

Techniques d'indexation et recherche multimédia

Chap3- Indexation des images

Amira BELHEDI

amira.belhedi@istic.ucar.tn

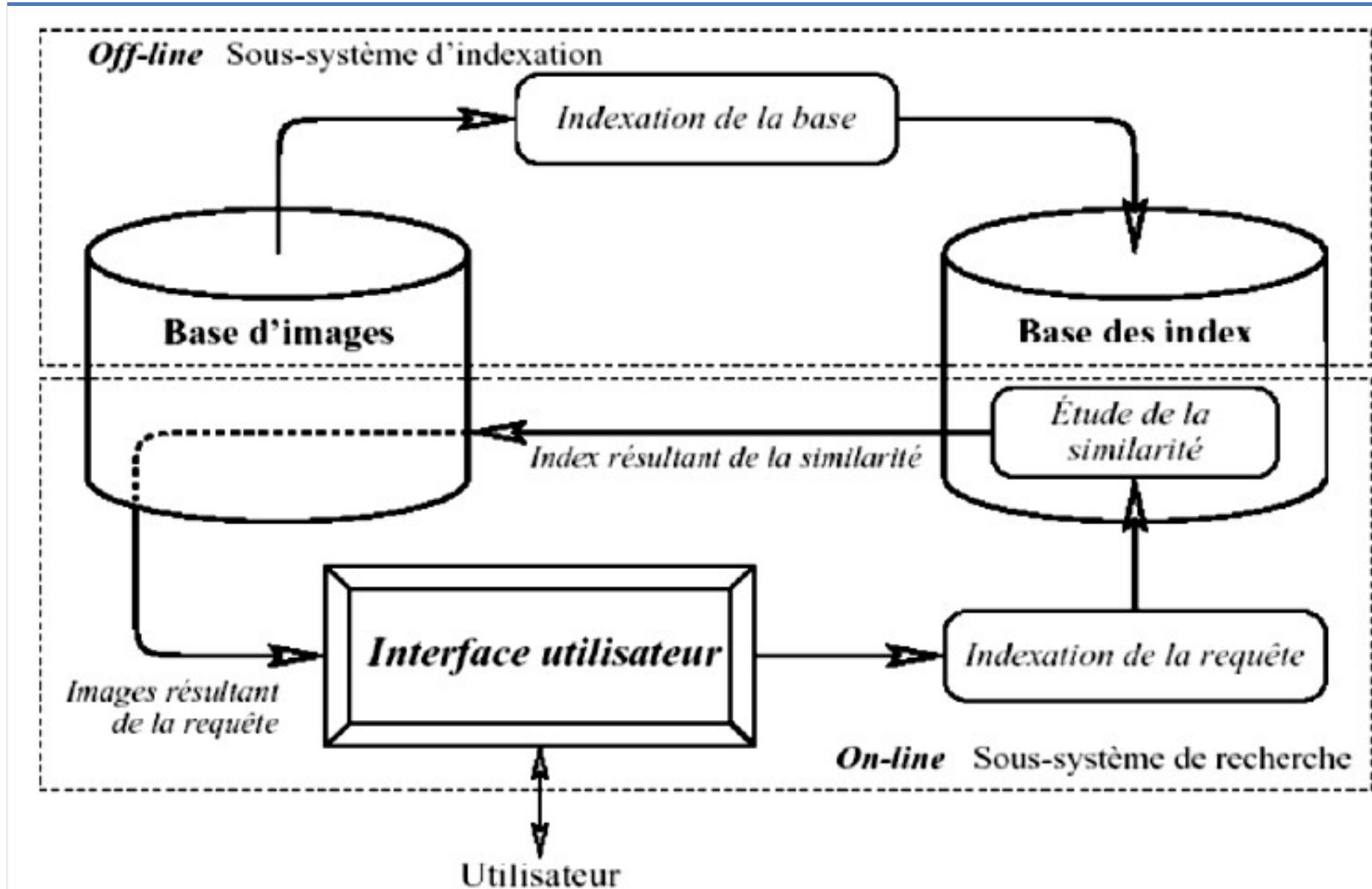
Fadoua MHIRI

fadoua.mhiri@istic.ucar.tn

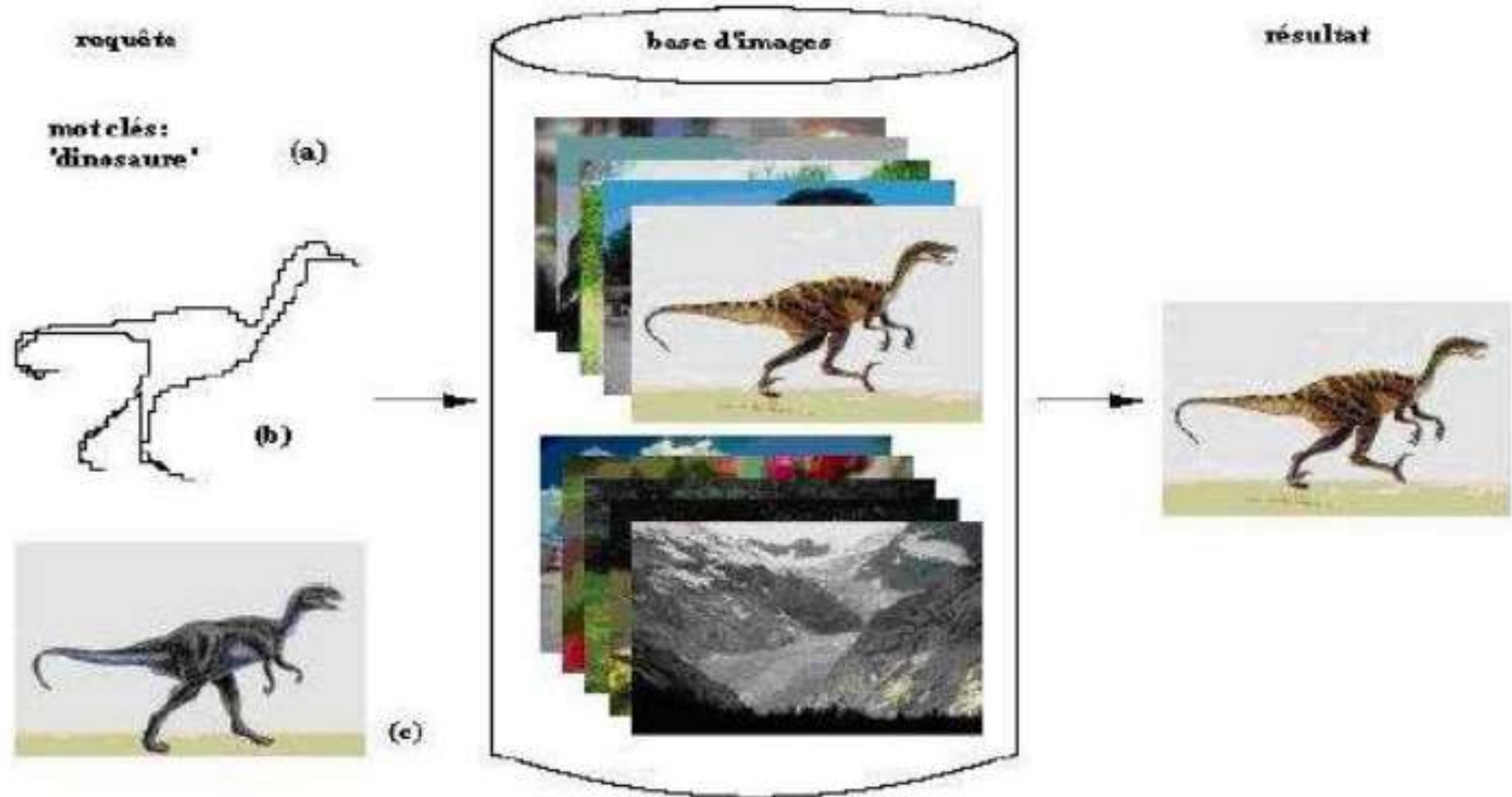
Wafa ABID

wafa.abid@istic.ucar.tn

Systeme d'Indexation et de Recherche d'Images



Trois façons de faire une requête pour la recherche d'images



requête par « ***mots clés*** » / requête par « ***esquisse*** » / requête par « ***exemple*** »

Plan

- Indexation textuelle des images
- Indexation des images par le contenu
 - Types de requêtes visuelles
 - Descripteurs visuels
 - Descripteurs globaux: extraits de l'image entière,
 - Histogramme
 - Texture
 - Descripteurs locaux: extraits d'une partie de l'image,
 - Régions de l'image: obtenues par segmentation
 - Points d'intérêt

Indexation textuelle des images

- Indexation manuelle
- Indexation automatique
 - à partir du texte associé à l'image
 - à partir du contenu visuel de l'image

Indexation textuelle manuelle

- Réalisée par un iconographe ou par les utilisateurs qui souhaitent décrire leurs images.
- Le principal problème est le **choix des termes**.
 - La probabilité pour une même image que le même terme soit choisi par deux individus est inférieur à 20%
 - Une même image peut avoir plusieurs sens (polysémie de l'image) et peut contenir plusieurs thèmes.
- L'indexation manuelle d'une image est donc **subjective** → plusieurs indexations sont donc possibles.

Indexation textuelle automatique

- L'indexation textuelle automatique des images consiste à:
 - associer des mots à une image
 - au moyen d'un système informatique
 - sans aucune intervention humaine.

Indexation textuelle automatique

- A partir de la transcription d'un message audio
 - Une manière originale et plus évidente pour les utilisateurs d'indexer textuellement des images est d'enregistrer un message audio qui décrit leur contenu
 - Il ne reste plus alors au système qu'à transcrire ce message en texte pour obtenir une indexation automatique.
- Il existe deux approches d'indexation textuelle automatique
 - à partir du **texte associé à l'image**,
 - à partir du **contenu visuel de l'image**.

Indexation textuelle à partir du texte associé à l'image

- Possible seulement lorsque les images sont associées à du texte
- Exemple pour les images du web → indexation utilisant le:
 - Texte de l'attribut ALT de la balise IMG d'une image (code HTML)
 - Nom de l'image → problème: la plupart des images ne sont pas nommées de façon pertinente, exemple img001.jpg
 - Texte de la page: titre/mots fréquents, etc. → problème: toutes les images d'une même page web sont indexées avec les mêmes mots.
 - Texte proche de l'image → problème: ce texte n'est pas forcément approprié pour décrire l'image
 - Une combinaison de toutes ces informations

Indexation textuelle à partir du contenu visuel de l'image

- Appelée auto-annotation par le contenu
- Utilisation des méthodes de classification
 - Apprentissage sur des images
 - Prédiction sur de nouvelles images.
- L'auto-annotation d'images est un problème toujours ouvert → même les meilleurs systèmes ne sont pas capables d'apporter un gain de plus de 40% sur le modèle empirique.

Indexation textuelle des images

- Avantage
 - donne la possibilité à l'utilisateur de poser sa requête **sous forme de texte** (requête par mots clés) lui permettant ainsi d'exprimer son besoin d'information avec du texte

Indexation textuelle des images

- Limites
 - l'indexation manuelle est très coûteuse en terme de temps
 - l'indexation automatique : à partir du texte associé à l'image ou à partir du contenu visuel de l'image → commet beaucoup d'erreurs → Les images sont mal annotées

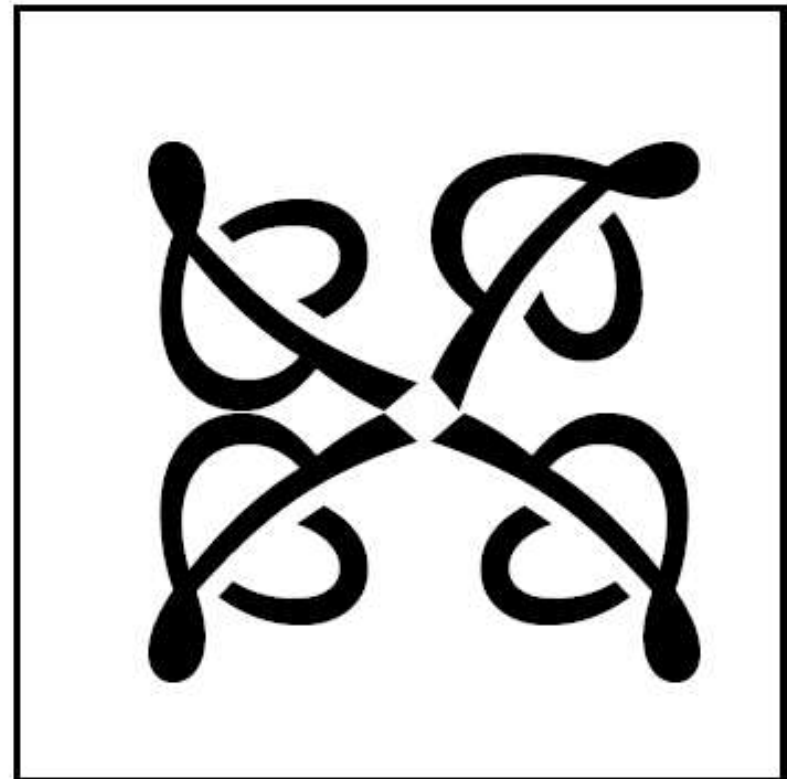
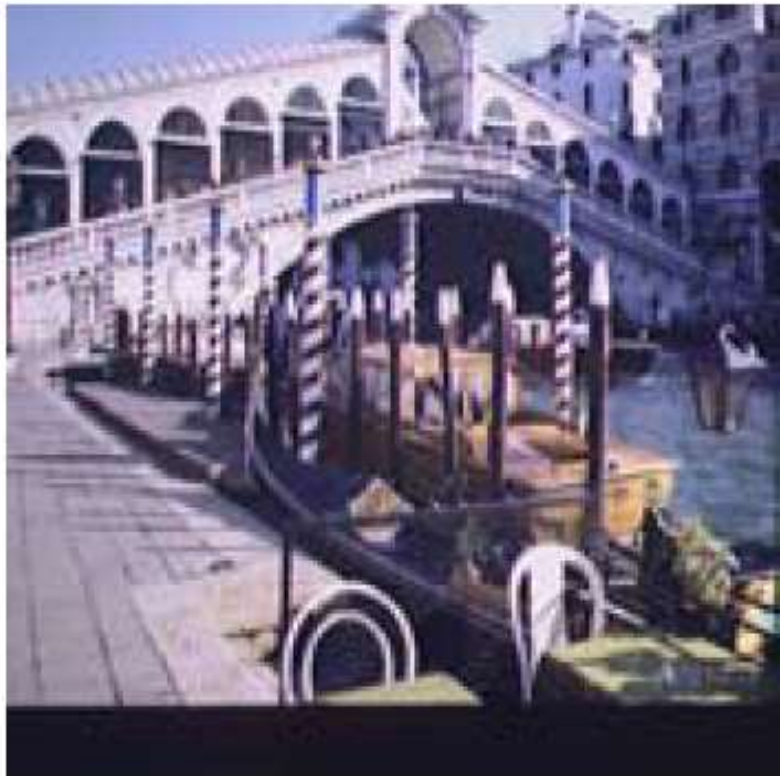
Limite de l'indexation textuelle des images

- Dépendance du contexte



Limite de l'indexation textuelle des images

- Quels mots-clés utilisés ??



Indexation des images par le contenu

Indexation des images par le contenu

- Utiliser des descripteurs multimédias
- Exemples de descripteurs pour l'indexations des images par le contenu
 - Descripteurs visuels pour les base d'images génériques
 - couleur, texture, forme, position et relations spatiales
 - Segmentation et points d'intérêt pour les base d'images spécifiques
 - Détection et signatures de *visages*,
 - Signatures d'*empreintes digitales*

Intérêt de l'indexation des images par le contenu

- Indépendante de la langue de recherche
- Description moins subjective que le texte
- Parfois plus riche que le texte
- Souvent plus efficace

Types de requêtes visuelles

- Un utilisateur peut avoir différents objectifs lorsqu'il effectue une recherche d'images.
 - rechercher une **image en particulier**
 - recherche de la photographie du sphinx
 - rechercher une **catégorie d'images**,
 - recherche des beaux couchers de soleil.
 - rechercher une image avec des **caractéristiques non-définies**
 - rechercher un fond d'écran avec des couleurs pastels
- ➔ **Navigation**: l'utilisateur n'a pas une idée précise de ce qu'il veut, il parcourt les images sans objectif précis, et peut même changer d'idée au cours de la recherche.

Types de requêtes visuelles

- Pour trouver l'information qu'il recherche, l'utilisateur dispose d'au moins trois types de moyens :
 1. Les requêtes **par exemple visuel**
 2. Les requêtes **directes**
 3. La **navigation** dans les images.

Requêtes par exemple visuel

- L'utilisateur fournit un exemple, le système en extrait des descripteurs visuels, puis recherche les images ayant des descripteurs similaires. L'exemple peut être :
 - une **image** : la requête est composée d'une image entière,
 - une **région** : recherche des régions d'images qui ressemble à la région de l'image sélectionnée par l'utilisateur
 - une **composition de régions** : l'utilisateur choisit les régions parmi des régions construites à partir des régions de la base
 - une **composition spatiale de régions** : l'utilisateur recherche une image ayant la région en haut de la même couleur que l'objet A, et la région à droite de la même texture que l'objet B
 - un **croquis dessiné par l'utilisateur** : descripteur utilisé est la forme

Types de requêtes visuelles

- les requêtes directes: elles sont composées de descripteurs visuels, par exemple,
 - 25% de rouge, 30% de bleu et 45% de vert
 - la couleur a un poids de 75% et la texture de 25%
- la navigation dans les images.

Extraction des informations visuelles

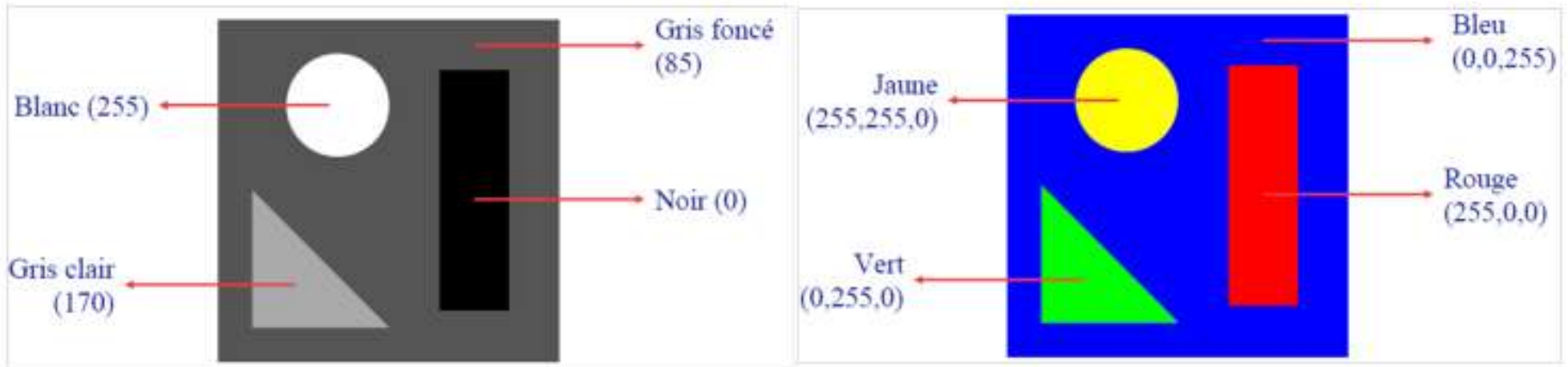
- La représentation des descripteurs visuels doit être :
 - **discriminante** : elle doit permettre de bien différencier les images différentes,
 - **invariante** : deux régions de tailles différentes ou prises avec des luminosités différentes doivent avoir des représentations très proches,
 - **compacte** : les espaces de grande dimension posent des problèmes de stockage des données → Il est donc intéressant de réduire au maximum la dimension des signatures.

Descripteurs visuels

- Descripteurs globaux: extraits de l'image entière, exemple:
 - Histogramme, texture...etc
- Descripteurs locaux: extraits d'une partie de l'image, exemple:
 - Les régions de l'image obtenues par segmentation de l'image entière, ou par recherche de régions d'intérêt,
 - Points d'intérêt

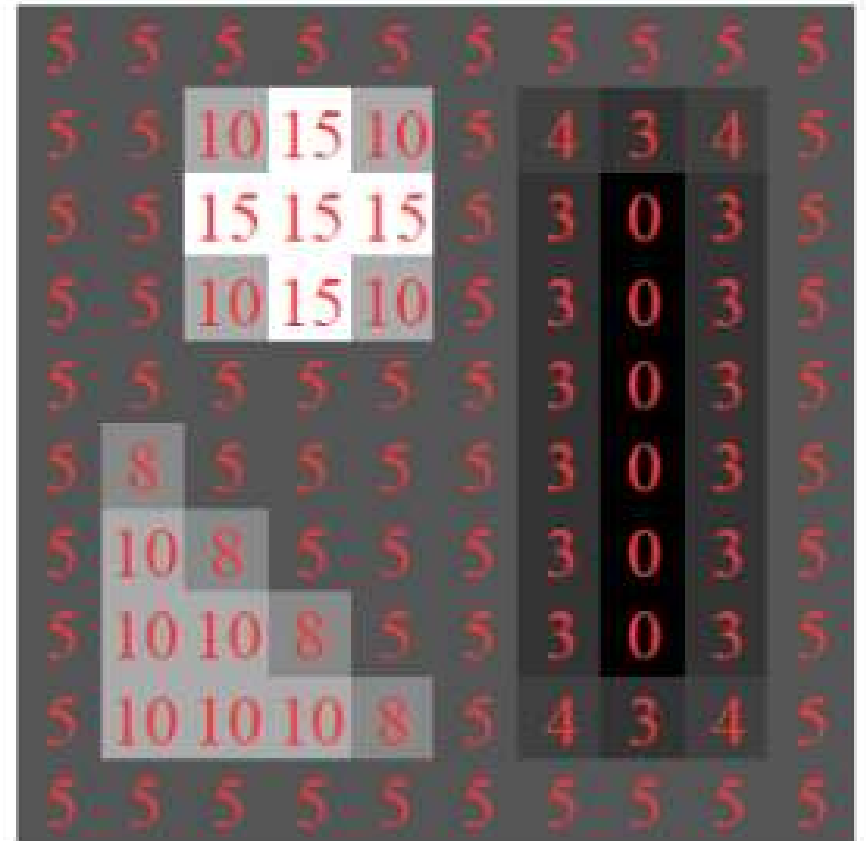
Histogramme

- L'histogramme est une représentation pratique de l'image
- Pour les images couleur, un histogramme par composante couleur, exemple, image RGB:
 - Histogramme du rouge R
 - Histogramme du vert G
 - Histogramme du bleu B

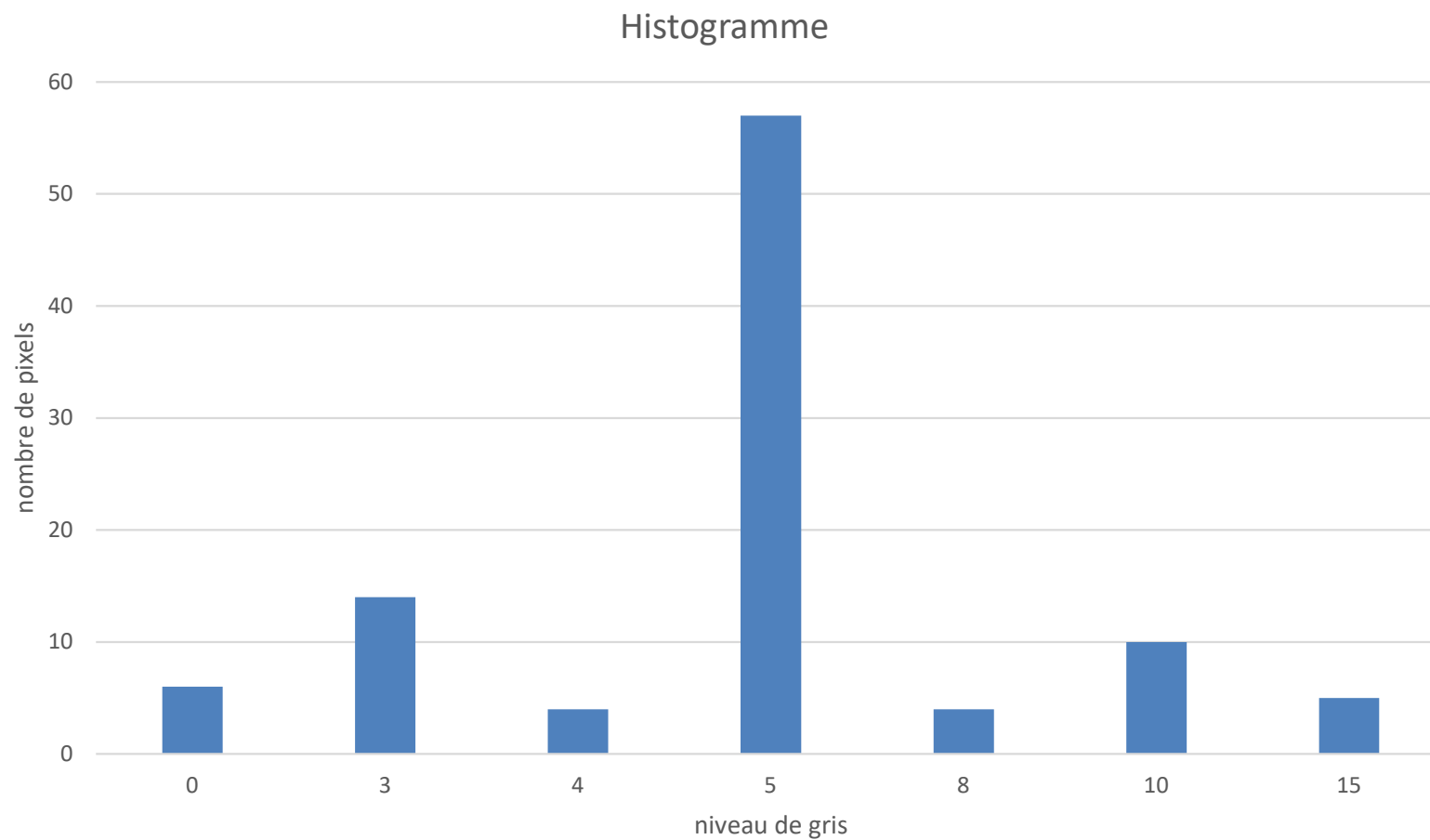


Exercice 1

1. Tracer l'histogramme et l'histogramme normalisé correspondant à cette image
2. Pour indexer une image, on utilise l'histogramme ou l'histogramme normalisé ? Dire pourquoi?
3. Quels sont les avantages et les limites de l'utilisation de l'histogramme dans l'indexation d'images, donner des exemples?

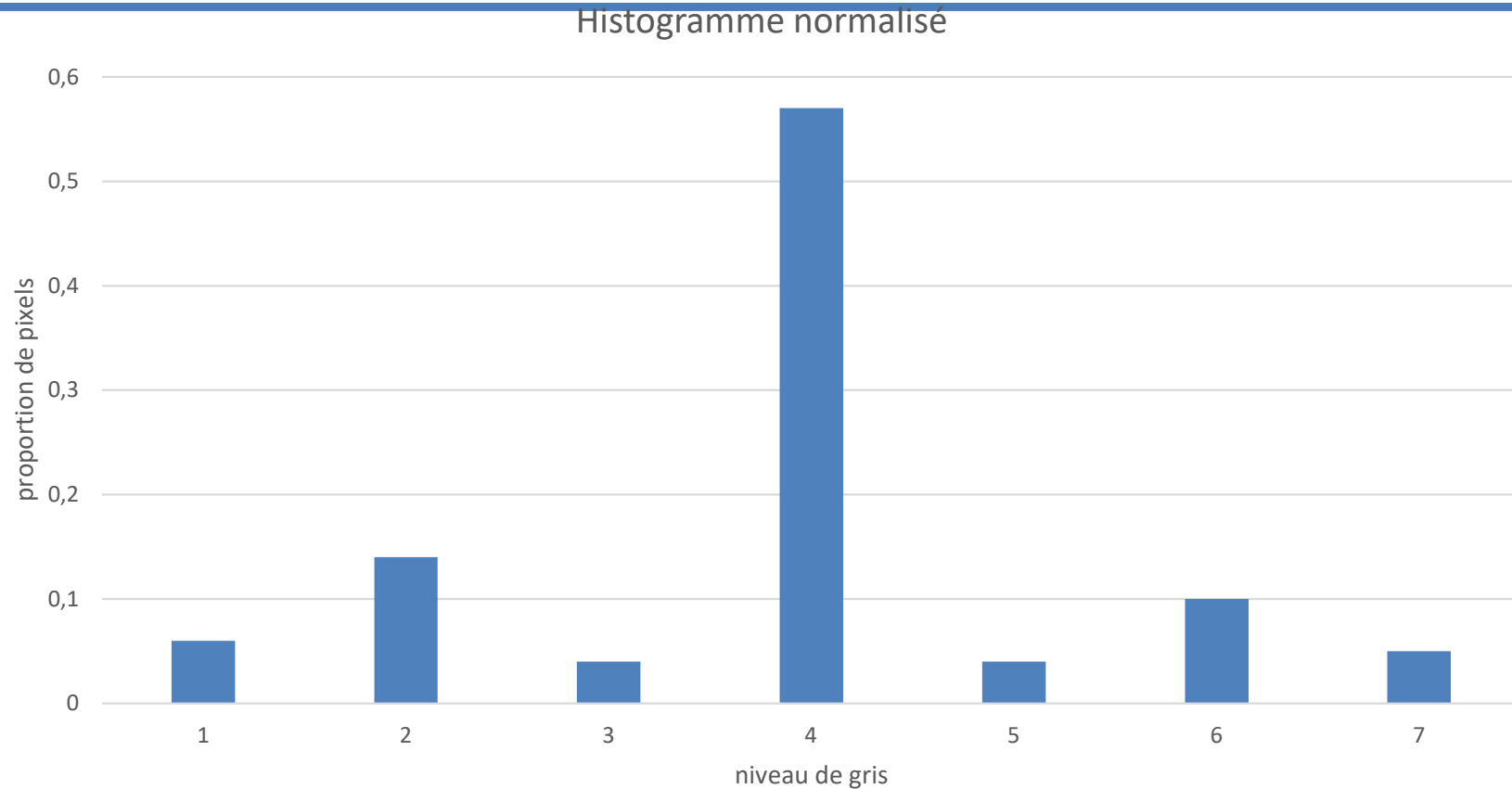


Corrigé de l'exercice 1



Histogramme

Corrigé de l'exercice 1

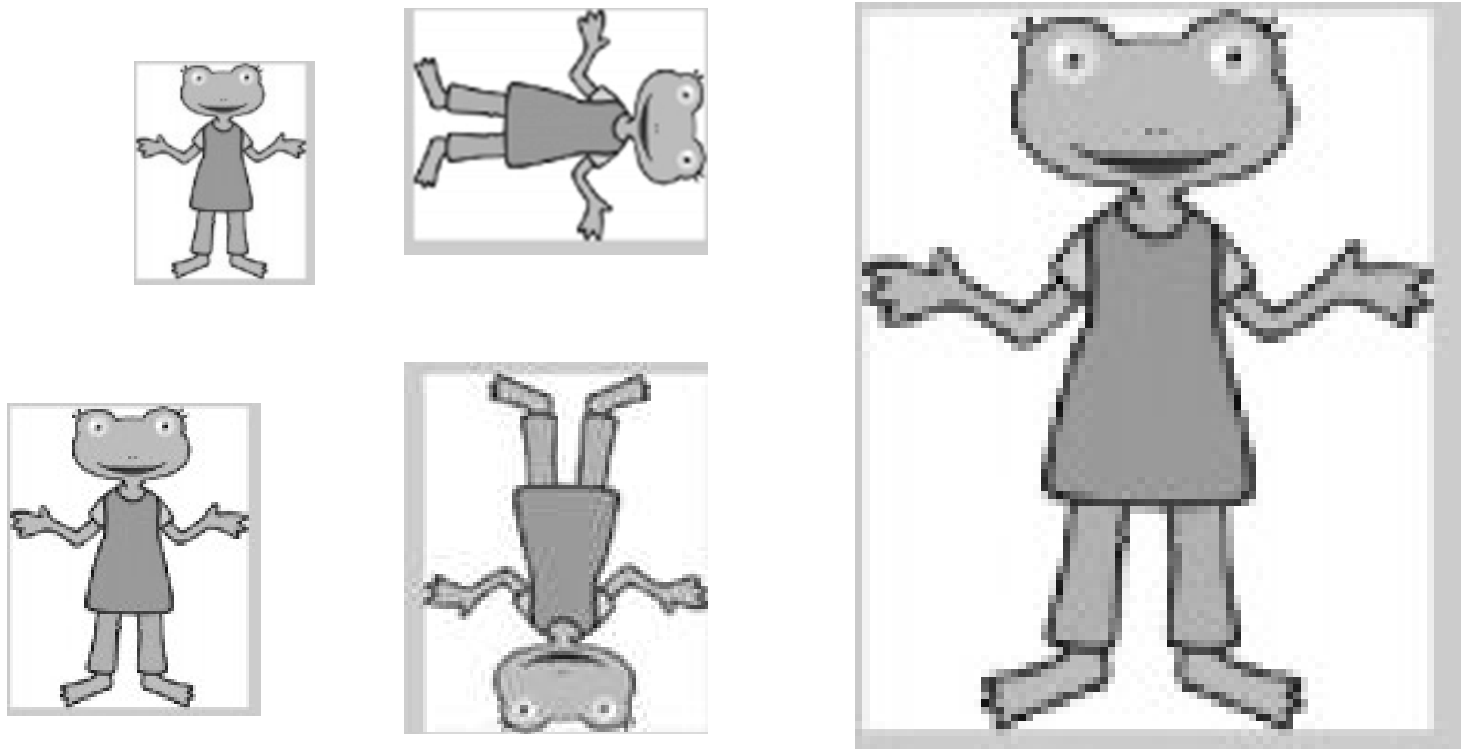


Histogramme normalisé

- Afin de rendre l'histogrammes invariant à la taille de l'image et des régions → **histogramme normalisé**

Question

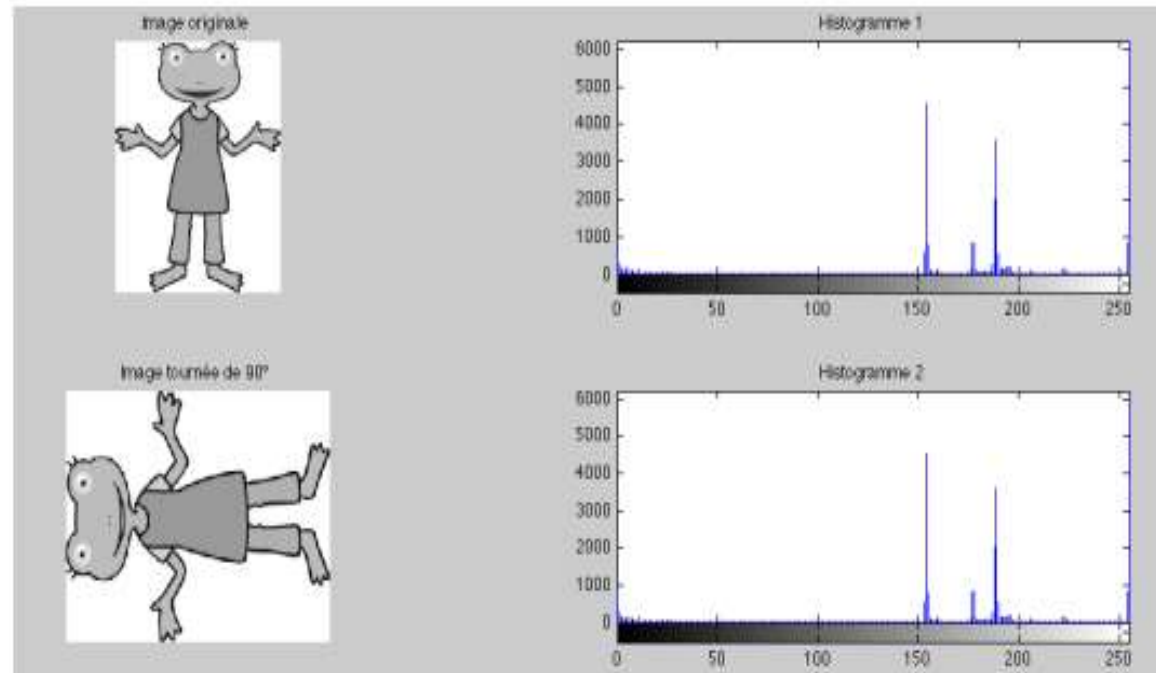
- Que pensez-vous de l'histogramme normalisé de ces images



- En déduire l'avantage d'utiliser l'histogramme normalisé dans l'indexation d'images

Avantage de l'histogramme

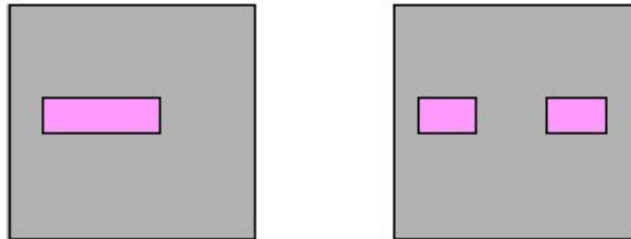
- Insensible aux changements d'orientation, de taille et de position des régions



- Autre avantage:
 - Réduction de la taille de la base d'indexes et du coût de la recherche

Limites de l'histogramme

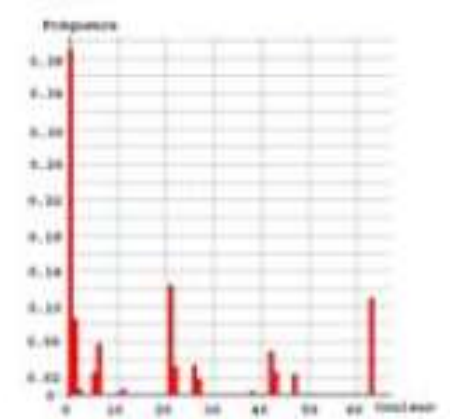
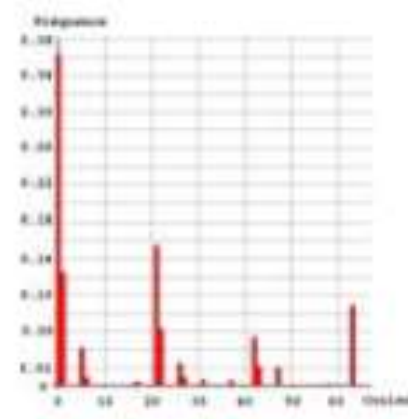
- Est-ce que ces deux images sont visuellement différentes ?
- Est-ce qu'elles ont des histogrammes différents?



- En déduire la limite de l'histogramme

Limites de l'histogramme

- Il ne peut pas distinguer des surfaces couleurs de taille équivalentes mais de répartition différentes.



- L'histogramme ne permet pas une recherche efficace d'objets dans l'image (ne capture pas les relations spatiales entre les régions)
- Indice visuel insuffisant si utilisé seul, car insuffisamment discriminant : pomme rouge vs Ferrari

Limites de l'histogramme

- Solution:
 - Combiner différents types d'informations dans la même signature: exp: couleur et texture
 - => Permet de rendre compte de la répartition spatiale de la couleur
- ➔ Problème: taille de l'espace de description

Texture

- Il n'y a pas de définition complètement satisfaisante de la texture:
 - Zone homogène en un certain sens, assemblage d'entités élémentaires formant un tout.



Texture

- Il existe plusieurs méthodes pour analyser la texture
 - Méthodes statistiques
 - Matrice de cooccurrence
 - ...
 - Méthodes fréquentielles (traitement du signal)
 - Transformée de Fourier
 - Filtre de Gabor
 - Transformée en Ondelettes
 - ...

Matrice de cooccurrence

- Identifier les répétitions de niveaux de gris selon une distance (pas) et une direction
- Matrice de cooccurrence : matrice de taille $N_g \times N_g$
 - N_g étant le nombre de niveaux de gris de l'image, exemple (256 x 256) (réduction souvent a des tailles 8x8, 16x16 ou 32x32)
- Pour un voisinage (dx, dy), la matrice de cooccurrence $M(dx, dy)$ est donnée par :

$$M [dx, dy] (u, v) = \frac{1}{(N_x dx)(N_y dy)} \sum_{i,j} 1 [I(i, j) = u \& I(i + dx, j + dy) = v]$$

- avec (N_x, N_y) : taille de l'image et (u, v) : niveaux de gris de l'image

Matrice de cooccurrence

- Plusieurs matrices, pour chaque distance (pas) et direction
 - Distance : 1, 2, 3, 4...
 - Direction : 0° , 90° , 45° , 135° , ...
- Temps de calcul de ces matrices est assez long
- Méthode de calcul de la matrice de cooccurrence:
 - On parcourt l'image et pour chaque couple de pixels formé avec la distance et la direction données, on incrémente la matrice des cooccurrences de 1

Exercice 2

- Soit l'image I définie par:
- Donner la matrice de cooccurrence pour distance 1 et direction 0°

1	4	4	3
4	2	3	2
1	2	1	4
1	2	2	3

Image

	1	2	3	4
1	?	?	?	?
2	?	?	?	?
3	?	?	?	?
4	?	?	?	?

Matrice pour distance=1
et direction= 0°



	1	2	3	4
1	?	?	?	2
2	?	?	?	?
3	?	?	?	?
4	?	?	?	?

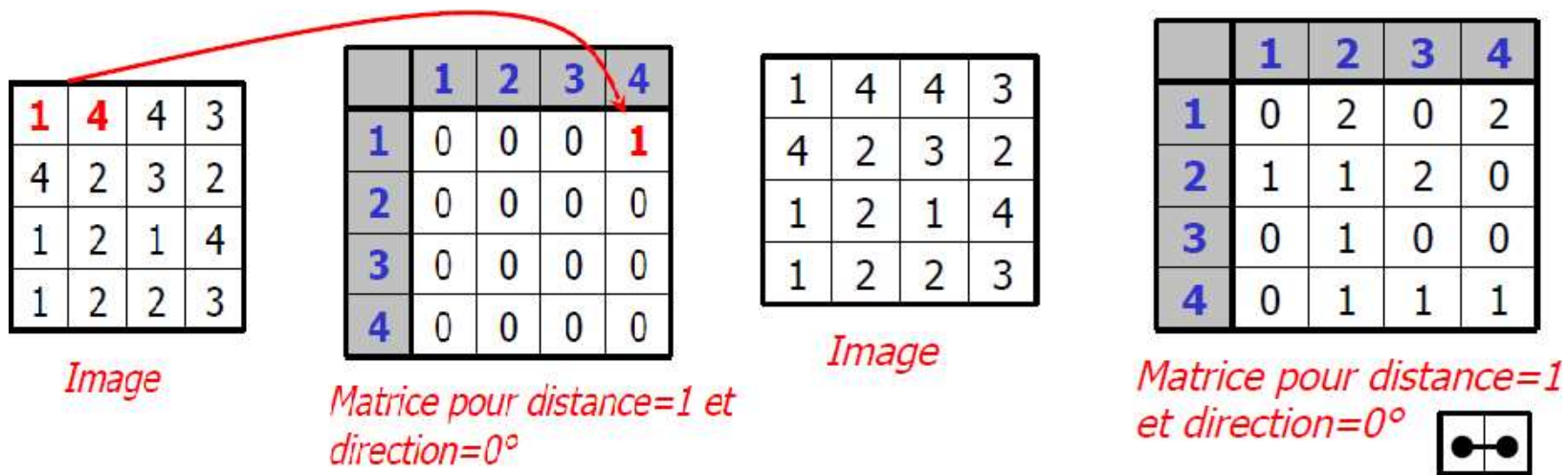
Matrice pour distance=1
et direction= 0°



- Exemple: le 2 de la matrice de cooccurrence (ligne 1 et colonne 4) signifie que l'on trouve deux fois un pixel de valeur 1 de distance 1 et de direction 0° d'un pixel de valeur 4.

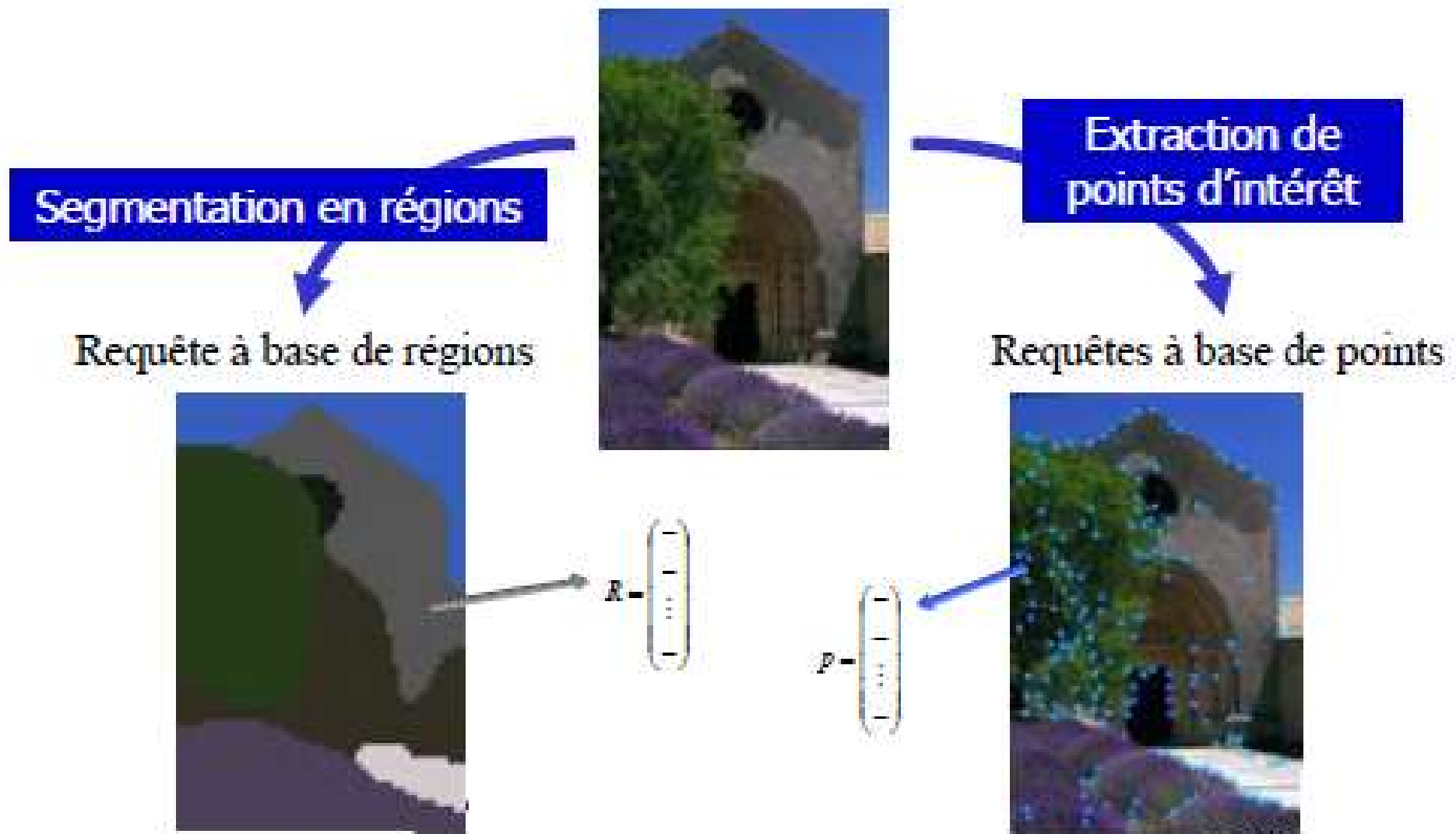
Corrigé de l'exercice 2

- Matrice de cooccurrence (non normalisée) pour distance 1 et direction 0° est:



- le 2 de la matrice de cooccurrence (ligne 1 et colonne 4) signifie que l'on trouve deux fois un pixel de valeur 1 de distance 1 et de direction 0° d'un pixel de valeur 4

Segmentation et points d'intérêts



Segmentation et points d'intérêts

- Segmentation en régions
 - Par approche région
 - Par approche contour
- Extraction des points d'intérêt
 - approches contours
 - approches intensité
 - approche à base de modèle

Segmentation

- Consiste à séparer, les divers composants visibles dans une image, en régions homogènes
- Région homogène: un ensemble de points (pixels) ayant des propriétés communes d'intensité, de texture, de couleur..., qui la différencie d'une région voisine.
- Il y a deux grandes familles d'algorithmes de segmentation d'images
 - Segmentation par approche «région»
 - Segmentation par approche «contour»

Segmentation par approche «région»

- Le principe est de trouver les régions en regroupant les pixels ayant des caractéristiques similaires et en séparant ceux qui sont différents
- Exemples de méthodes de segmentation basées région
 - Croissance de région
 - Split and Merge

➔ ces méthodes ont été vues en traitement d'image (semestre 3)

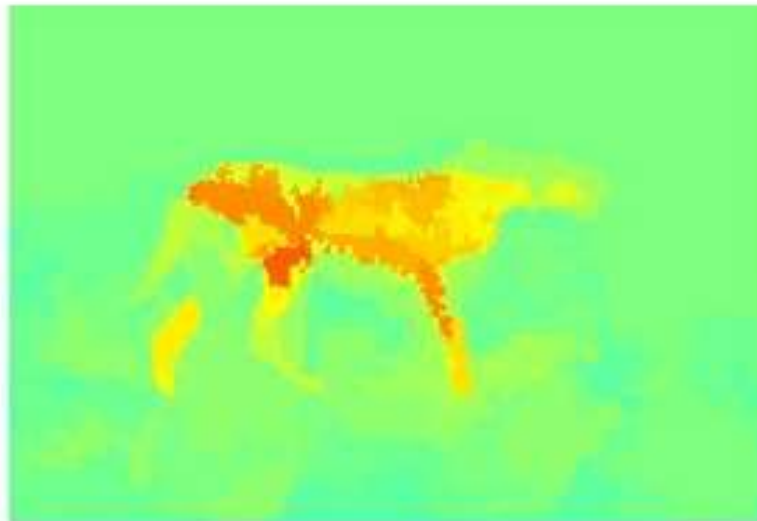
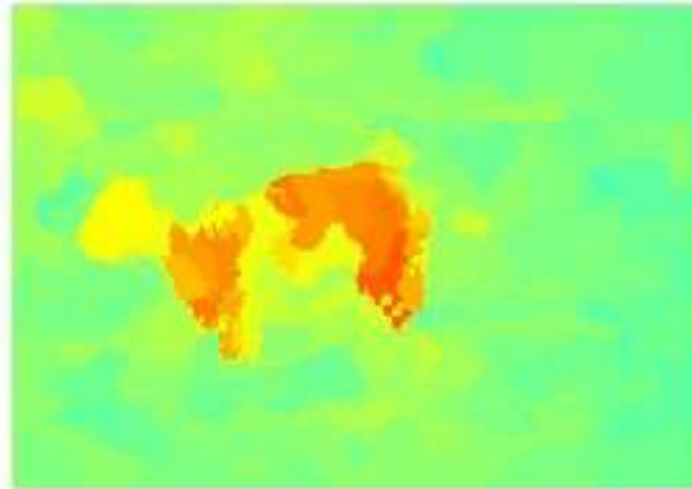
Segmentation par approche «région»

Ce qu'on aimerait



Segmentation par approche «région»

Ce qu'on a



Segmentation par approche région

- Exemple
- L'objet pingouin ne peut être représenté par une seule région d'images, car les parties blanches et noires ne peuvent être regroupées par un algorithme de segmentation



Segmentation par approche «contour»

- Un contour est une frontière entre deux milieux différents (2 couleurs, 2 niveaux de gris...).
- On utilise les méthodes de détection de contours, exemple:
 - Les méthodes de mise en évidence des variations de niveau (par produit de convolution):
 - Opérateur Gradient
 - Opérateur Laplacien
 - Les méthodes d'extraction de contours, exemple l'algorithme de Canny

Gradient et Laplacien

- Les détails correspondent généralement à des changements de niveau de gris = les contours de l'image
- Ils peuvent être mis en relief par la dérivée première (Gradient) ou bien par la dérivée seconde (Laplacien)
 - Pas de changement \rightarrow dérivée première est nulle
 - Transition \rightarrow dérivée première passe par un maximum et dérivée seconde passe par zéro

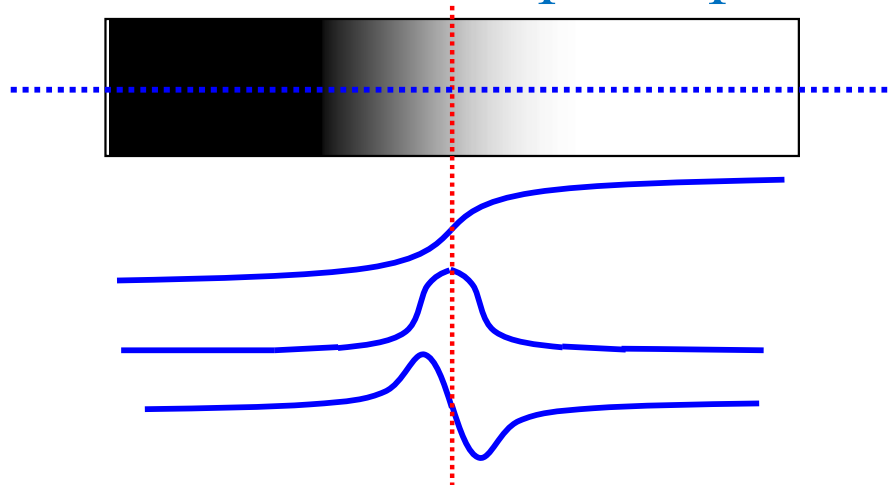


Image avec 2 régions

Profil

Dérivée première

Dérivée seconde

Gradient

- Le calcul des dérivées directionnelles en x et en y revient finalement à la convolution avec un noyau du Gradient, exemple:

Masques de Sobel

- $f_x(i,j) = (f * h_x)(i,j)$
- $f_y(i,j) = (f * h_y)(i,j)$

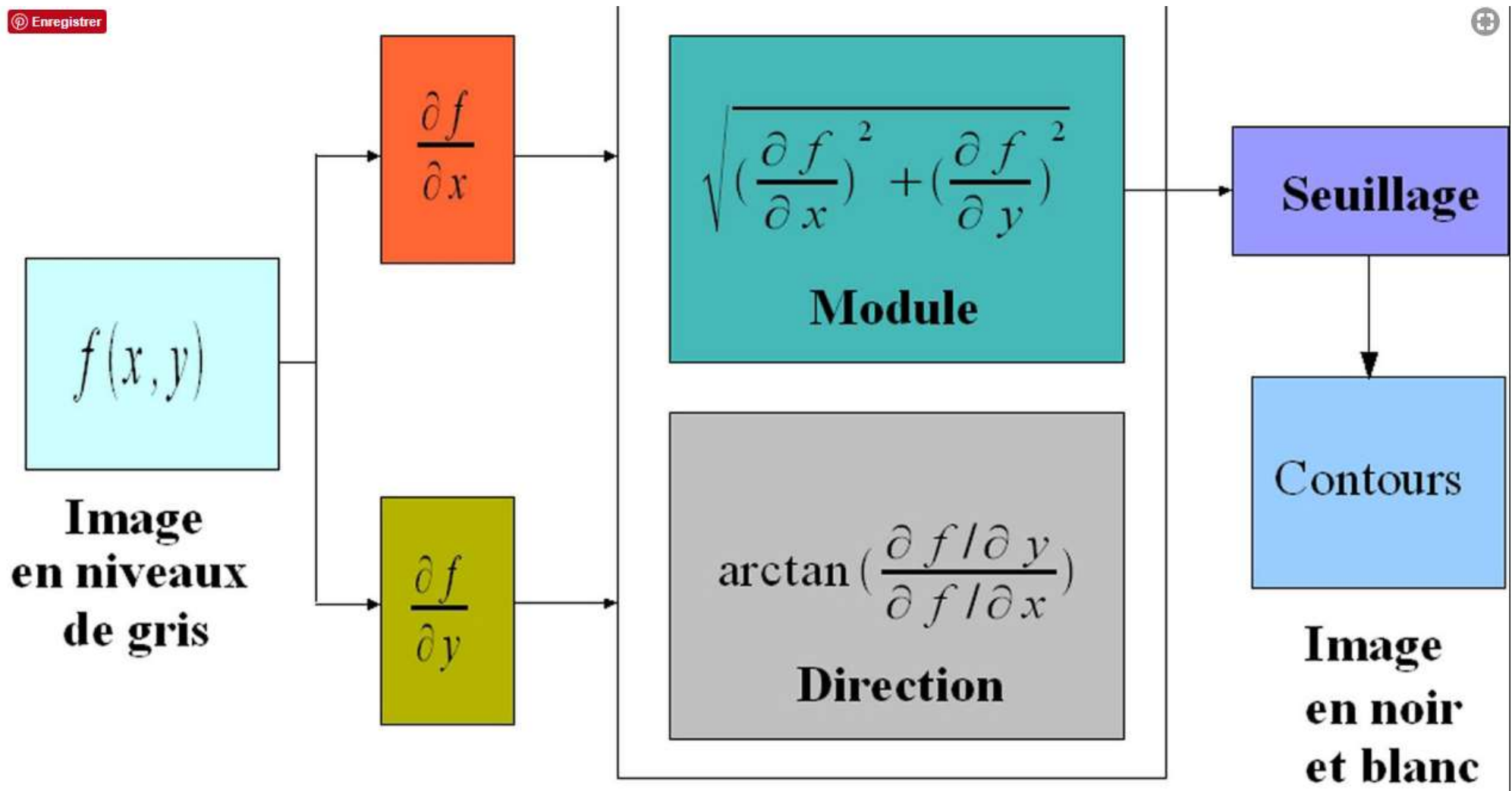
-1	0	1
-2	0	2
-1	0	1

Noyau h_x

-1	-2	-1
0	0	0
1	2	1

Noyau h_y

Détection des contours avec le Gradient



Gradient

- Exemple

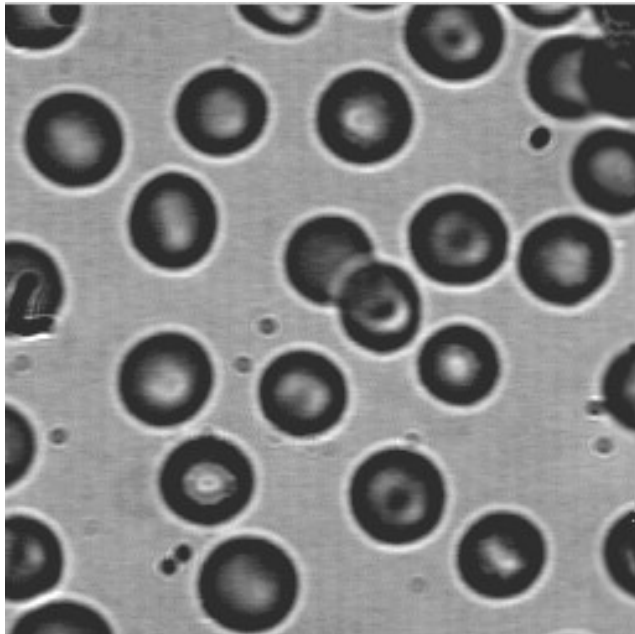


Image originale

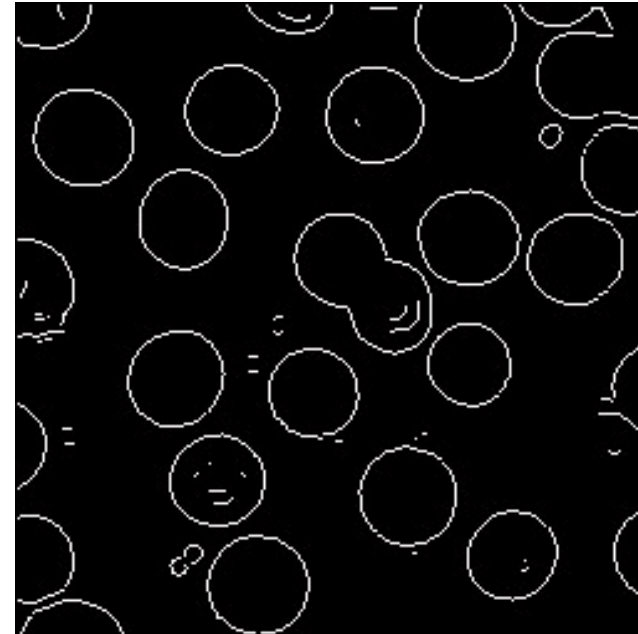


Image des contours
(**opérateur Sobel + seuillage**)

Exercice 5

- Utiliser les masques de Sobel pour déterminer les gradients dans les directions x (img1) et y (img2) de l'image I
- Remarque: mettre 0 pour les bords de l'image

- Calculer l'image module selon la formule:

$$IMM(i, j) = |img1| + |img2|$$

-1	0	1
-2	0	2
-1	0	1

Noyau hx

-1	-2	-1
0	0	0
1	2	1

Noyau hy

0	0	1	0	0
0	1	2	1	0
0	1	2	0	0
0	1	0	0	0
0	0	0	0	0

Image I

Corrigé de l'exercice 5

0	0	1	0	0
0	1	2	1	0
0	1	2	0	0
0	1	0	0	0
0	0	0	0	0

Image I

2	8	6	8	2
4	10	4	8	2
4	8	8	10	1
4	6	8	4	0
2	2	2	0	0

Image IMM

-1	0	1
-2	0	2
-1	0	1

Noyau hx

1	4	0	-4	-1
3	7	-1	-7	-2
4	6	-3	-6	-1
3	2	-3	-2	0
1	0	-1	0	0

Image Gradient
en x

-1	-2	-1
0	0	0
1	2	1

Noyau hy

1	4	6	4	1
1	3	3	1	0
0	-2	-5	-4	-1
-1	-4	-5	-2	0
-1	-2	-1	0	0

Image Gradient
en y

Segmentation

- La grande question en segmentation:
 - A quel niveau de segmentation doit-on s'arrêter ?
 - par exemple, un portrait sur fond uniforme → 2 régions
 - d'autres images nécessitent plus de régions
- Nombre de régions est différent selon l'image ??
→ Il n'existe pas de segmentation générique ou meilleure que les autres → en fonction de la tâche à accomplir certaines méthodes sont plus efficaces que d'autres.

Points d'intérêts

- Ce sont les points qui seront trouvés similaires dans les images similaires.
- Une manière de les déterminer est de prendre en compte les zones où le signal change, par exemple, les points d'intérêt peuvent être les coins, les jonctions en T ou les points de fortes variations de texture.
- Trois types d'approche pour l'extraction de points d'intérêt:
 - approches contours
 - approches intensité
 - approche à base de modèle

Points d'intérêts

- Approches **contours** :
 - Etape1: extraction des contours d'une image
 - Etape2: les points d'intérêt sont extraits le long des contours en considérant les points de courbures maximales ainsi que les intersections de contours.
- Approches **intensité** :
 - les points de discontinuité d'intensité sont les points d'intérêt
- Approche à **base de modèle** :
 - les points d'intérêt sont identifiés dans l'image par mise en correspondance de la fonction d'intensité avec un modèle théorique

Exercice 6

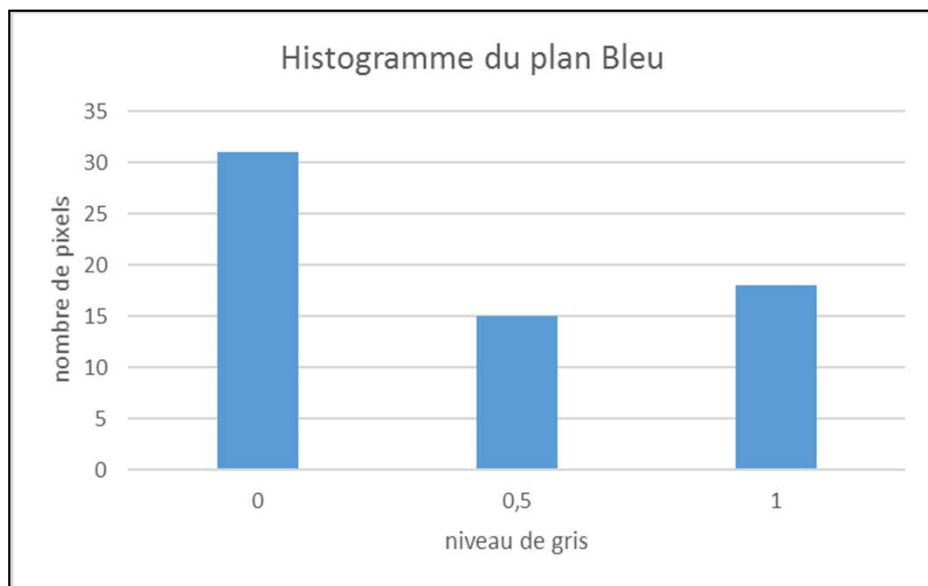
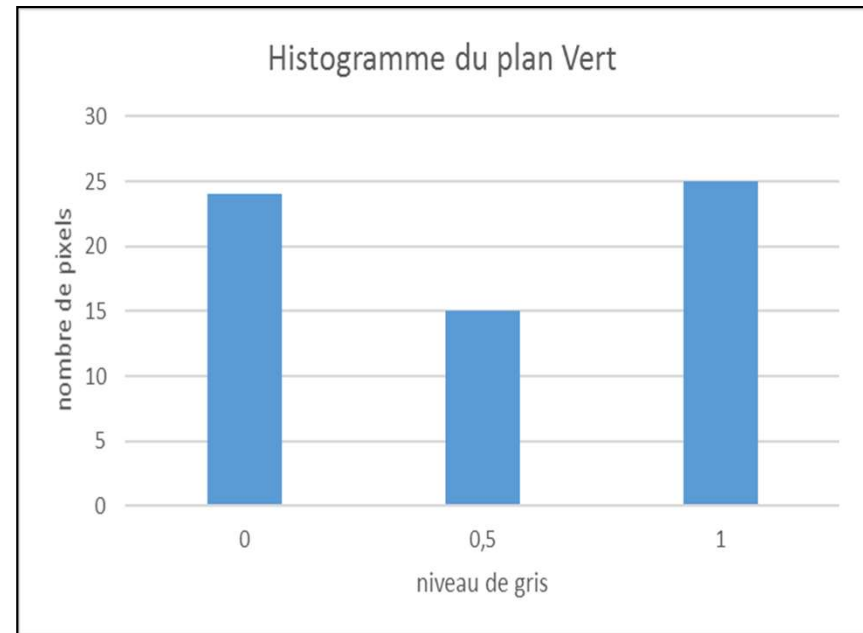
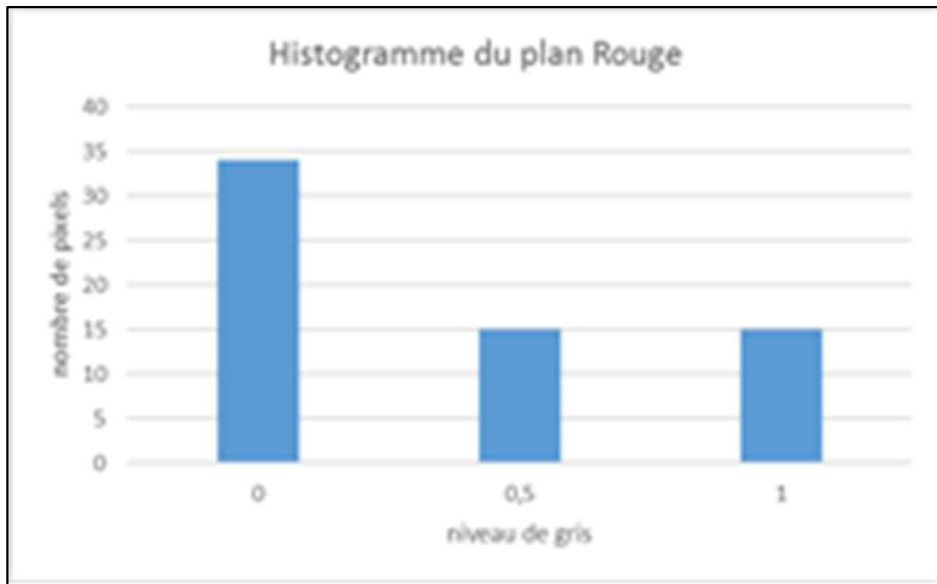
- Soit l'image couleur I définie par une matrice de chiffres et une matrice de couleurs (table t). Chaque chiffre de la première matrice I correspond à une couleur définie par table t.

Image I	Table de couleur
$\begin{bmatrix} 1 & 1 & 2 & 2 & 4 & 5 & 6 & 7 \\ 1 & 1 & 1 & 2 & 4 & 5 & 6 & 7 \\ 1 & 1 & 2 & 2 & 4 & 5 & 6 & 7 \\ 2 & 2 & 2 & 2 & 5 & 5 & 6 & 7 \\ 2 & 2 & 2 & 2 & 4 & 8 & 7 & 8 \\ 4 & 4 & 4 & 4 & 3 & 3 & 7 & 8 \\ 5 & 5 & 5 & 2 & 5 & 5 & 8 & 8 \\ 6 & 6 & 5 & 2 & 6 & 6 & 6 & 7 \end{bmatrix}$	$t = \begin{bmatrix} 0 & 0 & 0 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Exercice 6

1. Tracer l'histogramme l'image I
2. L'histogramme couleur est un des descripteurs les plus utilisés en indexation d'images est, pourquoi ?
3. Quel sont les inconvénients de ce descripteur ? Proposez des solutions.

Corrigé de l'exercice 6



Il faut trois histogrammes

→ un histogramme par couleur

Corrigé de l'exercice 6

2. L'histogramme couleur est un des descripteurs les plus utilisés en indexation d'images, pour différentes raisons:
- Réduit la taille de la base d'indexes et le coût de la recherche
 - Insensible aux changements d'orientation
 - Insensible à la position des régions
 - Insensible à la taille de l'image, dans le cas où on utilise des histogrammes normalisés

Corrigé de l'exercice 6

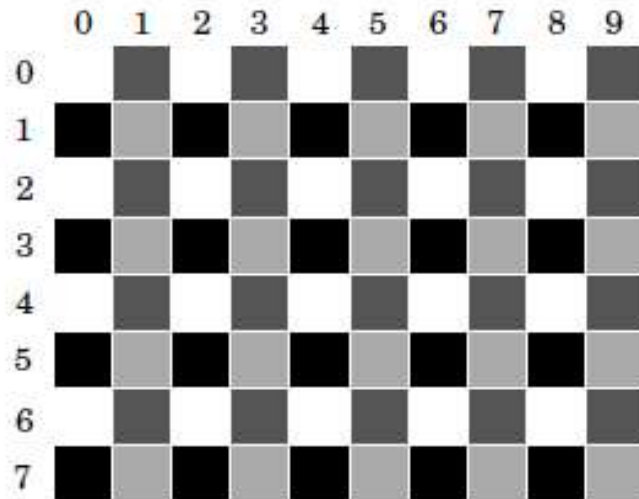
3. Inconvénients de ce descripteur:

- Il ne peut pas distinguer des surfaces couleurs de taille équivalentes mais de répartitions différentes.
- Il ne capture pas les relations spatiales entre les régions → Il ne permet pas une recherche efficace d'objets dans l'image
- Indice visuel insuffisant si utilisé seul → non discriminant, exemple : pomme rouge vs Ferrari (même nombre de pixels rouge dans les deux images)

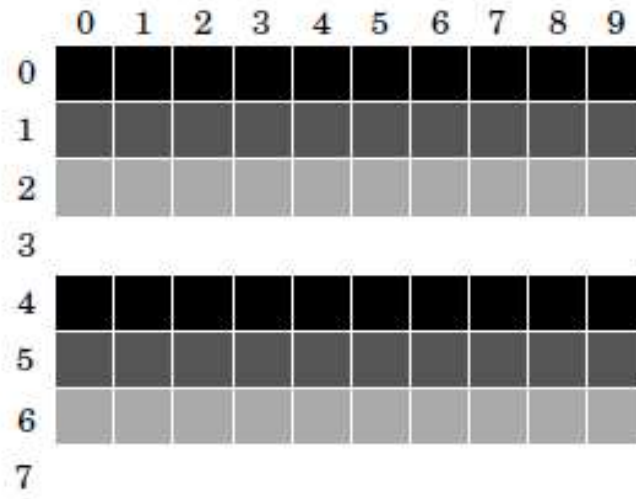
→ Solution : on peut combiner avec un autre descripteur, exemple : la texture

Exercice 7

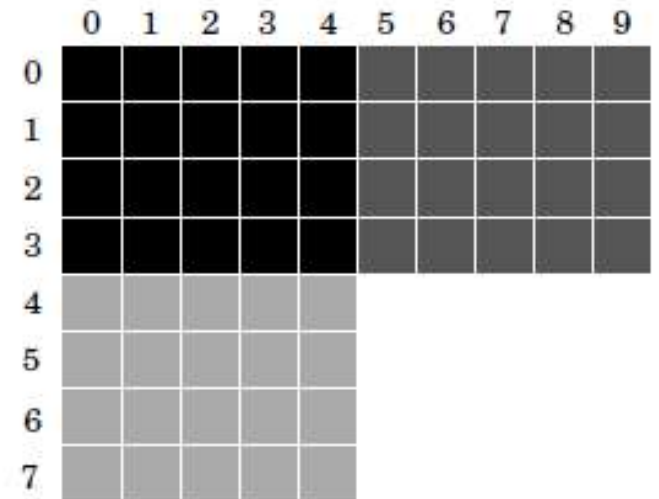
Considérant les trois mages ci-dessous dont le niveau de gris de chaque pixel peut prendre quatre valeurs distinctes : **0 pour le noir**, **1 pour le gris foncé**, **2 pour le gris clair** et **3 pour le blanc**. On cherche à caractériser les textures contenues dans ces trois images.



texture 1



texture 2



texture 3

Textures dans des images à 4 niveaux de gris

Exercice 7

1. Calculer les histogrammes de ces trois images et expliquer pourquoi ce descripteur ne peut pas être exploité pour décrire une texture.
2. On va utiliser comme descripteur la texture pour indexer ces images.
 - a) Calculer la matrice de cooccurrence pour la distance **d=1** et la direction horizontale **angle = 0°**
 - b) Est-ce que cette matrice de cooccurrence permet de différencier les différentes textures ?
 - c) Calculer maintenant la matrice de cooccurrence pour la distance **d=1** et la direction verticale (**angle = 90°**) pour la **texture 2** et la **texture 3**.
 - d) Est-ce que cette matrice permet de différencier la **texture 2** de la **texture 3**

Exercice 7

- e) Normaliser les matrices de cooccurrence afin que **la somme de leurs éléments soit égale à 1**.
- f) Pour caractériser ces matrices normalisées, on va utiliser l'indice qu'**Haralick** appelé **inertie** ou encore **contraste**, défini par l'expression ci-dessous. Calculer cet indice pour les matrices **normalisées** déterminées précédemment

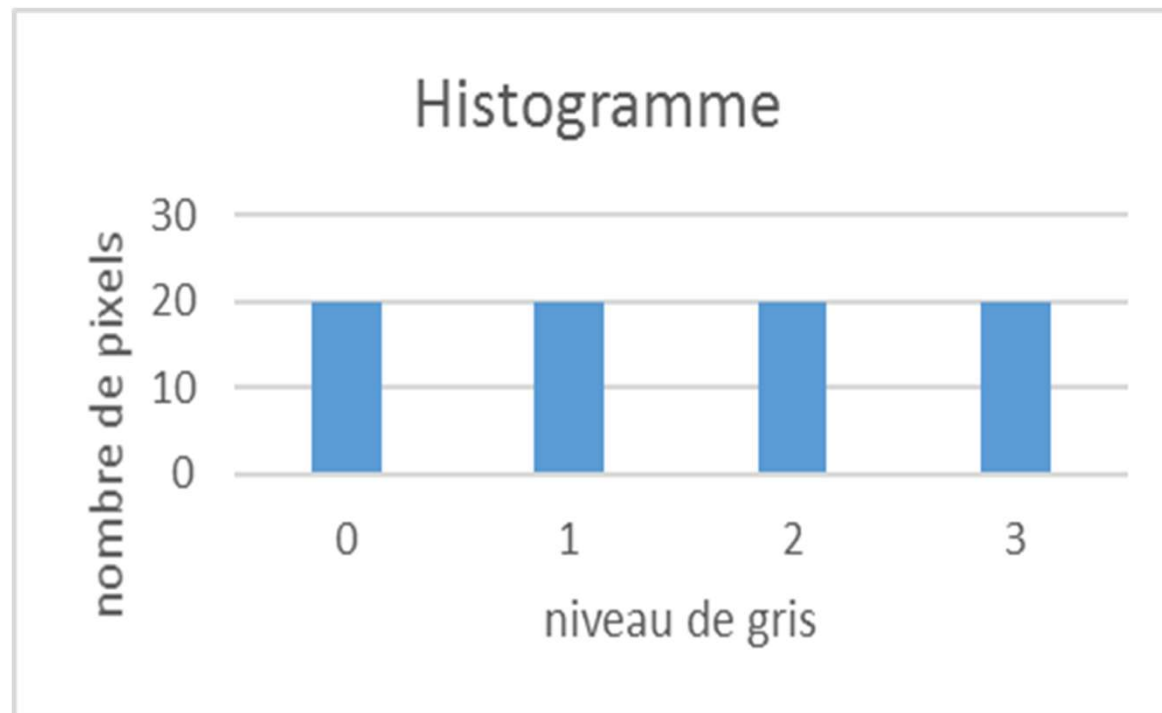
$$\sum_{u,v} (u - v)^2 \cdot M(u, v).$$

Mesure de Haralick (contraste)

avec (u,v) les
niveaux de gris
et M la matrice
de
cooccurrence

Corrigé exercice 7

- L'histogramme des trois images est identique → l'histogramme dans ce cas ne peut pas être utilisé comme descripteur car il ne permet pas de différencier les trois images (les textures des trois images)
- Il faut utiliser la texture comme descripteur: exemple la matrice de cooccurrence qui permet de caractériser la texture dans les images.



Corrigé exercice 7

Matrices de cooccurrence pour la distance $d=1$ et la direction horizontale $\text{angle} = 0^\circ$

	0	1	2	3
0	0	0	20/ 72	0
1	0	0	0	16/ 72
2	16/ 72	0	0	0
3	0	20/ 72	0	0

indice de contraste=4

Pour la texture 1

	0	1	2	3
0	18/ 72	0	0	0
1	0	18/ 72	0	0
2	0	0	18/ 72	0
3	0	0	0	18/ 72

indice de contraste=0

Pour la texture 2

	0	1	2	3
0	16/ 72	4/ 72	0	0
1	0	16/ 72	0	0
2	0	0	16/ 72	4/ 72
3	0	0	0	16/ 72

indice de contraste=0,1

Pour la texture 3

Corrigé exercice 7

- La matrice de cooccurrence pour la distance $d=1$ et la direction horizontale $\text{angle} = 0^\circ$ permet de différencier la texture 1 des deux autres textures (texture 2 et texture 3). Par contre, elle n'est pas suffisante pour différencier la texture 2 de la texture 3.
- ⇒ Solution : pour la texture 2 et la texture 3, calculer les matrices de cooccurrence pour la distance $d=1$ et la direction verticale $\text{angle} = 90^\circ$

Corrigé exercice 7

Matrices de cooccurrence pour la distance $d=1$ et la direction verticale $\text{angle} = 90^\circ$

	0	1	2	3
0	0	20/ 70	0	0
1	0	0	20/ 70	0
2	0	0	0	20/ 70
3	10/ 70	0	0	0

Indice de contraste= $150/70 = 2.14$

Pour la texture 2

	0	1	2	3
0	15/ 70	0	5/ 70	0
1	0	15/ 70	0	5/ 70
2	0	0	15/ 70	0
3	0	0	0	15/ 70

Indice de contraste= $40/70 = 0.57$

Pour la texture 3

⇒ La matrice de cooccurrence pour la distance $d=1$ et la direction verticale $\text{angle} = 90^\circ$ permet de différencier la texture 2 de la texture 3