LIRIS

CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE UMR 5205

Traitement et Analyse des données visuelles et sonores

****

Méthodes avancées en image et vidéo

Ecole Centrale de Lyon

Laboratoire d'InfoRmatique en Image et Systèmes d'information

LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon
Université Claude Bernard Lyon 1, bâtiment Nautibus
43, boulevard du 11 novembre 1918 — F-69622 Villeurbanne cedex
http://liris.cnrs.fr

INSA LYON

Université Claude Bernard Lyon 1

UNIVERSITÉ LUMIÈRE LYON 2

CENTRALE LYON

Speakers @ ECL

- **Mohsen Ardabilian**
  - mohsen.ardabilian@ec-lyon.fr
- **Emmanuel Dellandrea**
  - emmanuel.dellandrea@ec-lyon.fr

Summary

- Content Based Image Retrieval
- Image Processing: Super-Resolution
- Acquisition/Modelling: Range Face Acquisition
- Image Analysis: applied to Face Recognition
- Audio as a Modality

CBIR - content-based image retrieval

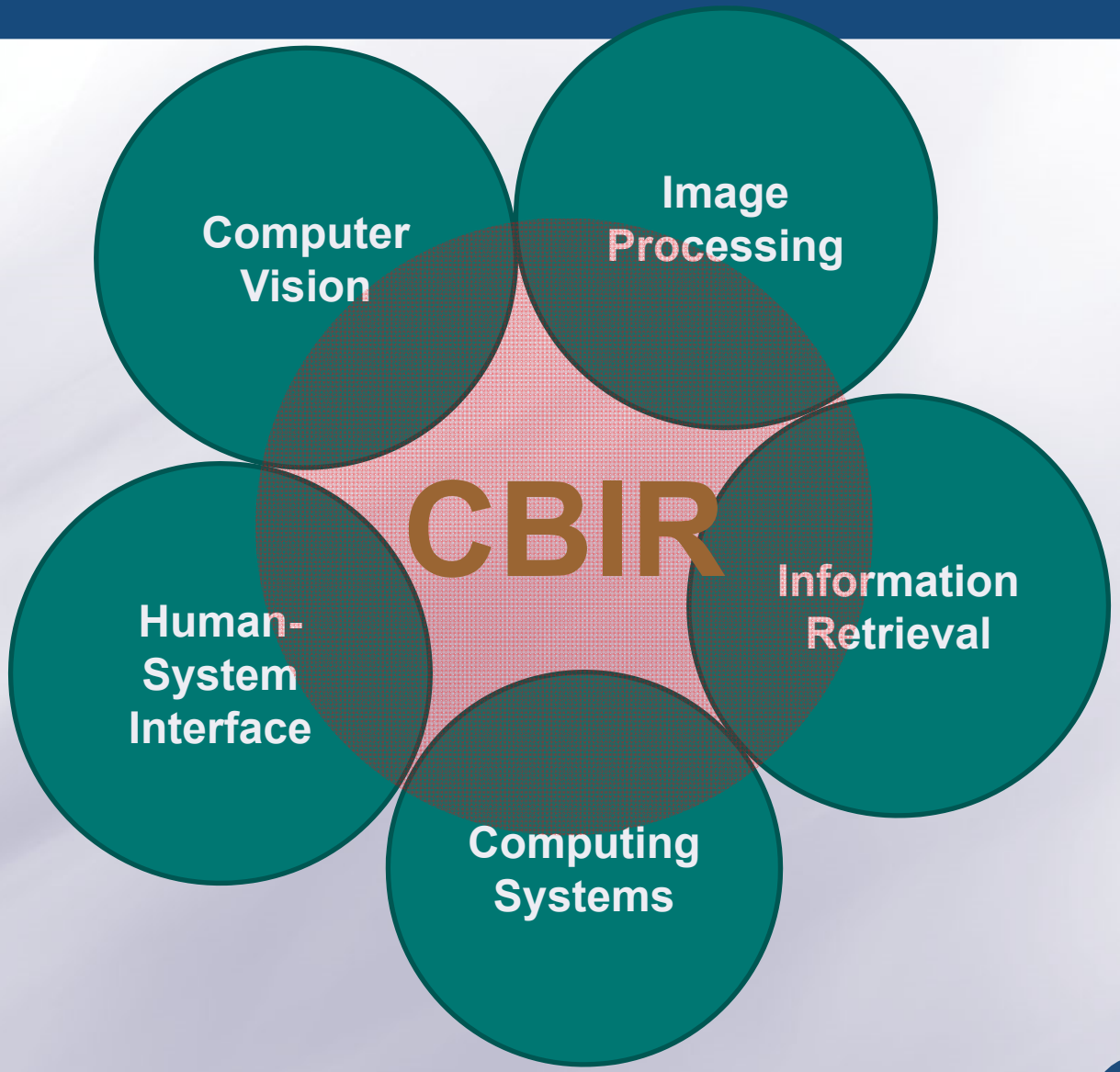
- [Kato 92] experiments into automatic retrieval of images from a database by colour and shape feature

- And later to describe the process of retrieving desired images from a large collection on the basis of

  - Features (primitive or semantic such as colour, texture and shape)
  - Automatically extracted from the images themselves
  - Retrieval of images by manually-assigned keywords is definitely not CBIR as the term is generally understood – even if the keywords describe image content

CBIR - content-based image retrieval

- Image databases are essentially unstructured: digitized images consist purely of arrays of pixel intensities, with no inherent meaning
- Need to extract useful information from the raw data before any kind of reasoning about the image's contents is possible
  - such as recognizing the presence of particular shapes or textures
- Raw material, words stored as ASCII character strings, has already been logically structured by the author [Santini and Jain, 1997]
- There is no equivalent of level 1, basic level (pixels), retrieval in a text database
- Big Data


Computer Vision
Image Processing
CBIR
Human-System Interface
Information Retrieval
Computing Systems

CBIR vs Images Processing & Computer Vision

- **CBIR draws its methods from image processing, computer vision to achieve image retrieval in big size data**
- **Image processing covers also enhancement, compression, transmission, and interpretation**
- **Grey areas such as object recognition by feature analysis**
- **Usually fairly clear-cut distinction between image analysis and CBIR**

Police forces use automatic face recognition systems in one of two ways: Firstly, the image in front of the camera may be compared with a single individual's database record to verify his or her identity. In this case, only two images are matched, a process few observers would call CBIR. Secondly, the entire database may be searched to find the most closely matching images. This is a genuine example of CBIR.

Image Retrieval

- Research and development issues in CBIR shared with image processing and information retrieval:
  - Understanding image users' needs and information-seeking behaviour
  - Identification of suitable ways of describing image content
  - Extracting such features from raw images
  - Providing compact storage for large image databases
  - Matching query and stored images in a way that reflects human similarity judgements
  - Efficiently accessing stored images by content
  - Providing usable human interfaces to CBIR systems

Video Retrieval

- Key research issues in video retrieval include:
  - automatic shot and scene detection
  - ways of combining video, text and sound for retrieval
  - effective presentation of search output for the user

CBIR - Content Based Image Retrieval

- Feature Extraction
  - Global Features
  - Local Features
- Search and Retrieval
  - Similarity measures
  - Region-based approaches
  - Classification / Clustering
- Annotation
  - Classification Models
  - Joint models
- Relevance Feedback and Learning
- Multimodal Fusion and Retrieval
- Hardware Support
- Interface and visualization

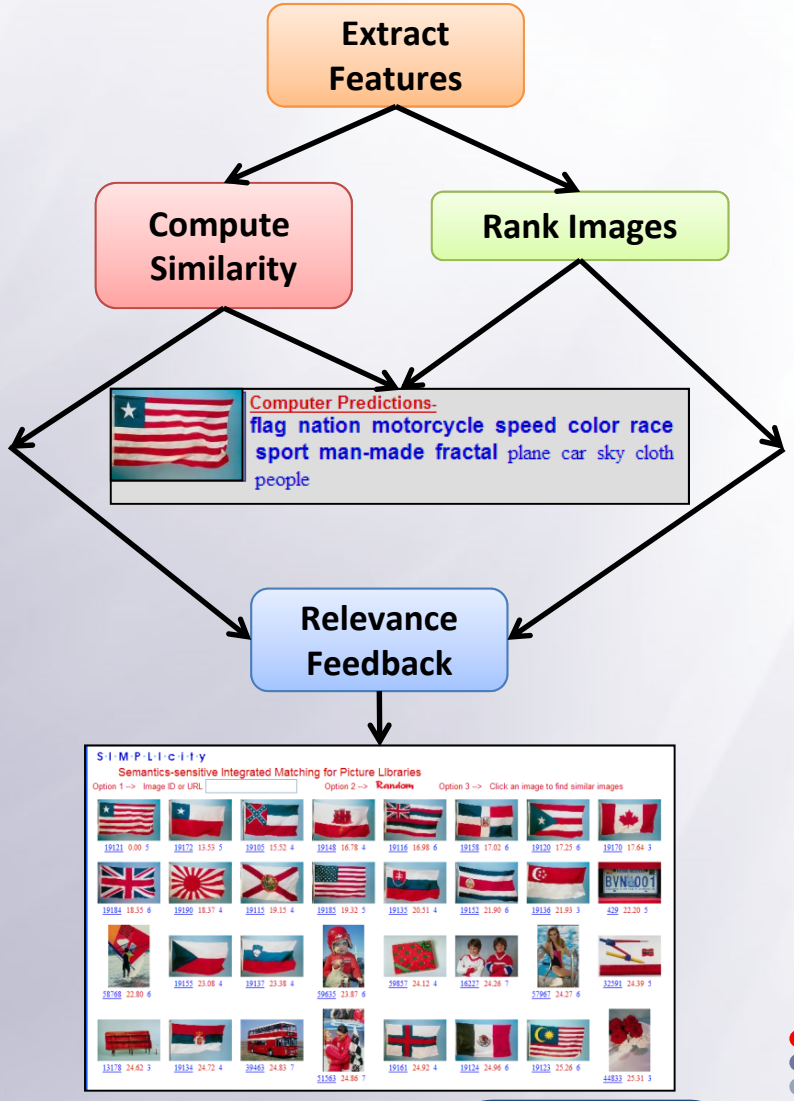
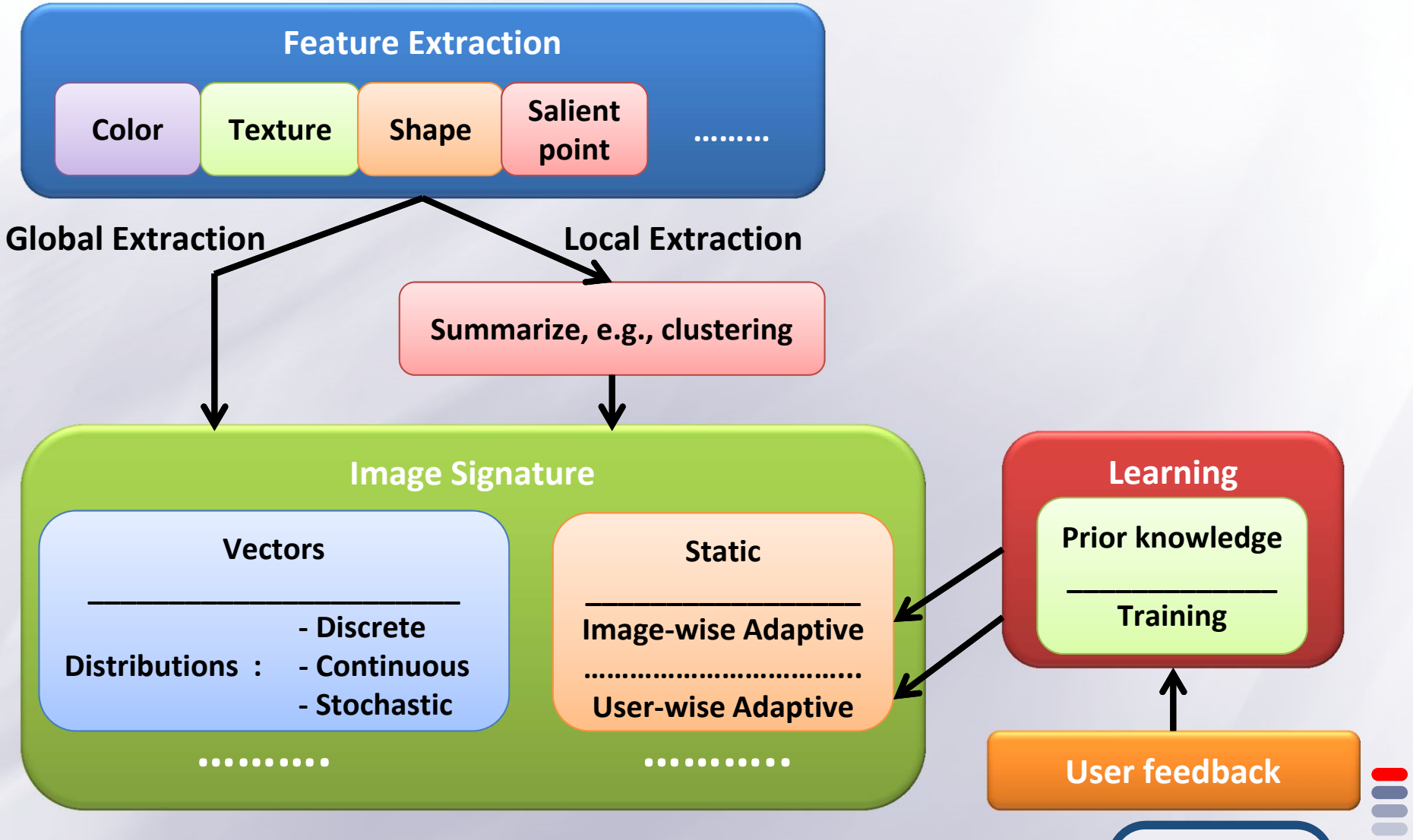


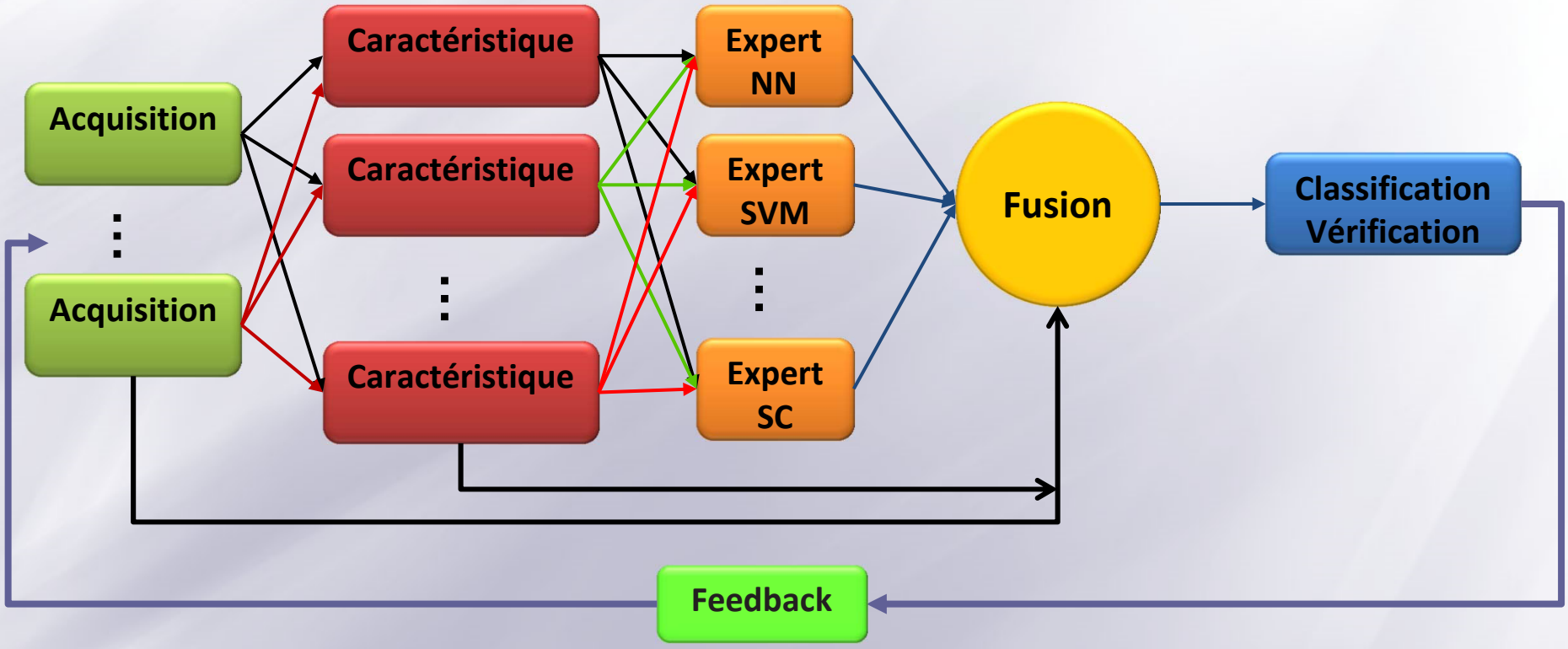
Image signature formulation



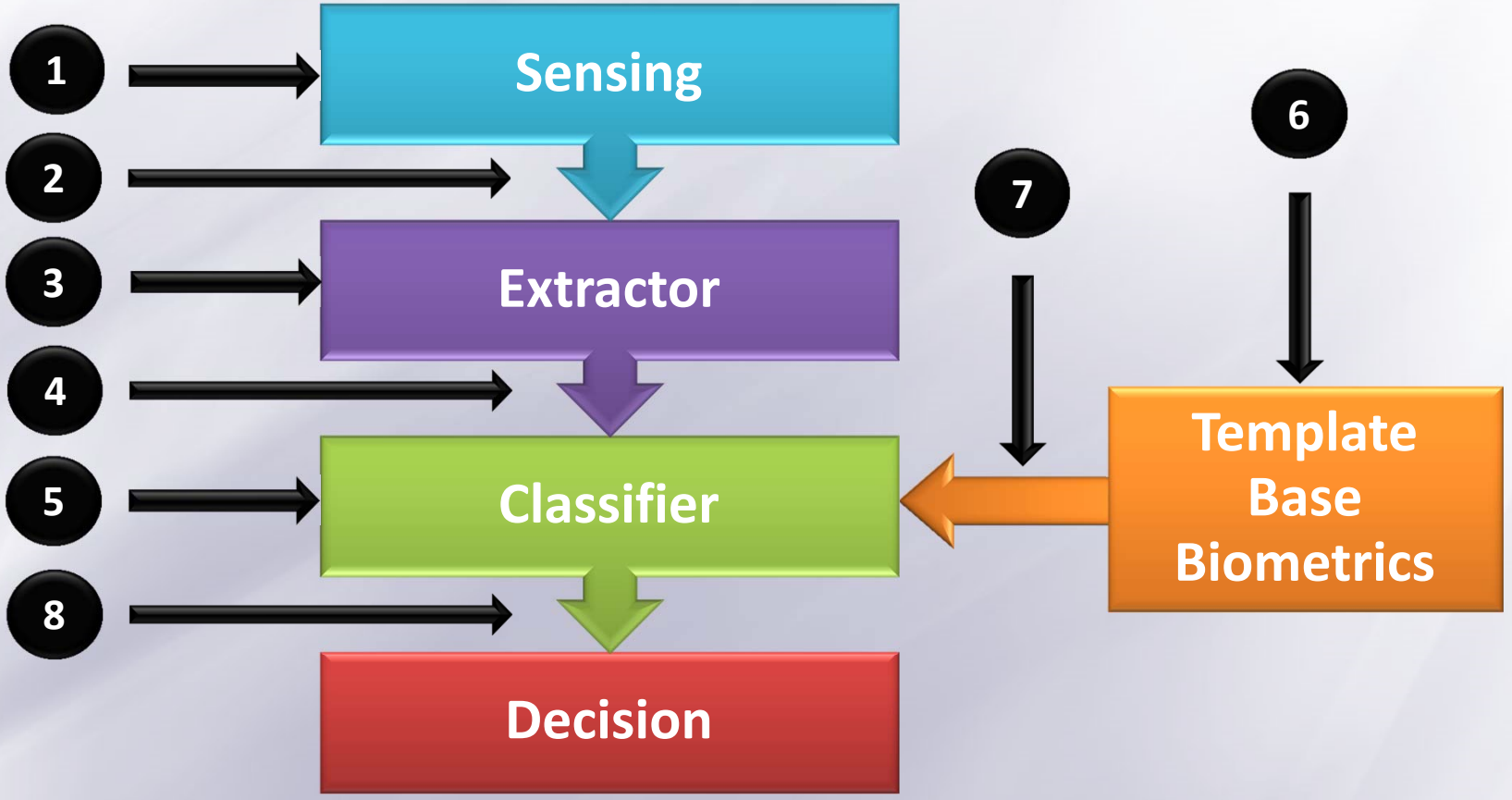
Query types

- Three levels of increasing complexity:
  - *Level 1:* Retrieval by *primitive* features such as colour, texture, shape or the spatial location of image elements
  - *Level 2:* Retrieval by *derived* features (sometimes known as *logical*), involving some degree of logical inference about the identity of the objects depicted in the image
    - Retrieval of objects of a given type (e.g. “find pictures of a double-decker bus”)
    - Retrieval of individual objects or persons (“find a picture of the Eiffel tower”)
  - *Level 3*: Retrieval by abstract attributes, involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted
    - Retrieval of named events or types of activity (e.g. “find pictures of Scottish folk dancing”)
    - Retrieval of pictures with emotional or religious significance (“find a picture depicting suffering”)

Classification et reconnaissance : Paradigme et architecture générale

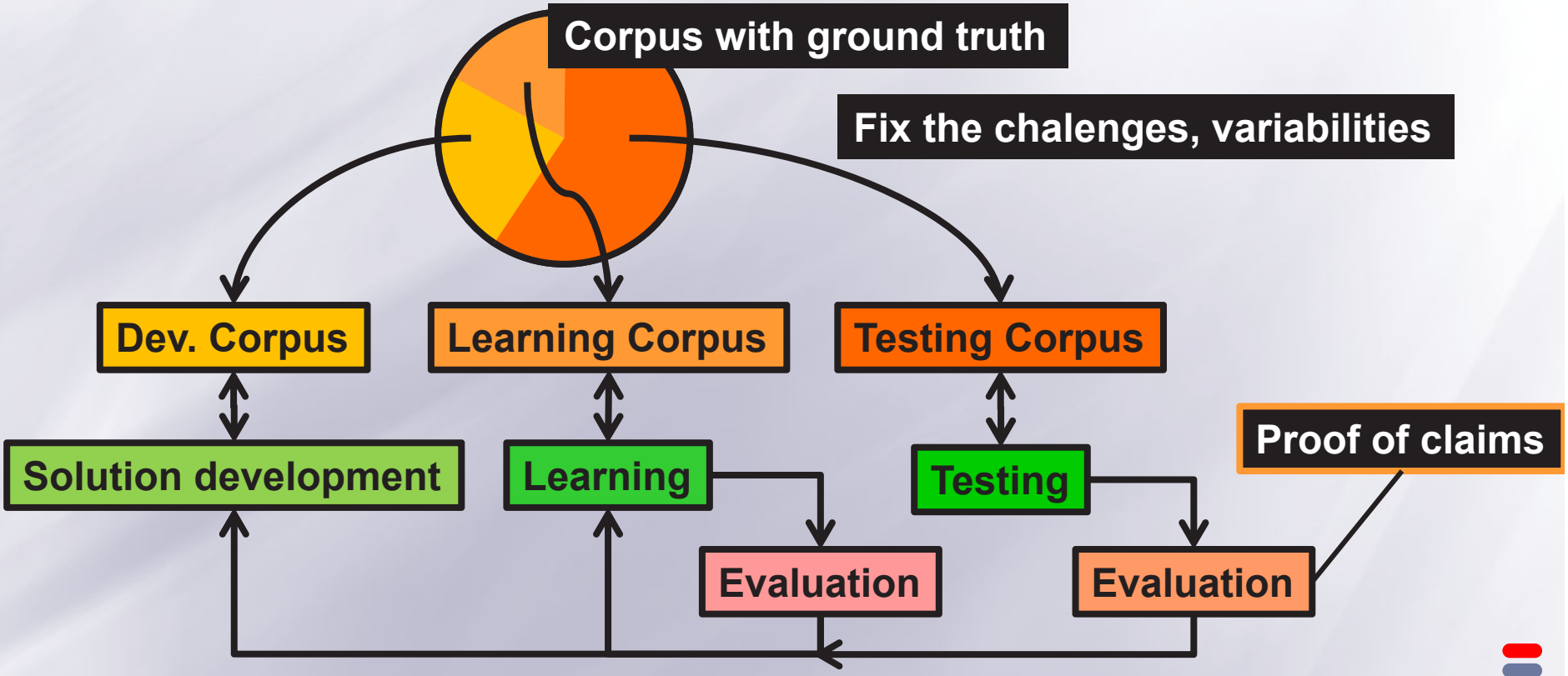


Biometrics Systems as a specific CBIR



Methodology for innovation

- **Development of novel component:** models, algorithms, features, classifiers, frameworks, etc.



What are the innovations ?

- Modeling of the problem, the solution, and the assessment
  - Emerging problem
    - Modeling
    - Solution
    - Assesment
  - Existing problem
    - Novel modeling
    - Innovative new solution, new and improved solution
    - Assesment
  - Dedicated dataset
  - Assesment and Certification

Evaluation : protocole et facteurs mesurables

Ecole Centrale de Lyon

Laboratoire d'InfoRmatique en Image et Systèmes d'information

LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon
Université Claude Bernard Lyon 1, bâtiment Nautibus
43, boulevard du 11 novembre 1918 — F-69622 Villeurbanne cedex
http://liris.cnrs.fr

Evaluation

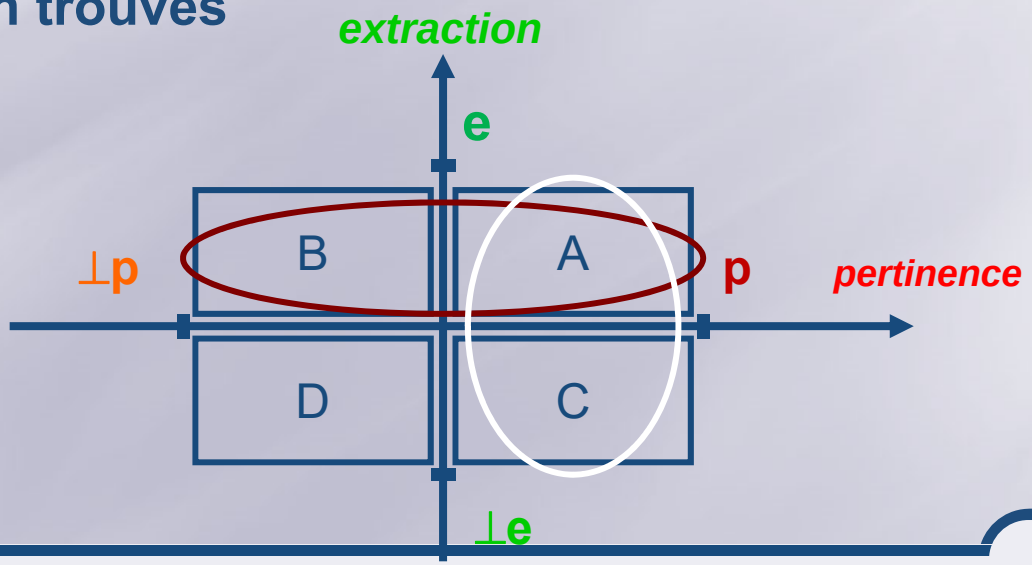
- Corpus de développement, d'apprentissage et de test
  - Evaluer une solution de manière objective
    - Représente le phénomène
    - Enferme des variabilités
    - Des objectifs précis sont visés : précision, robustesse, rapidité
  - Comparer des solutions de manière objective
- Critères d'évaluation
  - Mesure d'efficacité
    - Taux de rappel, de précision, de bruit et de silence, etc.
  - mais aussi, la précision quantitative, les complexités spatiale et temporelle

Comment mesurer ?

- **Vrais positifs (*True Positive*) => A : Nb éléments pertinents trouvés**
- **Faux positifs (*False Positive*) => C : Nb éléments pertinents non trouvés**
- **Vrais négatifs (*True Negative*) => B : Nb éléments non pertinents trouvés**
- **Faux négatifs (*False Negative*) => D : Nb éléments non pertinents non trouvés**

Méthode :

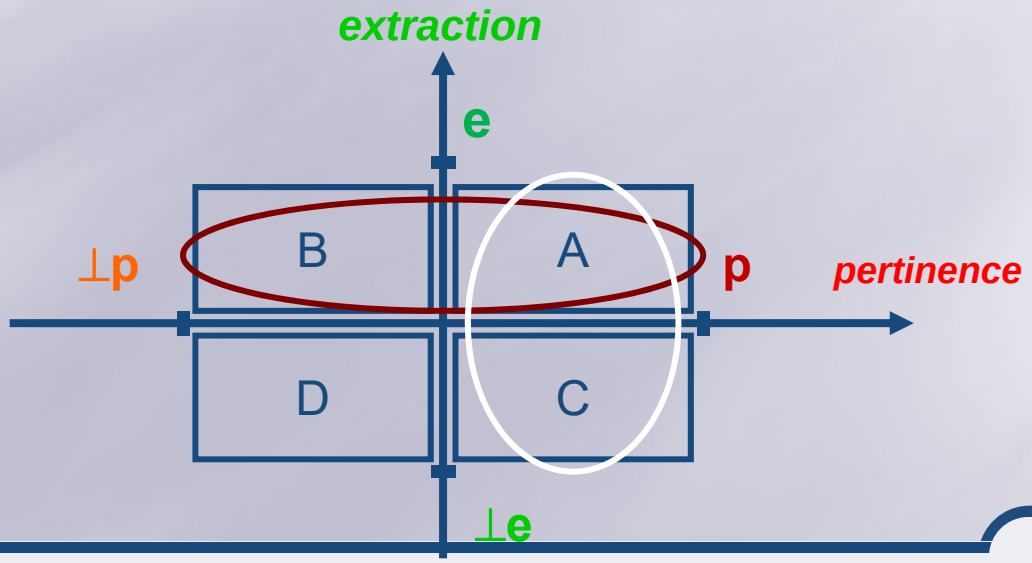
1. ***Définition***
2. ***Visualisation***
3. ***Équation***



Taux de précision (P)

La précision : elle mesure le taux d'éléments pertinents trouvés (classification, reconnaissance, segmentation) par rapport à l'ensemble d'éléments trouvés

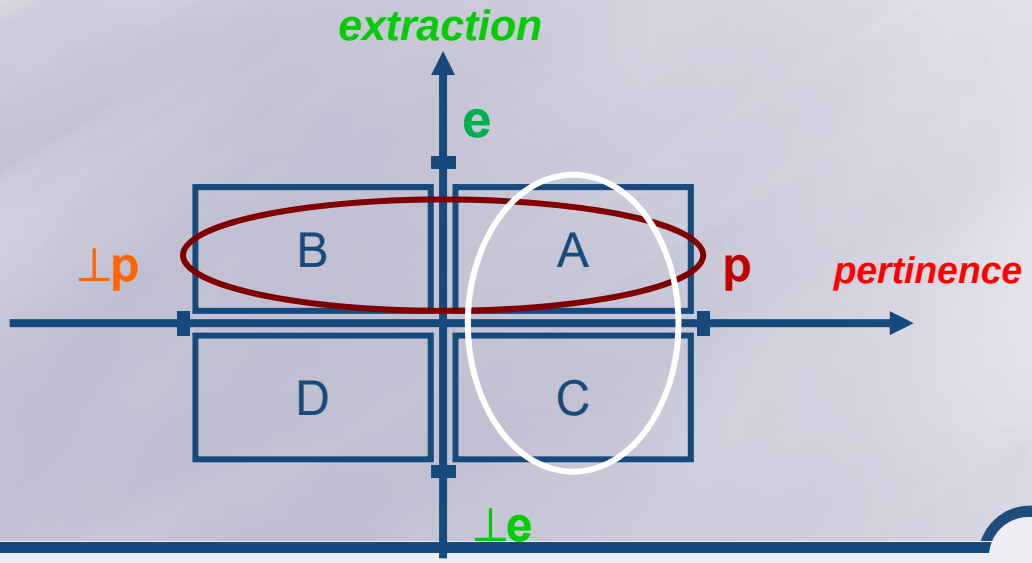
$$Précision = \frac{nombre\ (éléments\ pertinents)}{nombre\ (éléments\ trouvés)} = \frac{A}{A+B}$$



Taux de rappel

Le rappel : elle mesure le taux d'éléments pertinents trouvés (classification, reconnaissance, segmentation) par rapport à l'ensemble d'éléments pertinents existants dans le corpus

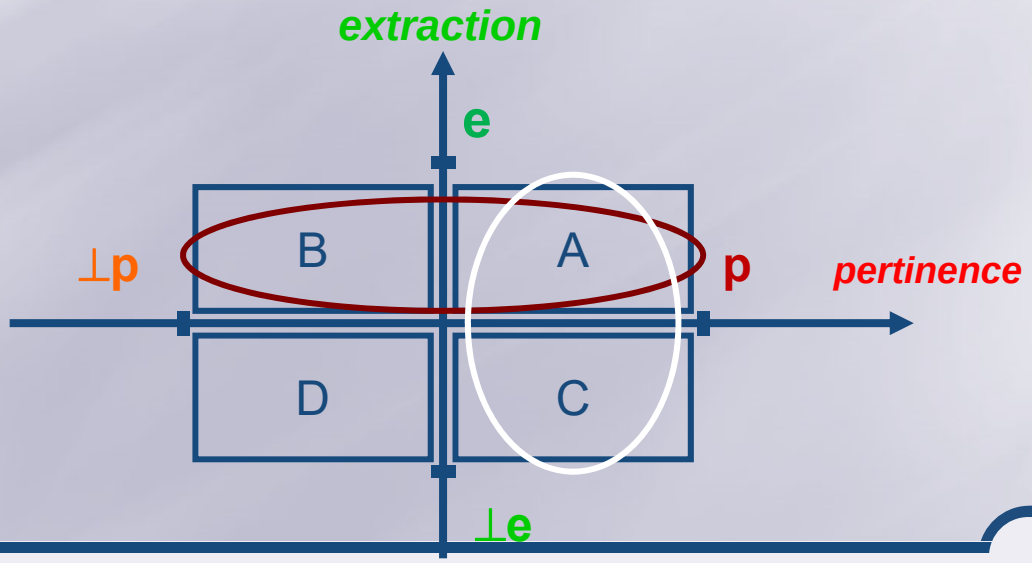
$$Rappel = \frac{nombre\ (éléments\ pertinents\ trouvés)}{nombre\ (éléments\ pertinents\ dans\ le\ corpus)} = \frac{A}{A+C}$$



Bruit

Le bruit : il mesure le taux d'éléments non pertinents trouvés par rapport à l'ensemble d'éléments trouvés

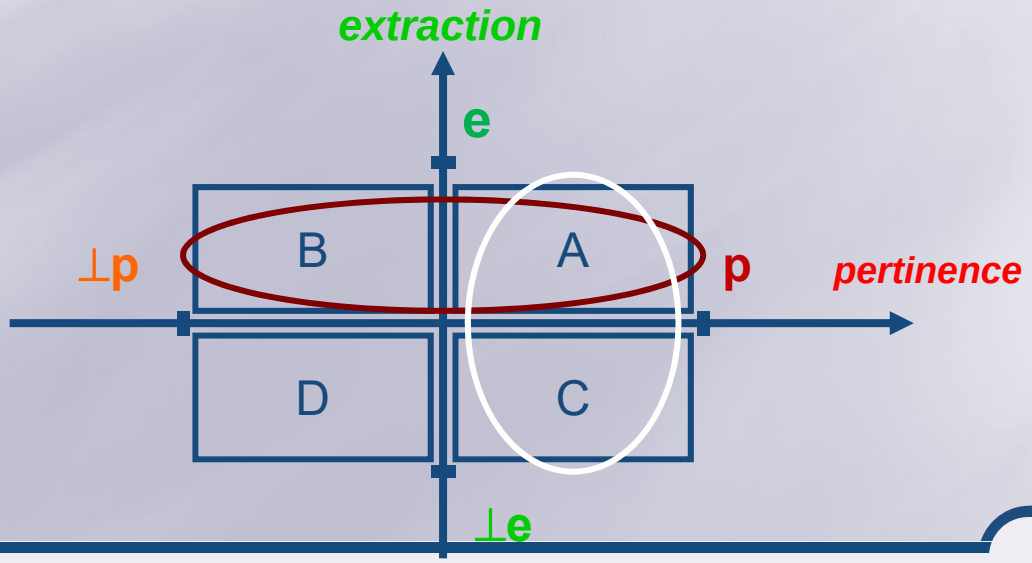
$$Bruit = \frac{nombre\ (éléments\ non\ pertinents\ trouvés)}{nombre\ (éléments\ trouvés)} = \frac{B}{A+B}$$



Silence

Le silence : il mesure le taux d'éléments pertinents non trouvés par rapport à l'ensemble d'éléments pertinents dans le corpus

$$Silence = \frac{nombre\ (éléments\ pertinents\ non\ trouvés)}{nombre\ (éléments\ pertinents\ dans\ le\ corpus)} = \frac{C}{A+C}$$

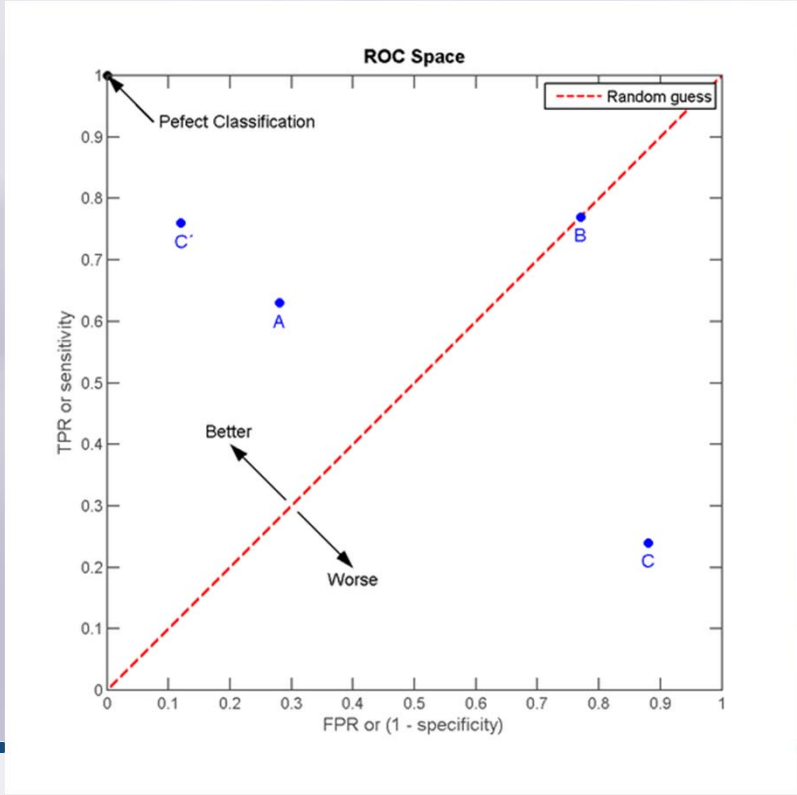


Terminology

- *True Positive (TP)*
  - eqv. with hit
- *True Negative (TN)*
  - eqv. with correct rejection
- *False Positive (FP)*
  - eqv. with false alarm, Type I error
- *False Negative (FN)*
  - eqv. with miss, Type II error
- *Sensitivity* or *True Positive Rate (TPR)*
  - eqv. with hit rate, Recall
  - TPR=TP/P=TP/(TP+FN)
- *False Positive Rate (FPR)*
  - eqv. With Fall-Out
  - FPR=FP/N=FP/(FP+TN)
- *Accuracy (ACC)*
  - ACC=(TP+TN)/(P+N)
- *Specificity (SPC) or True Negative Rate*
  - SPC=TN/N=TN/(FP+TN)=1-FPR
- *Positive Predictive Value (PPV)*
  - eqv. With Precision
  - PPV=TP/(TP+FP)
- *Negative Predictive Value (NPV)*
  - NPV=TN/(TN+FN)
- *False Discovery Rate (FDR)*
  - FDR=FP/(FP+TP)
- *ROC*
- *DET*
- *Rank*

Courbe ROC - Receiver Operating Characteristic

- Le graphe des performances d'un système quand son seuil de discrimination varie. Obtenu en traçant ***TPR = True Prositif Rate*** en fonction de ***FPR = False Positive Rate*** pour différentes valeurs de seuil.



Courbe DET - Detection Error Tradeoff

- Le graphe d'erreur d'un système. ***FNR =False Negative Rate*** VS. ***FPR = False Positive Rate***.