{B}{B + C}$ , Rappel =  $\frac{B}{A + B}$
- ▶  $A + B$  : images pertinentes
- ▶  $B + C$  : images retrouvées
- ▶  $\widetilde{Rang}$  = Normalisation du rang moyen des images pertinentes [MÜLLER99]



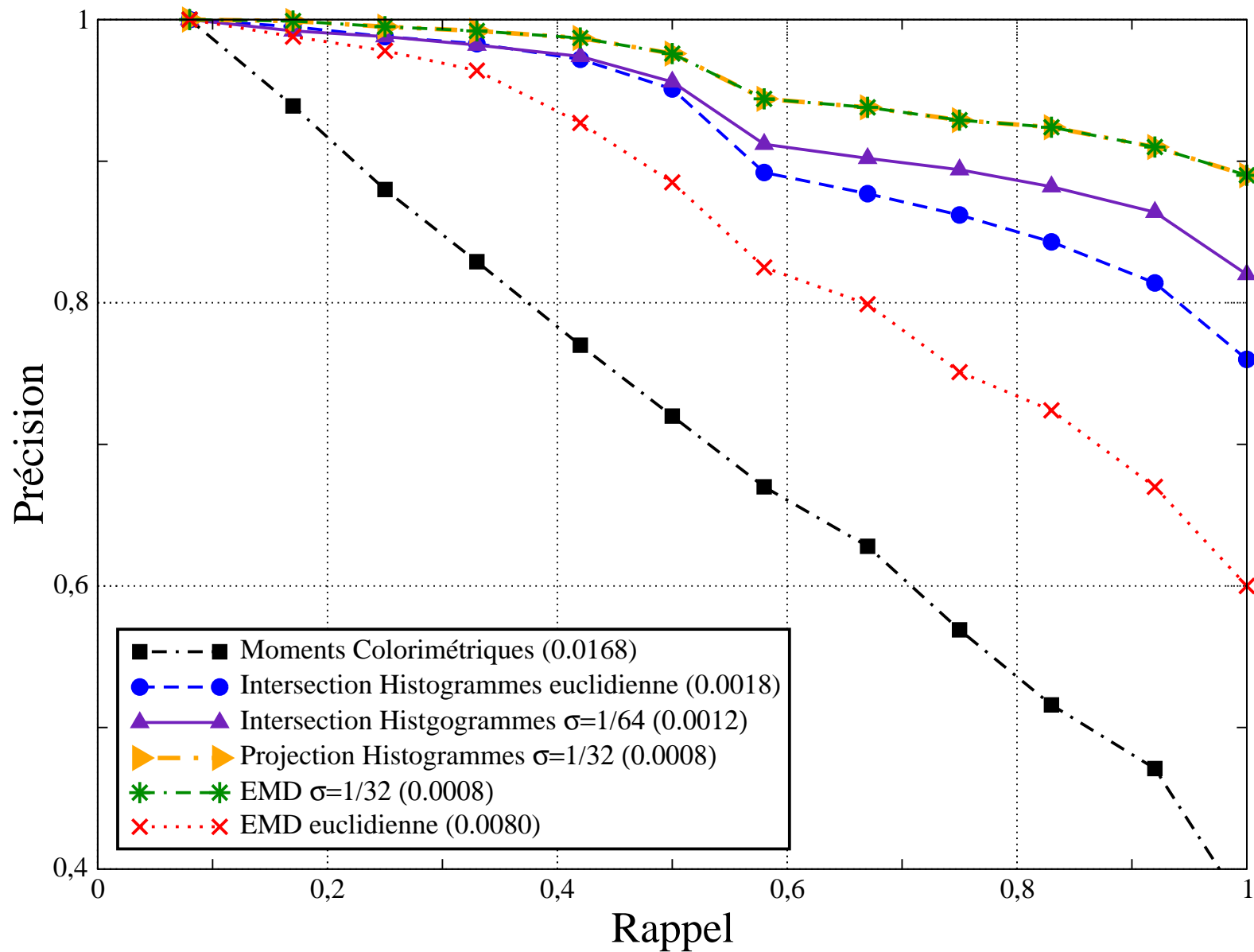
# Résultats - Couleur

Base Columbia en  $L^*a^*b^*$  avec 32 couleurs

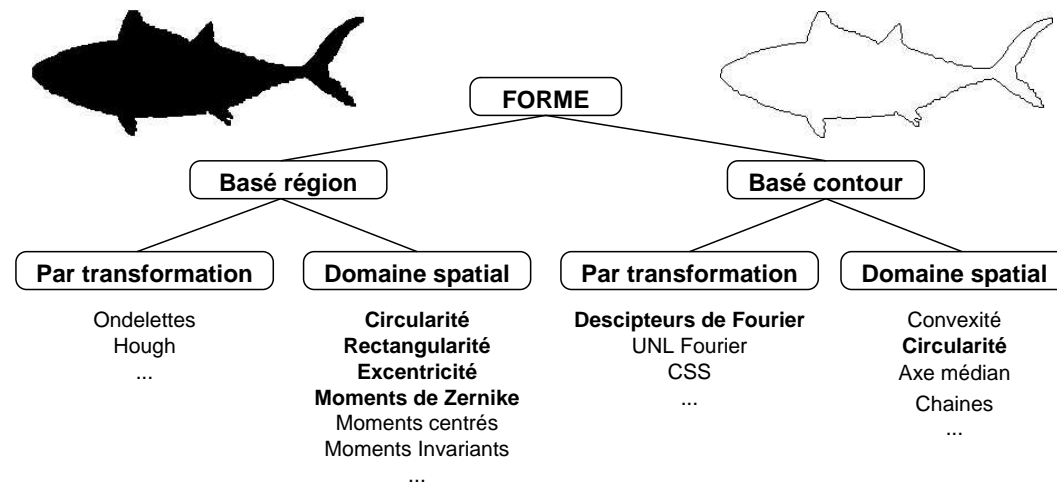


# Résultats - Couleur

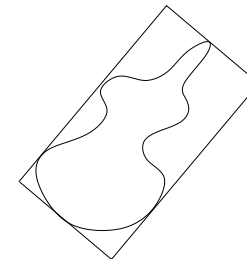
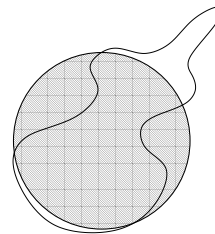
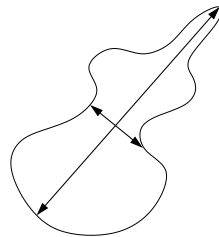
Base locale en  $L^*a^*b^*$  avec 32 couleurs



# Description Forme

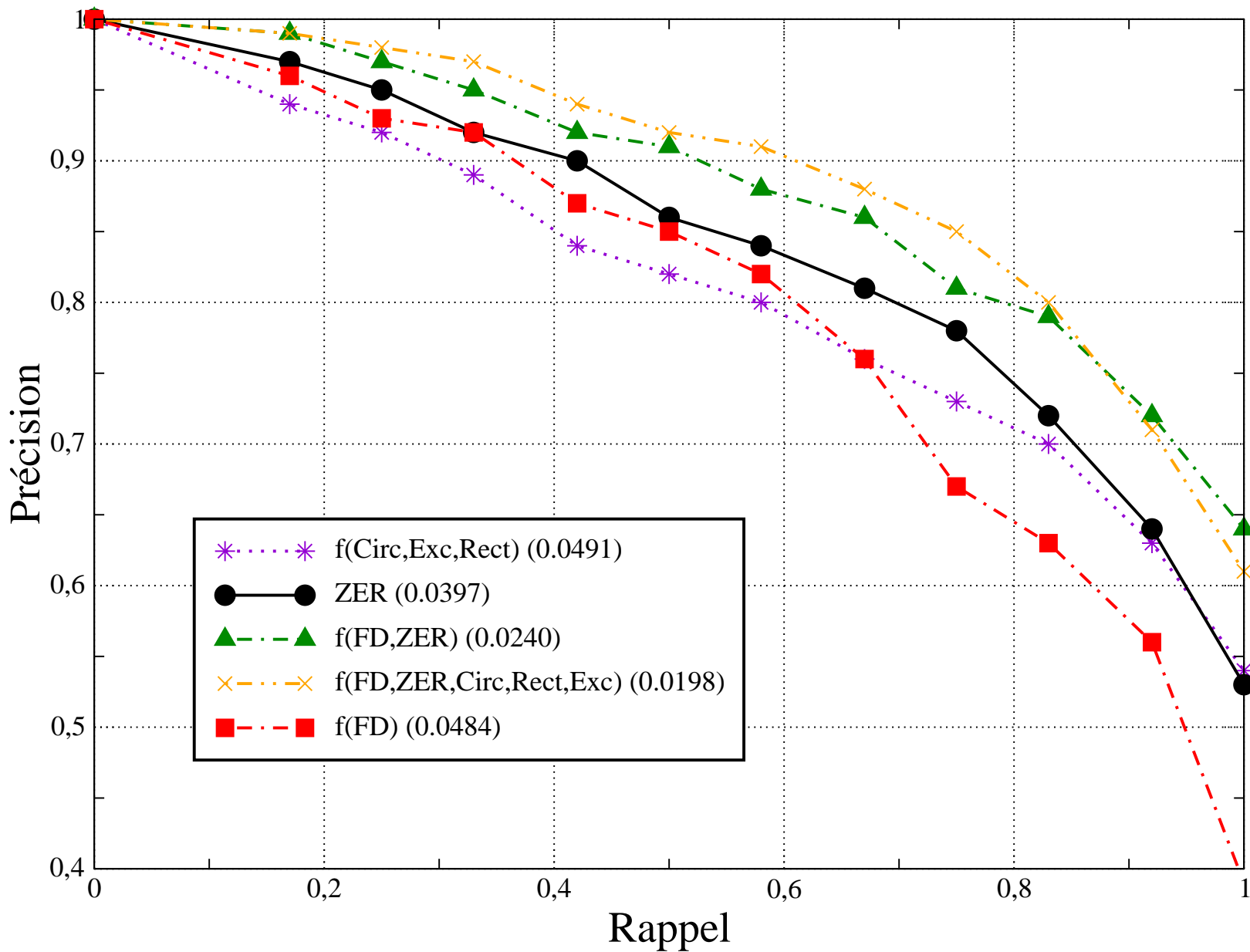


- ▶ Diverses techniques testées avec résistance à l'occultation
- ▶ Combinaison d'informations complémentaires :
  - ▶ Région : moments de ZERNIKE
  - ▶ Contour : descripteurs de FOURIER
  - ▶ Globale : excentricité, circularité, rectangularité



# Résultats - Forme

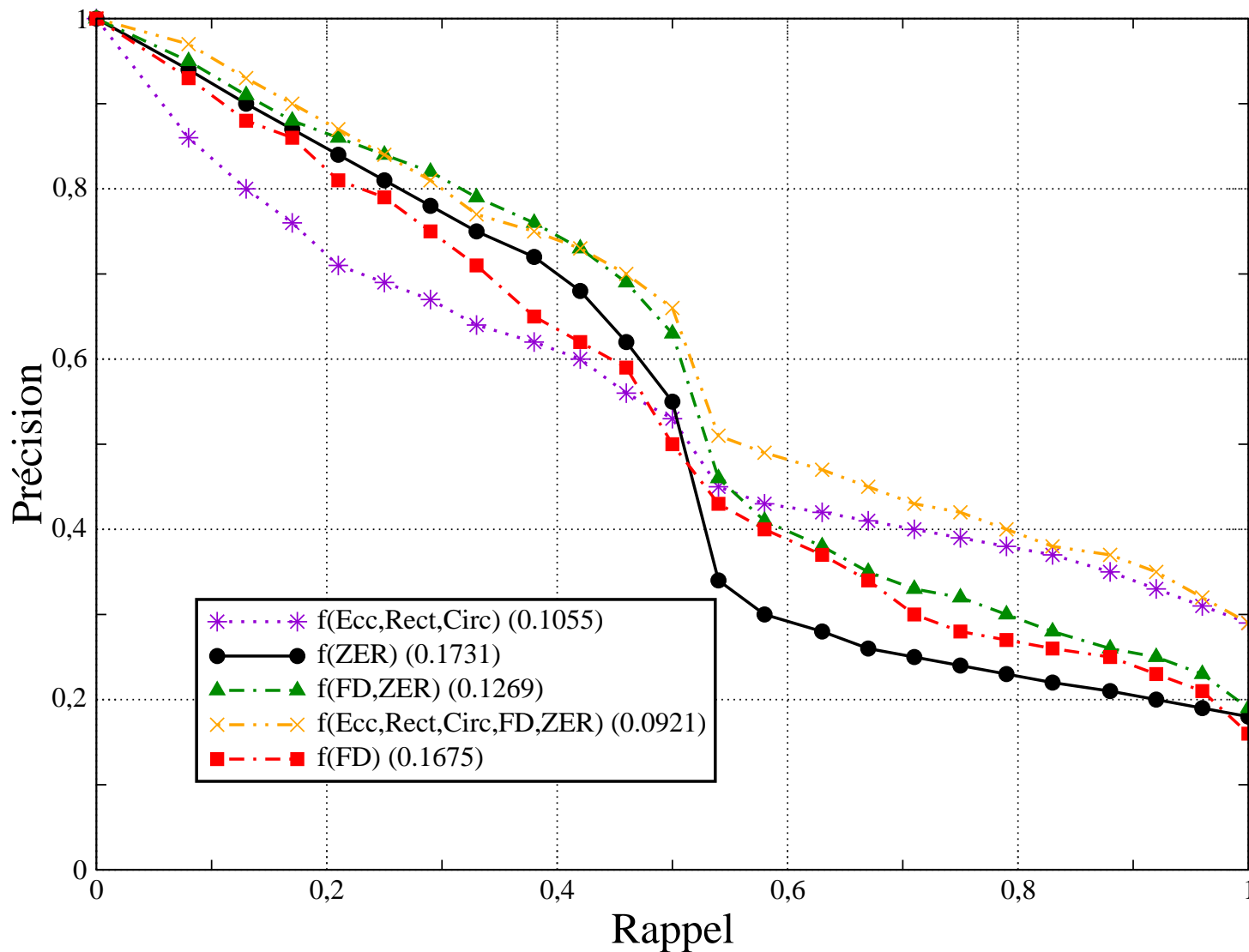
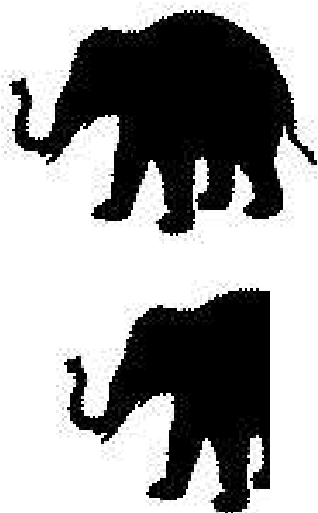
Base SIID

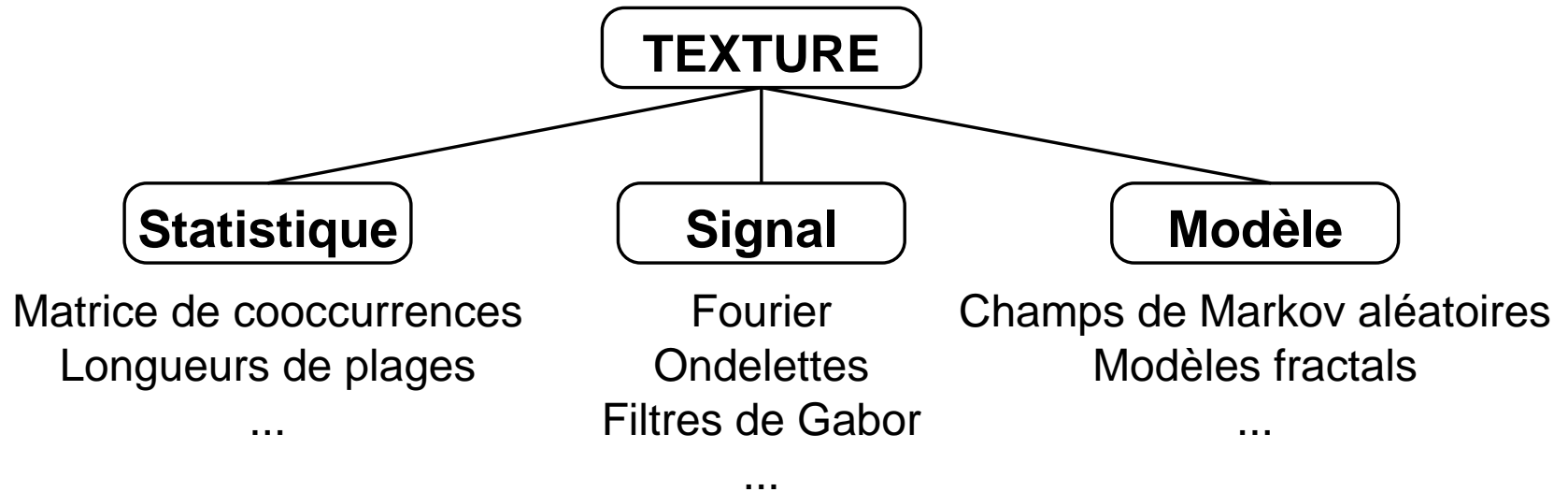




# Résultats - Forme

Base SIID - occultation de 33%





- ▶ Images médiévales : texture peu pertinente
- ▶ Étude rapide des descripteurs (MEASTEX)
- ▶ Forte variabilité des techniques aux types de textures étudiées
- ▶ Choix : matrice de cooccurrences
  - ▶ Simplicité
  - ▶ Additivité



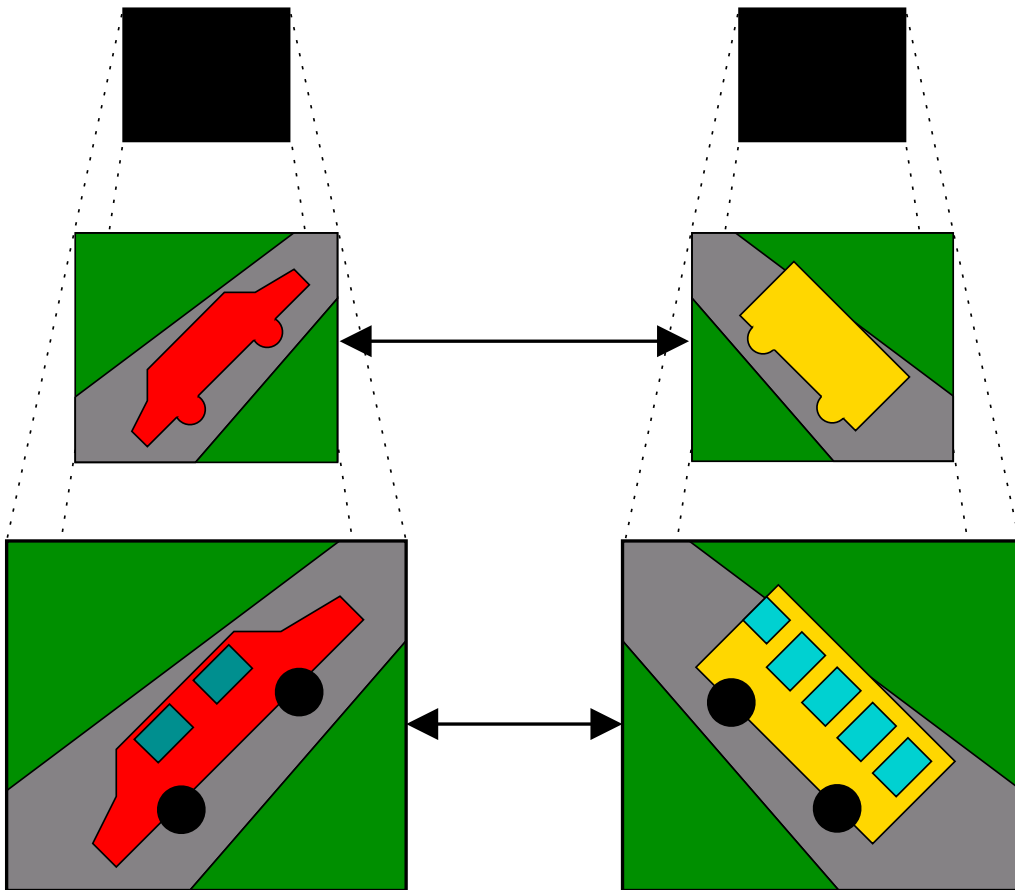
# Plan

---

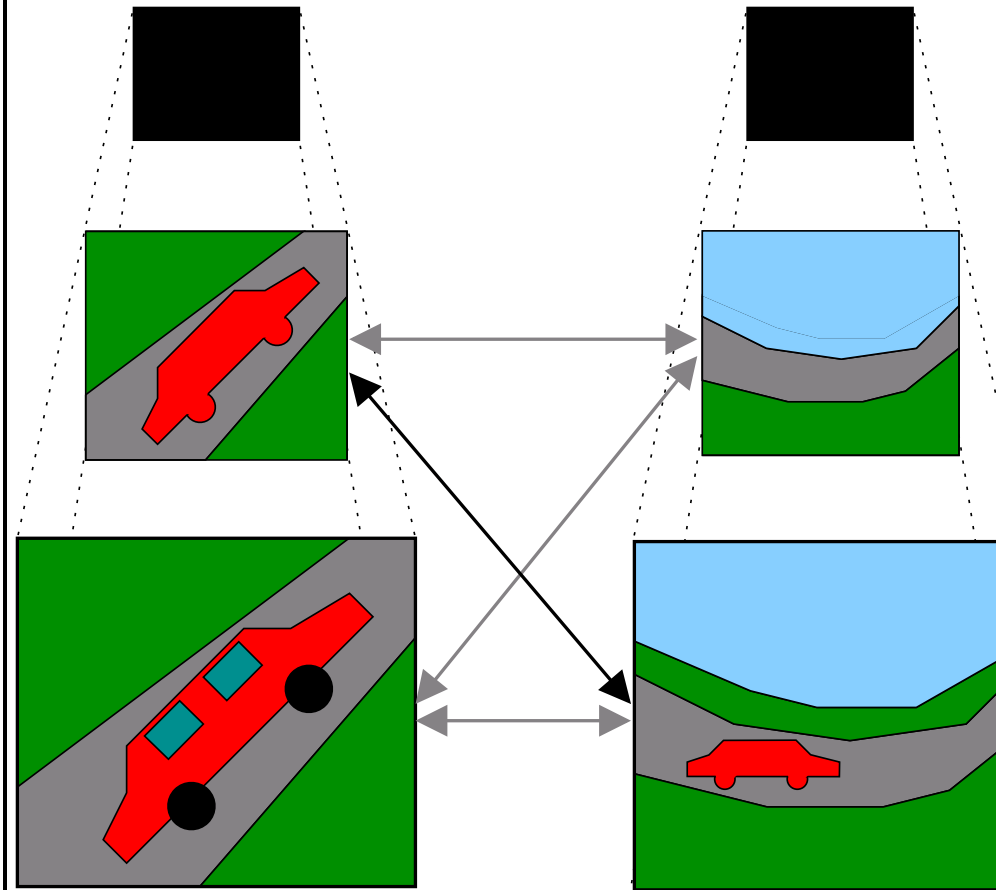
1. Introduction
2. Décrire les images pour indexer
3. Segmentation multi-échelles
4. Description des régions
5. Utilisation de la structure pour l'indexation
6. Conclusion / Perspectives

# Utilisation des échelles

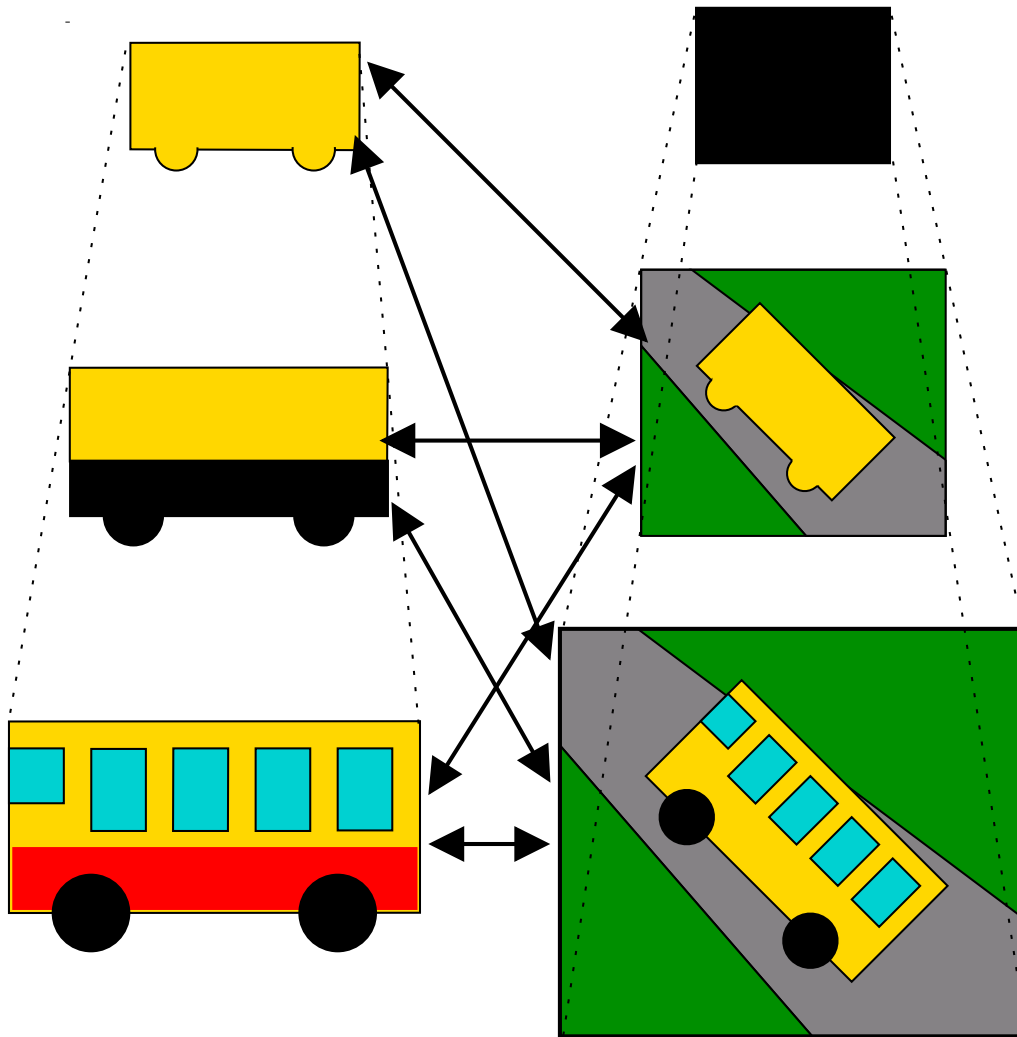
- ▶ Bases spécialisées
- ▶ Cohérence entre les niveaux



- ▶ Bases généralistes
- ▶ Non cohérence entre les niveaux



# Recherche d'objets



- ▶ Extraction de la sous-pyramide correspondante à l'objet
- ▶ Comparaison avec la pyramide de l'image

# Comparaison des niveaux

- ▶ Sans prise en compte de l'arrangement spatial :

- ▶  $R = \{R_i / i \in [1 \dots |R|]\}$  et  $T = \{T_j / j \in [1 \dots |T|]\}$

- ▶  $[D_{ij}] = [d(R_i, T_j)]$ , avec  $i \in [1 \dots |R|]$  et  $j \in [1 \dots |T|]$

- ▶  $p_{NivSim}(R, T) = \frac{1}{|R|} \sum_{i=1}^{|R|} \left[ \min_{j \in [1 \dots |T|]} D_{ij} \right]$

$$\rightarrow S_{NivSim}(R, T) = \frac{p_{NivSim}(R, T) + p_{NivSim}(T, R)}{2}$$

- ▶ Pas de pondération suivant la taille des régions

- ▶ Peu pertinent pour la recherche d'objets - Projections désordonnées

- ▶ Prise en compte de l'arrangement spatial : *Graph Matching*

- ▶ Complexité plus importante

- ▶ Cohérence pour la recherche d'objets



# Recherche d'objets

Recherche utilisant la structure pyramidale



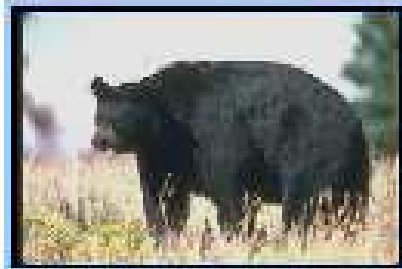
0 - 0  
im251.png



1 - 0.185874  
im245.png



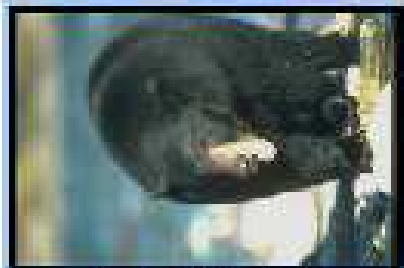
2 - 0.212248  
im246.png



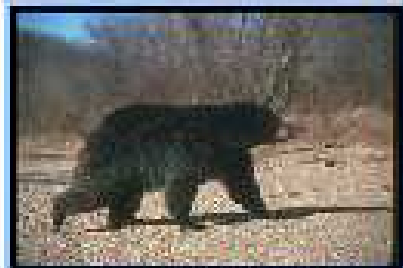
3 - 0.216301  
im242.png



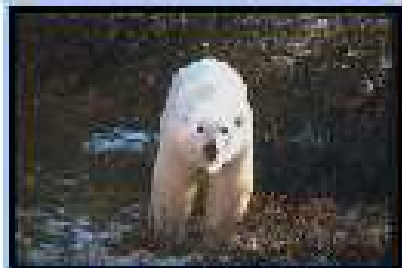
4 - 0.221216  
im252.png



5 - 0.236487  
im261.png



6 - 0.244637  
im250.png



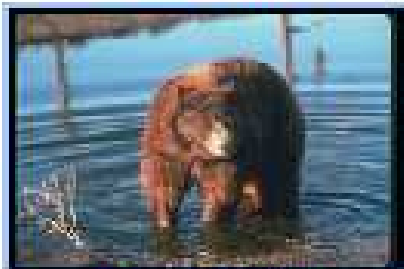
7 - 0.246484  
im229.png



8 - 0.248414  
im212.png



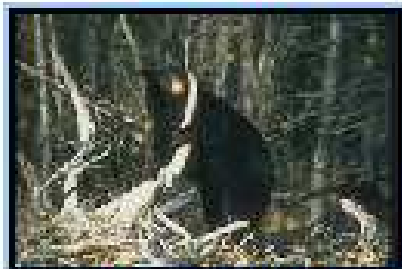
9 - 0.249736  
im243.png



10 - 0.251046  
im265.png



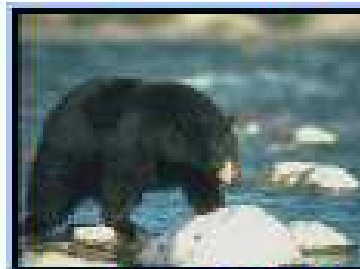
11 - 0.253829  
im240.png



12 - 0.254868  
im266.png



13 - 0.256119  
im237.png



14 - 0.256919  
im253.png

# Fresques médiévales

Recherche utilisant la structure pyramidale



0 - 0  
apotres/00000862.png



1 - 0.118877  
apotres/00000860.png



2 - 0.12311  
apotres/00000865.png



3 - 0.131774  
apotres/00000869.png



4 - 0.134906  
apotres/00000920.png



5 - 0.138701  
apotres/00000872.png



6 - 0.138892  
apotres/00000885.png



7 - 0.139018  
apotres/00000939.png



8 - 0.142028  
apotres/00000852.png



9 - 0.142307  
frise/00000934.png



10 - 0.142812  
apotres/00000958.png



11 - 0.144051  
frise/00000896.png



12 - 0.14559  
frise/00000932.png



13 - 0.14567  
apotres/00000922.png



14 - 0.1465  
apotres/00000972.png



# Détection de sites

---

- ▶ Retrouver la provenance d'une image inconnue
- ▶ Pour chaque site : détermination d'un ensemble restreint d'images caractéristiques
- ▶ Mise en concurrence de deux approches complémentaires :
  - ▶ Approche globale fondée sur la couleur
  - ▶ Approche pyramidale complète
- ▶ Problème de zone d'intérêt dans les images - tout ce qui n'est pas peinture murale ne doit pas être pris en compte

# Détection de sites

Sélection de l'image : baptistere\_saint\_jean/00000852.png

Nombre d'images à voir : 10

GO

RANDOM

Image recherchée :



00000915.png

Le site trouvé est **baptistere\_saint\_jean** (à 100%)

Les images les plus proches sont :



1-

baptistere\_saint\_jean/  
00000916.png



2-

baptistere\_saint\_jean/  
00000869.png



3-

baptistere\_saint\_jean/  
00000919.png



4-

baptistere\_saint\_jean/  
00000885.png



5 - baptistere\_saint\_jean/  
00000920.png



6-

baptistere\_saint\_jean/  
00001025.png



7-

baptistere\_saint\_jean/  
00000965.png



8-

baptistere\_saint\_jean/  
00000922.png



9-

baptistere\_saint\_jean/  
00000910.png



10 - baptistere\_saint\_jean/  
00000909.png





# Plan

---

1. Introduction
2. Décrire les images pour indexer
3. Segmentation multi-échelles
4. Description des régions
5. Utilisation de la structure pour l'indexation
6. **Conclusion / Perspectives**



# Conclusion

---

- ▶ Structure :
  - ▶ Nouvelle structure de représentation multi-échelles pour l'indexation :
    - ▶ Notion d'arrangement des régions
    - ▶ Notion de composition
  - ▶ Algorithme de segmentation multi-échelles
- ▶ Description :
  - ▶ Étude des descripteurs de régions : couleur/forme/texture
  - ▶ 3 propositions pour la description couleur
- ▶ Indexation :
  - ▶ Méthodes d'utilisation de la structure pour la recherche d'images ou d'objets
  - ▶ Résultats encourageants
- ▶ Applications CESC M :
  - ▶ Recherche d'objets
  - ▶ Détection de provenance d'images inconnues



- ▶ Structure de représentation :
  - ▶ Outil de segmentation multi-échelles « parfait »
  - ▶ Étude plus poussée des descripteurs de forme et de texture :
    - ▶ Recherche de l'additivité
    - ▶ Amélioration de la précision de la description suivant l'application
  - ▶ Graphes d'adjacences ne décrivent pas complètement la topologie de l'image (doubles adjacences, ordonnancement)
    - Extension vers les cartes généralisées
- ▶ Utilisation de la structure :
  - ▶ *Graph matching* → appariement de pyramides
  - ▶ Définition de critères simples de rejet
- ▶ Projets applicatifs :
  - ▶ Extension à la vidéo avec un suivi d'objets
    - Réduction de la taille de la description des scènes
  - ▶ CESC M : étude et caractérisation des mains

