

A thick black L-shaped frame is positioned on the left and right sides of the slide, framing the central text.

# BASES DE DONNÉES MULTIMÉDIA

Partie 2

# CONTENU ET MÉTA-DONNÉES

# Contenu et méta-données

Les données “brutes” (fichier image, fichier son) provenant d’objets multimédia contiennent des **informations sémantiques**.

On en distingue

1. La donnée brute (image, signal sonore)
2. Les métadonnées décrivant le contenu

## DEFINITION:

Le processus d’abstraction permet de décrire le contenu d’objets multimédia par des métadonnées, soit manuellement, soit par des procédés (semi-)automatiques

# Métadonnées / Méta information

- En BD relationnelles classiques

- ✓ métadonnées = schéma de la BD = informations structurées sur la structure des données = dictionnaire de données

- En MM

- ✓ métadonnées = données structurées sur les éléments individuels de la BD
- ✓ permettent de pallier, en partie, la difficulté d'associer un sens aux données MM
- ✓ peuvent être acquises automatiquement (difficile et risque d'erreur) ou générées manuellement (coût élevé)
- ✓ peuvent aussi impliquer une description technique (et non sémantique) des données MM : date de la prise de vue, paramètres techniques de la prise de vue
- ✓ peuvent concerner la structure de données MM complexes

# 3 classes de métadonnées

Classe	Exemple	Utilisation
Indépendant du contenu	Nom du photographe	ABR
Décrit le contenu	Le nom des personnes	TBR
Dépend du contenu	Forme d'un papillon	CBR

# Exemples de métadonnées

Média	Indépendant du contenu	Décrit le contenu	Dépend du contenu
Texte	Date de mise-à-jour	Mots clés, format	Frontière des sujets
Son	Durée	Personne qui parle	Reconnaissance vocale
Image	Titre, date	Format	Caractéristiques (traits)
Vidéo	Distributeur, durée	Angles des prises de vues	Frontières des séquences

# ID3 Tags

- Conçu en 1996 et devenu *de facto* un standard
- 128 octets

Position	Taille	Contenu
0	3	Marqueur 'TAG'
3	30	Titre
33	30	Interprète
63	30	Album
93	4	Année
97	30	Commentaire
127	1	Genre musical

Valeur numérique	Genre
0	Blues
1	Classic rock
2	Country
3	Dance
4	Disco
5	Funk
6	Grunge
7	Hip-Hop
8	Jazz
9	Metal
10	New Age
11	O/dies (rétro)
12	Autre
13	pop

# Exchangeable image file format (Exif)

- Spécification pour les formats d'images des appareils numériques
  - ✓ non géré par une organisation internationale, mais largement utilisé
- Pour JPEG, TIFF, RIFF. Ne supporte pas JPEG2000, PNG ou GIF
- Le format supporte:
  - ✓ Date et heure, enregistrés par l'appareil
  - ✓ Les paramètres de l'appareil
- **Dépendent du modèle** : inclus la marque et des informations diverses telles que le temps d'ouverture, l'orientation, la focale, l'ISO, etc.
  - ✓ Une vignette de pré-visualisation
  - ✓ La description et les informations de copyright
- Supporté par de nombreuses applications



<u>General</u>		<u>P</u> ermissions	<u>M</u> eta Info	<u>P</u> review
<b>JPEG Exif</b>				
Comment:	<div></div>			
Creation Date:	05-01-14			
Creation Time:	12:38:36 am			
Dimensions:	2560 x 1920 pixels			
Exposure Time:	0.100 (1/10)			
JPEG Quality:	Unknown			
Aperture:	f/3.3			
Color Mode:	Color			
Date/Time:	05-01-14 12:38:36 am			
Flash Used:	Off			
Focal Length:	6.3 mm			
ISO Equiv.:	100			
JPEG Process:	Baseline			
Camera Manufacturer:	PENTAX Corporation			
Metering Mode:	Pattern			
Camera Model:	PENTAX Optio WP			
Orientation:	1			

OKCancel

# Annotations

- Une annotation textuelle sera toujours trop restrictive
- Malgré cela, l'approche la plus utilisée reste l'annotation textuelle et manuelle

**Avantage :** recherche indépendante du type de media

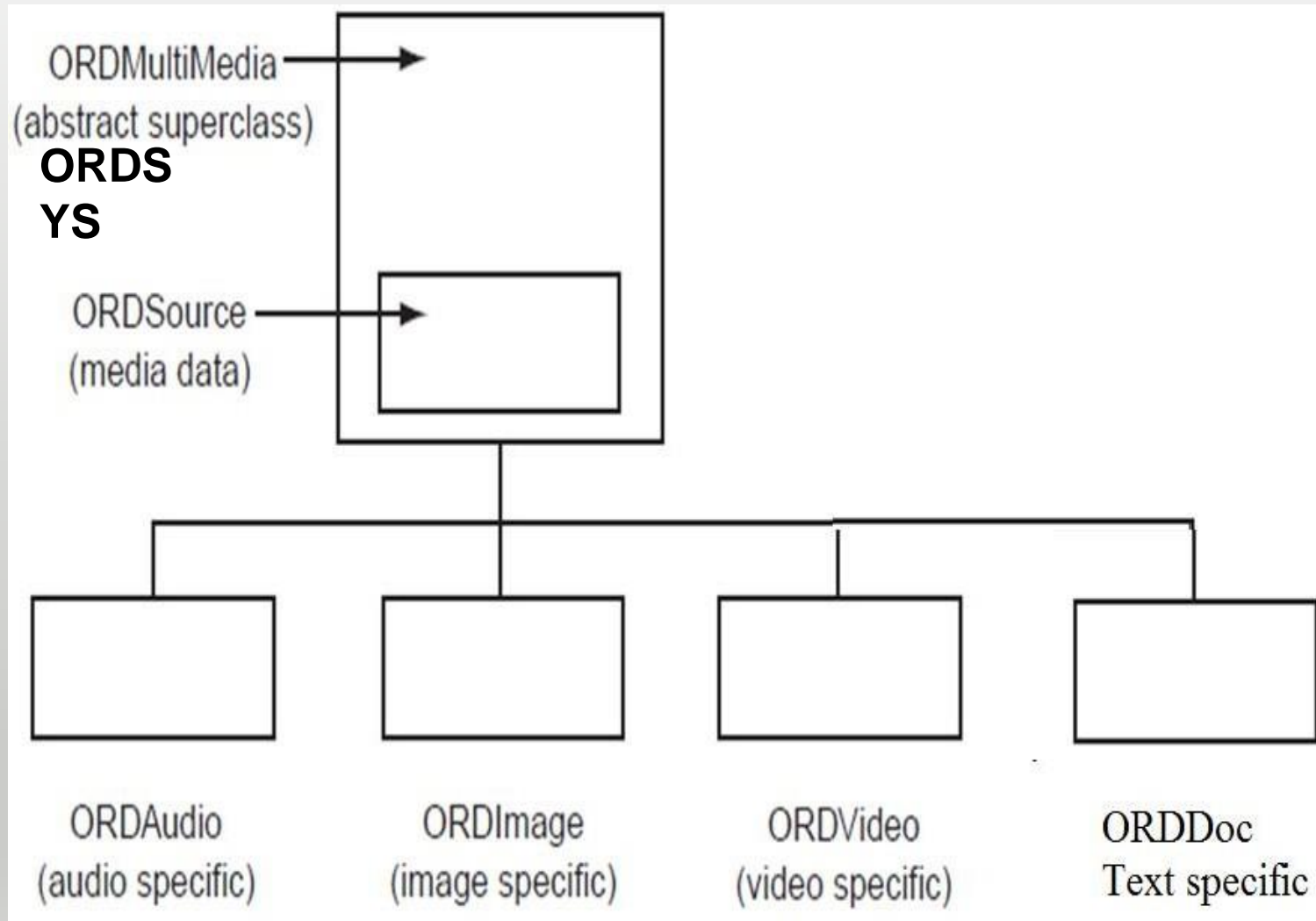
**Inconvénients :**

1. Le coût d'annotation manuel est très important
2. Différentes personnes utilisent un vocabulaire différent pour signifier la même chose (ex: clair, lumineux).  
L'annotation est subjective.
3. Le **non-verbal** ne peut être exprimé sans ambiguïté.

# intermedia

- Oracle *InterMedia* est une extension d'Oracle 10g pour stocker, gérer, manipuler et rechercher des images, des fichiers audio, vidéo ainsi que d'autres sources de données hétérogènes.
- Oracle fourni des **classes** (relationnel-objet) pour stocker les documents multimédia : ORDImage, ORDAudio, ORDVideo et ORDDoc.
- Ces quatre types stockent les sources de données dans une classe d'**objet** appelée ORDSOURCE.

# intermedia



# example

```
CREATE TABLE album
```

```
(
```

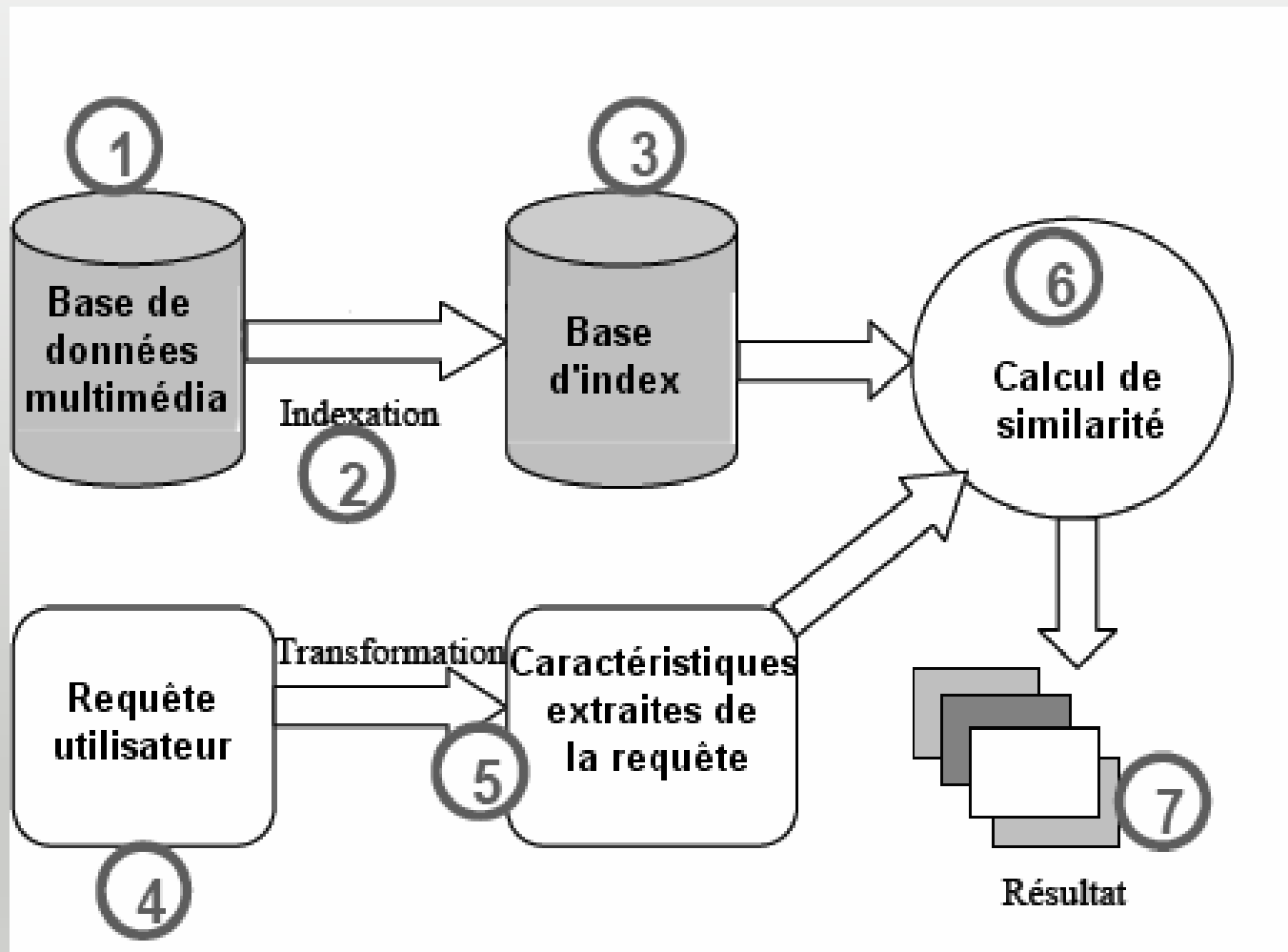
```
photo_id INTEGER, photographer VARCHAR2(64),
```

```
annotation VARCHAR2(255), PIC ORDSYS.ORDImage
```

```
);
```

```
SELECT t. photo_id id, t.PIC.getHeight() height, t.PIC.getWidth()  
width, t.PIC.getMimeType() mimetype, FROM album t;
```

# Systeme de recherche d'informations multimedia SRIM



# Indexation

- Structurer la base pour ne pas avoir à la parcourir entièrement à chaque recherche, comme dans un dictionnaire
- Nécessite de décrire le contenu avec des index portant sur les médias
- Difficulté : les médias en question sont représentés par des vecteurs de grande dimension

# Niveaux d'interrogation de BDMM

- **Niveau1**: extrait automatique des **caractéristiques physiques** tel que couleur, forme, texture, localisation spatiale et mouvement des objets
- Exemple de La requête :
  - ✓ « **Trouver des objets qui possèdent du violet** »
  - ✓ « **Trouver des images avec une région bleu ciel dans le haut de l'image** »



# Niveaux d'interrogation de BDMM

- **Niveau 2** : extraction automatique des caractéristiques **logiques** qui sont reliées à l'identité d'un objet dans le média
- Exemple de requête:
  - ✓ « **Trouver une vidéo d'un avion en approche** »
  - ✓ « **Trouver une image de baleine bleue** »

# Niveaux d'interrogation de BDMM

- **Niveau3** :extraction automatique des attributs abstraits associés avec la **compréhension (sens)** de la nature et l'objectif de l'objet dans le média
- Exemple de requête:
  - ✓ « **Trouver une photo d'un acte terroriste** »

# Approche d'interrogation MM

- Approche actuelle de BD = **ABR** (Attribute Based Retrieval)
- Approche d'annotation avec du texte = **TBR** (Text Based Retrieval)
- Approche d'interrogation avec le contenu = **CBR** (Content Based Retrieval)

# Interrogation des attributs (ABR)

- Utilise un ensemble d'attributs (champs d'une table) structurés comme dans les SGBD traditionnels
- Doit pouvoir traiter des objets de grande taille, tel que des vidéos (CLOB, BLOB)
- Efficace pour des données textuelles
- N'utilise pas la richesse du contenu des images
- Requêtes limitées au niveau conceptuel 1
- Exemple:  
"Afficher les image des voitures de la **marque "Renault"**"

# Interrogation du texte (TBR)

- Il faut ajouter des annotations (courtes descriptions dans des attributs structurées)
- Utilise aussi des champs de l'approche ABR
- Difficile à réaliser en pratique :
- Annotations générées et saisies manuellement, donc
  - dispendieuses.
- Annotations faites par des professionnels mais les requêtes formulées par des utilisateurs.
- Formulation dépend de l'interprétation de l'utilisateur
  - Exemple:
    - ✓ Afficher les images du : "match Chelsea vs Manchester United"
    - ✓ Description des vidéos sur Youtube !!!

# Interrogation du contenu(CBR)

- Alternative au TBR
- L'extraction de plusieurs éléments d'information peut être automatisée
- Repérage plus intuitif car utilise des éléments en relation naturelle avec les objets: texture, couleur, forme (attributs physiques du média)
- Exemple
  - ✓ Afficher les images des "voitures rouge"
    - Forme + couleur

Nature de l'interrogation	Techniques employées
Attribut (ABR)	Techniques et structures de données typiques du SGBD et de classification des attributs
Texte (TBR)	Techniques d'indexation (et classification) et des extensions SQL (contains, context, score, sim, ...) pour les BD textuelles
Contenu (CBR)	Techniques d'extraction des traits automatiques et de classification de ses traits en regroupements logiques

# Modes de recherche

- Linguistique - Linguistique (LL)
  - ✓ Utilise les langages de recherche des bases de données  
ex. SQL
  - ✓ Recherches dans les métadonnées
- Visuelle - Visuelle (VV)
  - ✓ CBR (Content based retrieval) – utilise les
    - caractéristiques physiques de l'image pour retrouver des semblables
- Visuelle - Linguistique (VL)
  - ✓ Présente une image et fait la requête avec ses métadonnées
- Linguistique - Visuelle (LV)
  - ✓ Utilise un index des attributs visuels

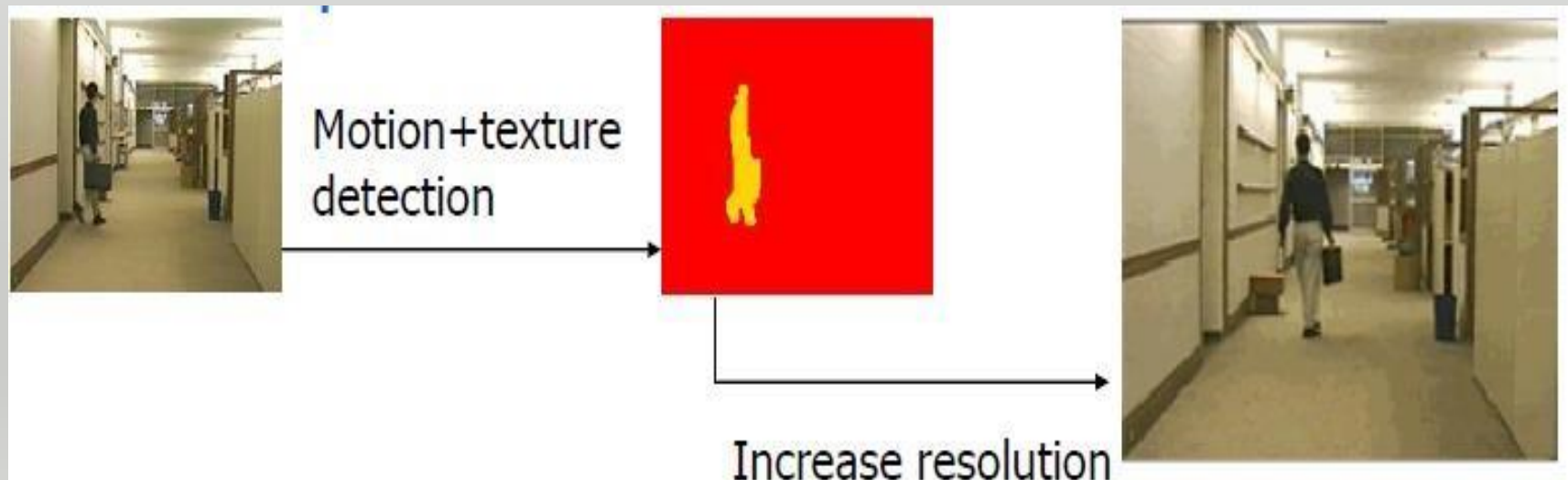


# Exemples d'applications multimédias

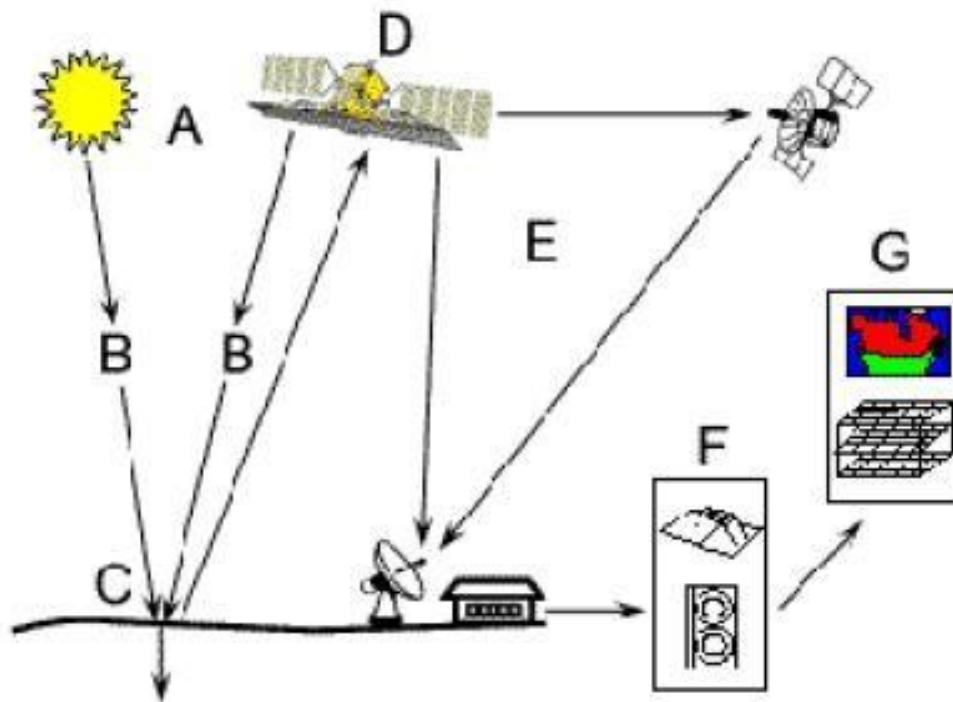
- Entertainment systems – Video on Demand, Audio on Demand, Digital Photo albums
- Public Protection
- Medical Information Systems
- Remote Sensing (Télédétection)
- News/Films à la demande
- Commerce électronique
- Informations médicales
- Systèmes d'informations géographiques
- Architecture/Design
- Protection du copyright
- Géo-localisation
- Enquêtes policières
- Militaire
- Expérimentation scientifiques
- Enseignement
- Archivage, gestion des bases de données de contenu (personnelles ou professionnelles)
- Moteur de recherche (ex: Internet)
- Applications industrielles

# Context-Aware Multimedia Adaptation

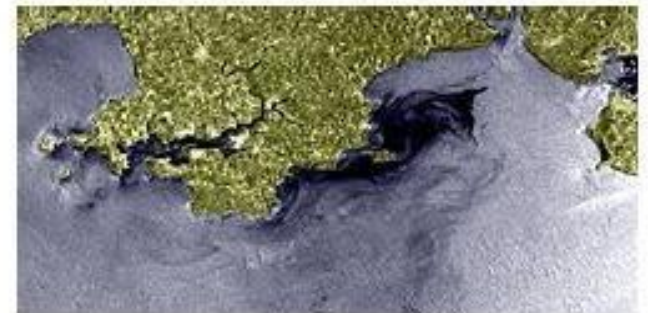
- Écoute le contenu et ses changements
- S'adapte quand il y a un "évènement" détecté
- Example:



# Télédétection, Remote Sensing (Gouv. Canada)



© CCRS / CCT



# Agriculture Canada

- Classification des produits agricoles
- Évaluation de la condition des récoltes
- Caractéristiques de terrain
- Application des normes de production agricole



# Quelques systèmes commerciaux

- QBIC (Query By Image Content) :
  - Sans être exactement le premier système de recherche/indexation d'images, QBIC marque en 1995 le début des systèmes commerciaux. Il est aujourd'hui l'un des systèmes les plus cités et décrits de la littérature.
- Les requêtes proposées par QBIC sont :
  - Requête par l'exemple,
  - Manipulation directe des caractéristiques de bas niveau,
  - Requête par croquis.



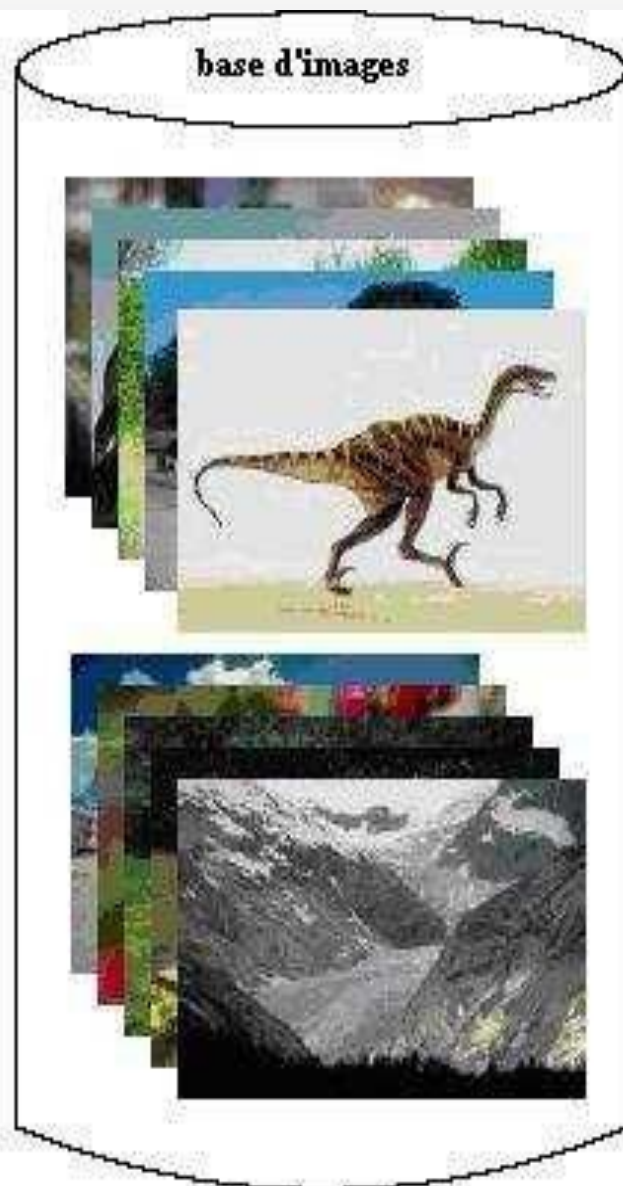
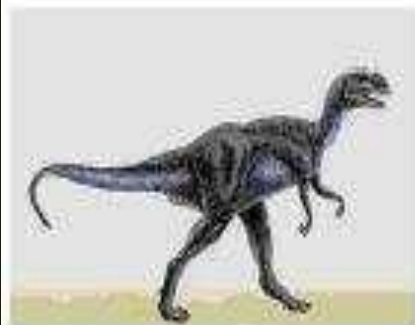
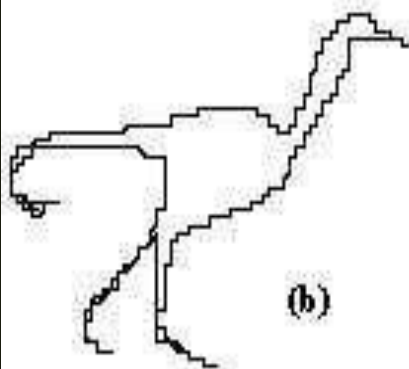
requête

mot clés:  
'dinosaur'

(a)

base d'images

résultat



# VisualSEEK :

- C'est un système de recherche d'image développé par l'université de Columbia (New York).
- Son système de requête combine entre les **relations spatiales et le bouclage de pertinence**.
- **Nota : La pertinence c'est l'adéquation entre le contenu effectif des documents et l'information recherchée par un utilisateur. Les images pertinentes sont obtenues via des mesures de similarité. Le jugement de pertinence de l'utilisateur n'est celui du système c'est pourquoi il y a bouclage.**

Browse images

Add images

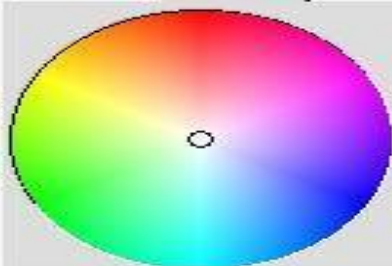
Search image

Options

Local database

Import

Draw



Brush size



8

Color



Results

10



Search!



54.54%



54.32%



54.26%



53.92%



53.51%



53.34%



53.31%



53.25%



53.22%



52.77%