

Cours Vision Artificielle

Master 2 Electronique des Systèmes Embarqués

Faculté de Technologie



Chap : 01

Introduction à la vision Artificielle

L'objectif de cours

Ce cours a pour objectif la :

- ➔ Découverte de techniques parmi les plus récentes en **traitement d'images** et en *vision artificielle*.
- ➔ Compréhension théorique des **principaux problèmes du traitement d'image** et **de la vision artificielle**.
- ➔ Connaissance pratique des outils principaux pour résoudre ces problèmes.

Plan du Cours

Chapitre 1 : Introduction à la vision artificielle

Traitement d'Images et Vision.

Les problèmes propres au Traitement d'Images.

Echantillonnage et Quantification.

Chapitre 2 : Techniques de base de traitements de l'image

Notion d'histogramme et de contraste

Correction de la dynamique de l'image

Egalisation d'histogramme et correction gamma

Opérations logiques et arithmétiques sur les images

Chapitre 3 : La restauration d'images

Convolution 2D.

Filtrage Spatial.

Transformée de Fourier 2D.

Filtrage Fréquentiel.

Chapitre 4 : Détection des contours et Segmentation

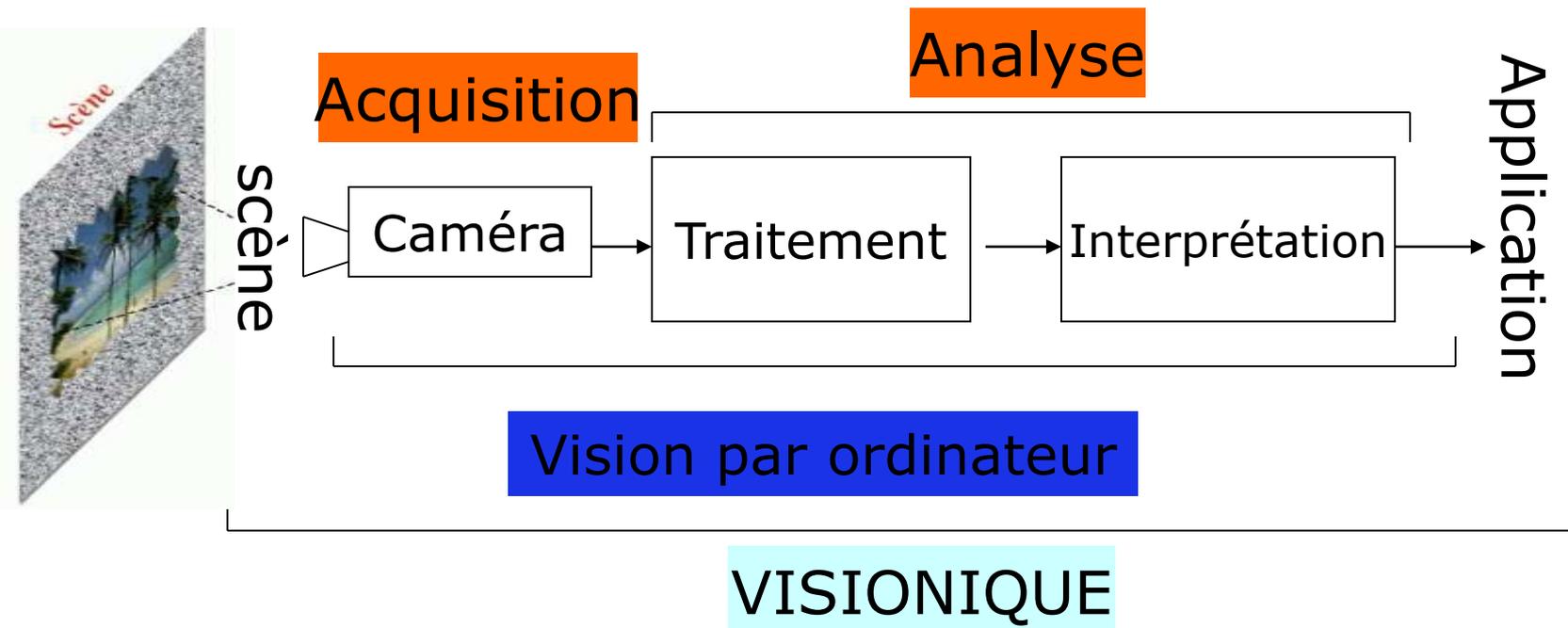
Approche contours.

Approche régions.

Méthodes de segmentation et de classification.

Constitution d'un système de vision

Un système de vision artificielle : une machine associant un capteur d'image et un système de traitement qui transforme et interprète cette image pour fournir une information pertinente sur la scène observée.



Vision artificielle (Computer Vision)?

□ Qu'est ce que c'est ?

- La **vision artificielle** désigne la compréhension d'une scène ou d'un phénomène à partir d'informations « image », liant intimement *perception*, *comportement* et *contrôle*.

□ Premier niveau : acquisition

- acquiert une image grâce à une chaîne intégrant des éléments optiques et un capteur

□ Deuxième niveau : traitement de l'image

- modifie le contenu de l'image afin de mettre en évidence des éléments d'intérêt (objets, contours)

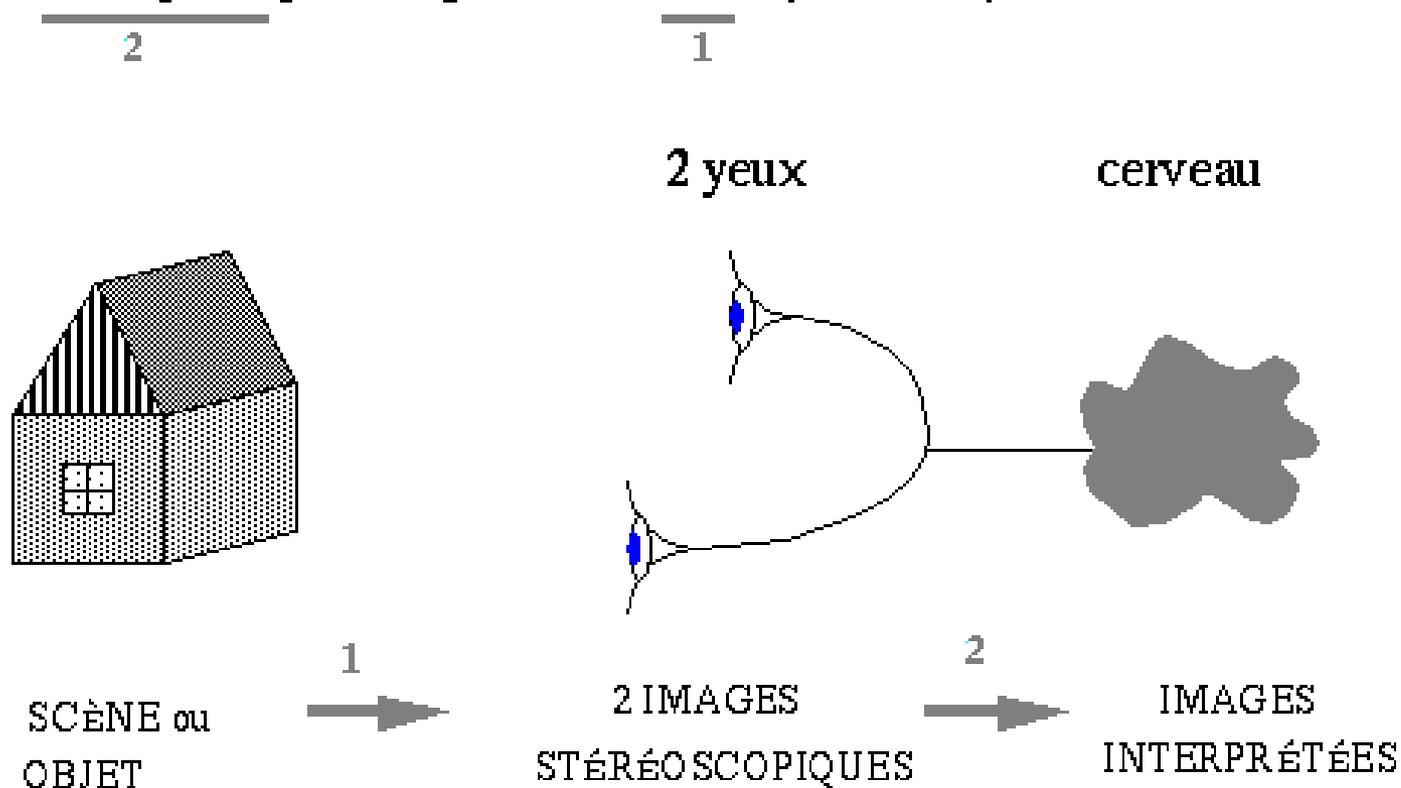
□ Troisième niveau : reconnaissance

- utilise des techniques d'intelligence artificielle pour identifier des formes connues dans l'image

Qu'est-ce que la vision ?

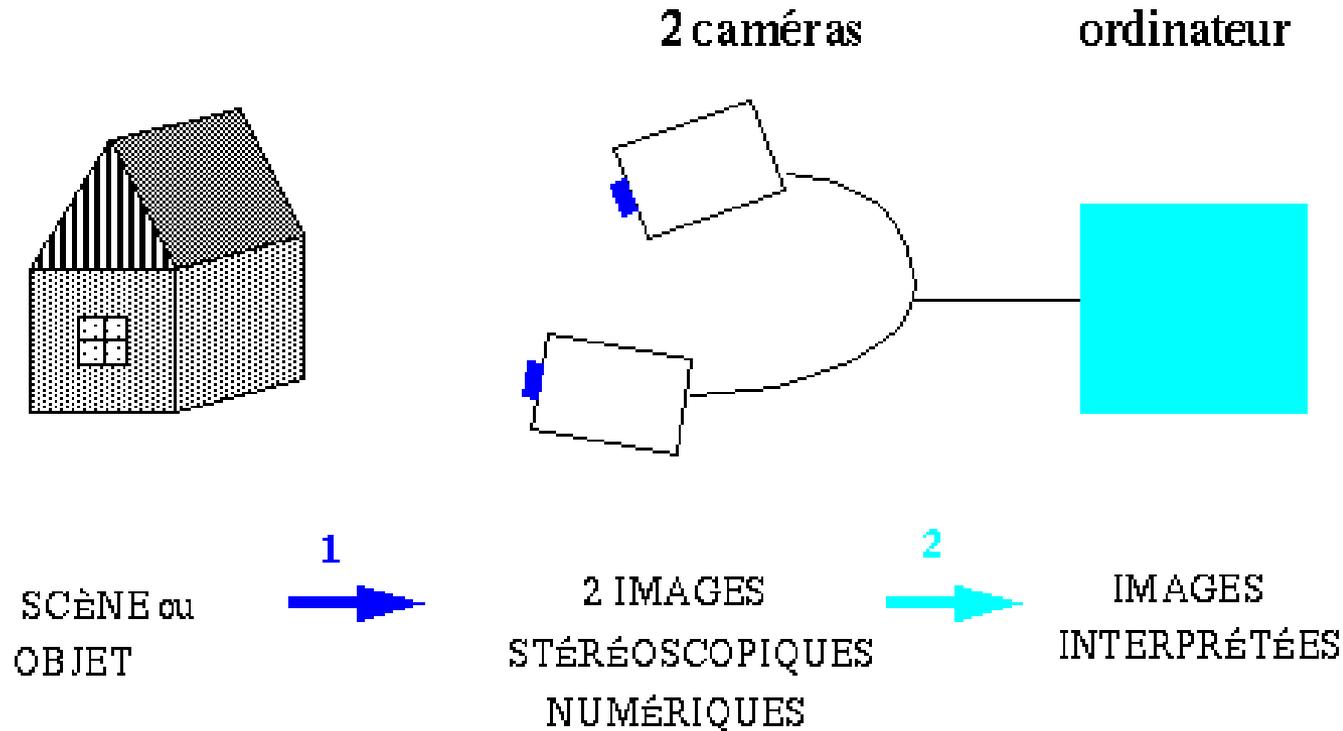
1. La vision humaine (Définition)

“Perception par l'organe de la vue” (Larousse)



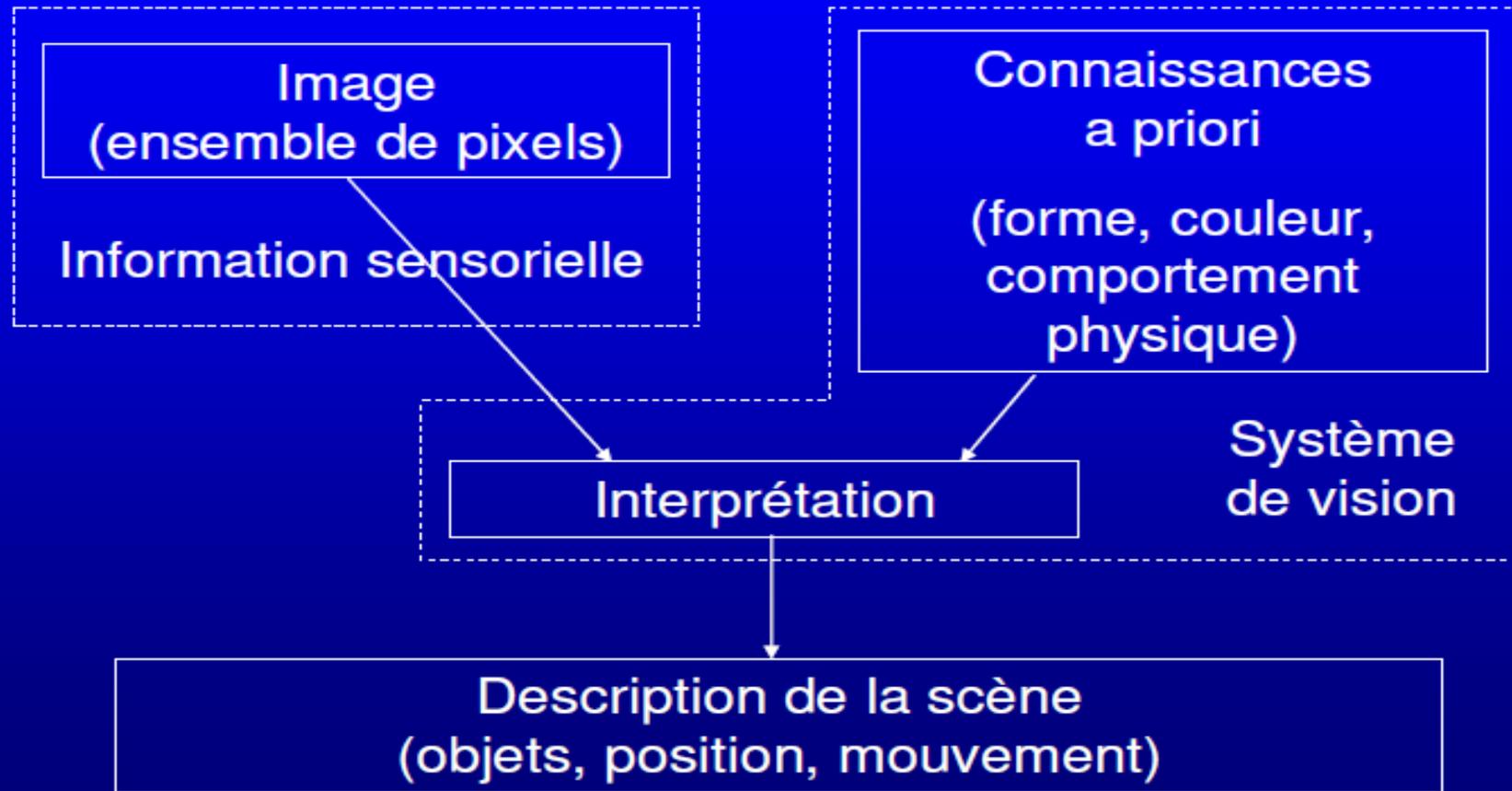
vision = vue + perception

2. La vision par ordinateur (Définition)

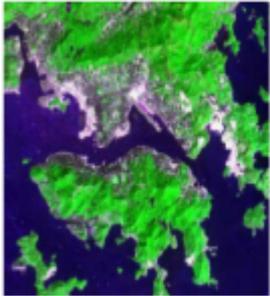


On essaie de reproduire la vision humaine pour donner de l'autonomie à un robot équipé de 2 caméras

Une théorie de la vision



Extension de la perception humaine ? Oui

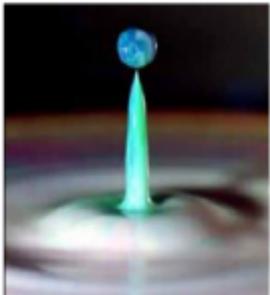


- Capteurs meilleurs que l'oeil humain
 - voient ce que nous ne voyons pas directement, du fait de la **limitation** de notre système visuel



- Autres propriétés optiques
 - problème d'**échelle**, de **résolution**, de point de vue (ex : images satellitaires)

- Autres gammes de longueurs d'onde
 - caméras **multi-spectrales**, infrarouge proche, thermiques (infrarouge lointain)



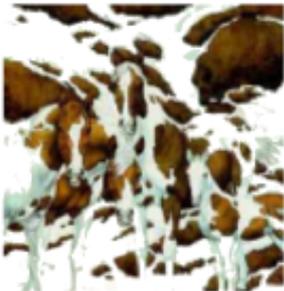
- Autres cadences
 - caméras à **haute cadence**, plusieurs milliers d'images par seconde.

Mieux que la perception humaine ? Non



■ Moins « intelligent » que l'humain

- même si les images initiales sont parfois très riches, le **résultat** de leur traitement par un système de VA reste souvent **basique**



■ Puissance de calcul

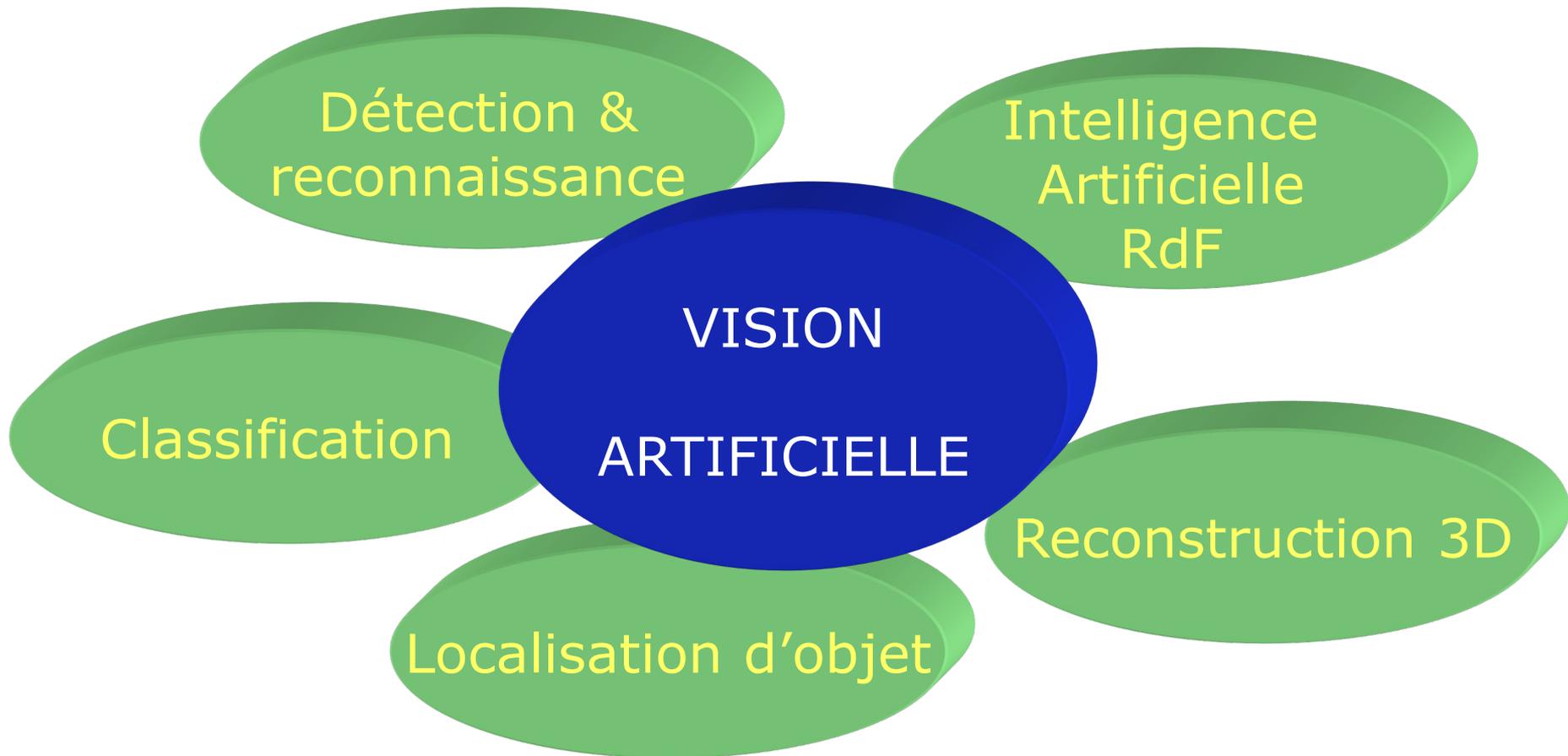
- les systèmes de VA ont une **capacité de calcul** infiniment plus limitée que celle du **cerveau** humain, ou de celui des animaux évolués



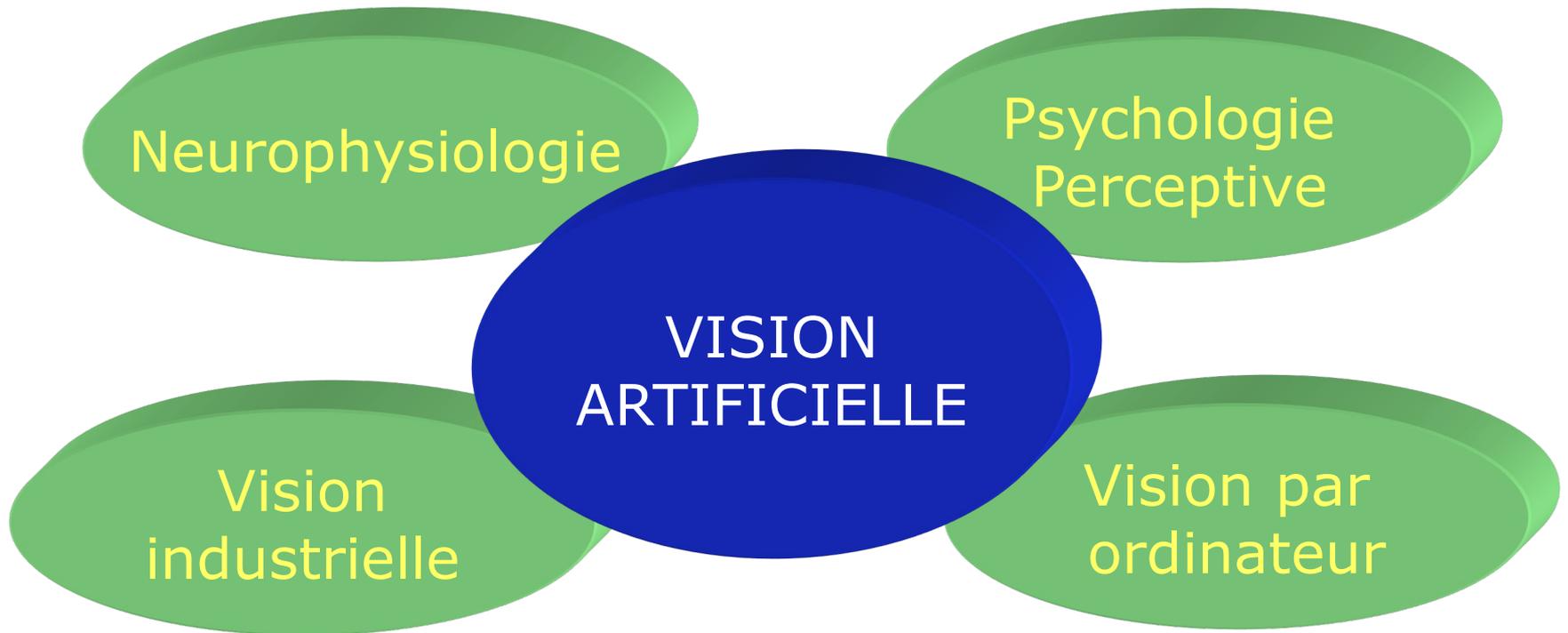
■ Connaissance et reconnaissance

- nous exploitons nos **connaissances** pour **interpréter** le contenu de l'image : on reconnaît parfois des objets alors qu'ils ne sont pas **visibles**

Vue d'ensemble d'utilisations



Vue d'ensemble de la spécialité



Exemple d'application de la VA : imagerie médicale



■ Objectifs

- aider le médecin lors du diagnostic, le chirurgien lors de la réalisation d'un geste opératoire

■ Amélioration des images

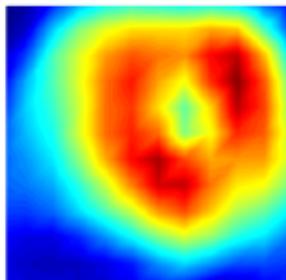
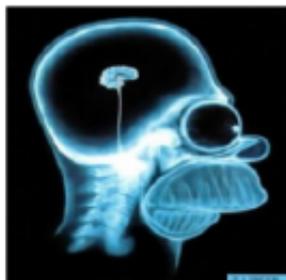
- rehaussement du contraste, élimination du bruit, **mise en évidence** des détails

■ Détection et localisation

- **positionnement** des organes, **détection** des tumeurs, **mesure** de dimensions et de volumes

■ Imagerie interventionnelle

- assistance **en-ligne** au praticien : opérations réalisées sur les images en **temps-réel**



Vision industrielle : contrôle de la qualité, de l'aspect, ...



■ Objectifs

- éviter le contrôle **visuel** par un opérateur (tâche répétitive peu valorisante)

■ Contrôle dimensionnel

- le système de vision détermine la **dimension**, la **forme**, la **position** de l'objet qu'il observe



■ Contrôle d'aspect

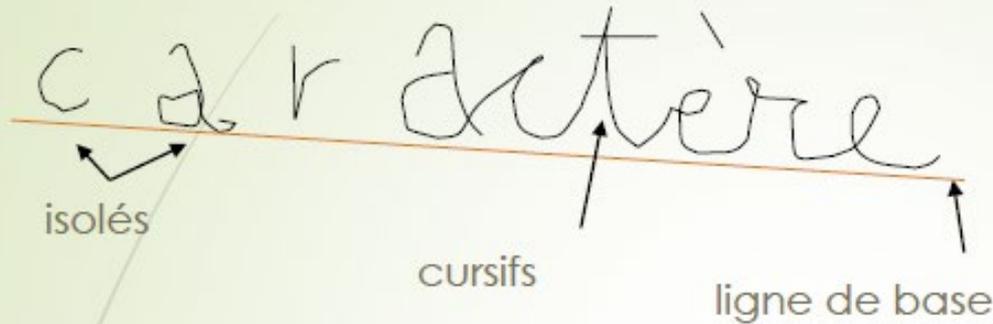
- le système détermine la **couleur**, la **texture** des objets observés



■ Contrôle de la qualité

- à partir des données précédentes, le système détermine la **qualité d'un produit**

Reconnaissance de texte

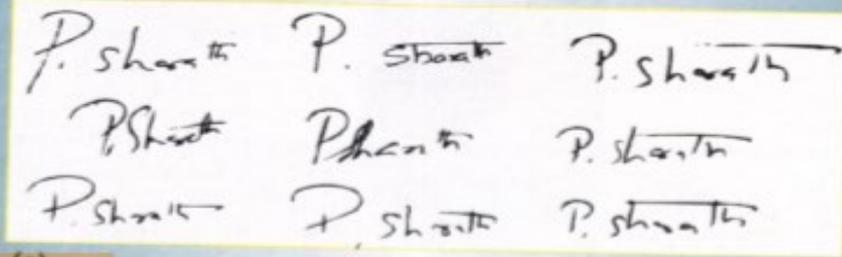


Variabilité entre scripteurs

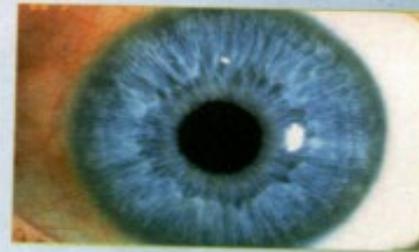


Biométrie

Signature



Iris



(Courtesy of IrisScan Inc.)

Iris

(a)
Visage

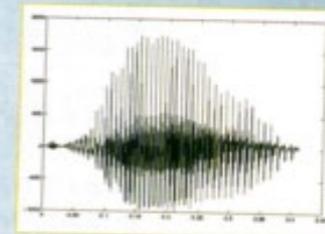


(Courtesy of Andrew Senior)



(Courtesy of Andrew Senior)

Empreinte
vocale



(Courtesy of Kluwer Academic)

Voice print

(b)
Empreinte
digitale



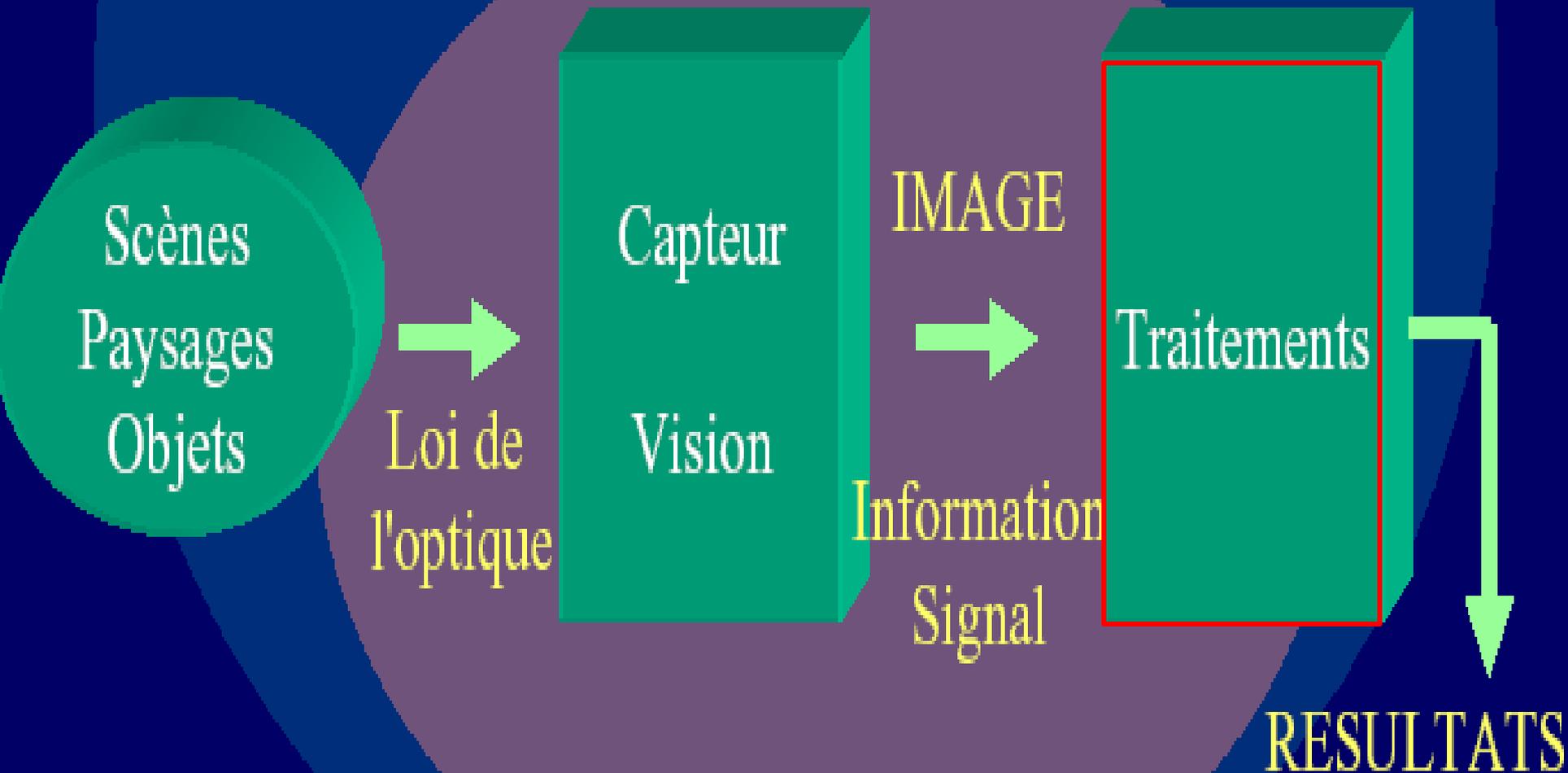
(c)



Introduction aux traitement d'images

I) A) système de vision

Image = information issue d'un capteur de vision (œil, caméra)

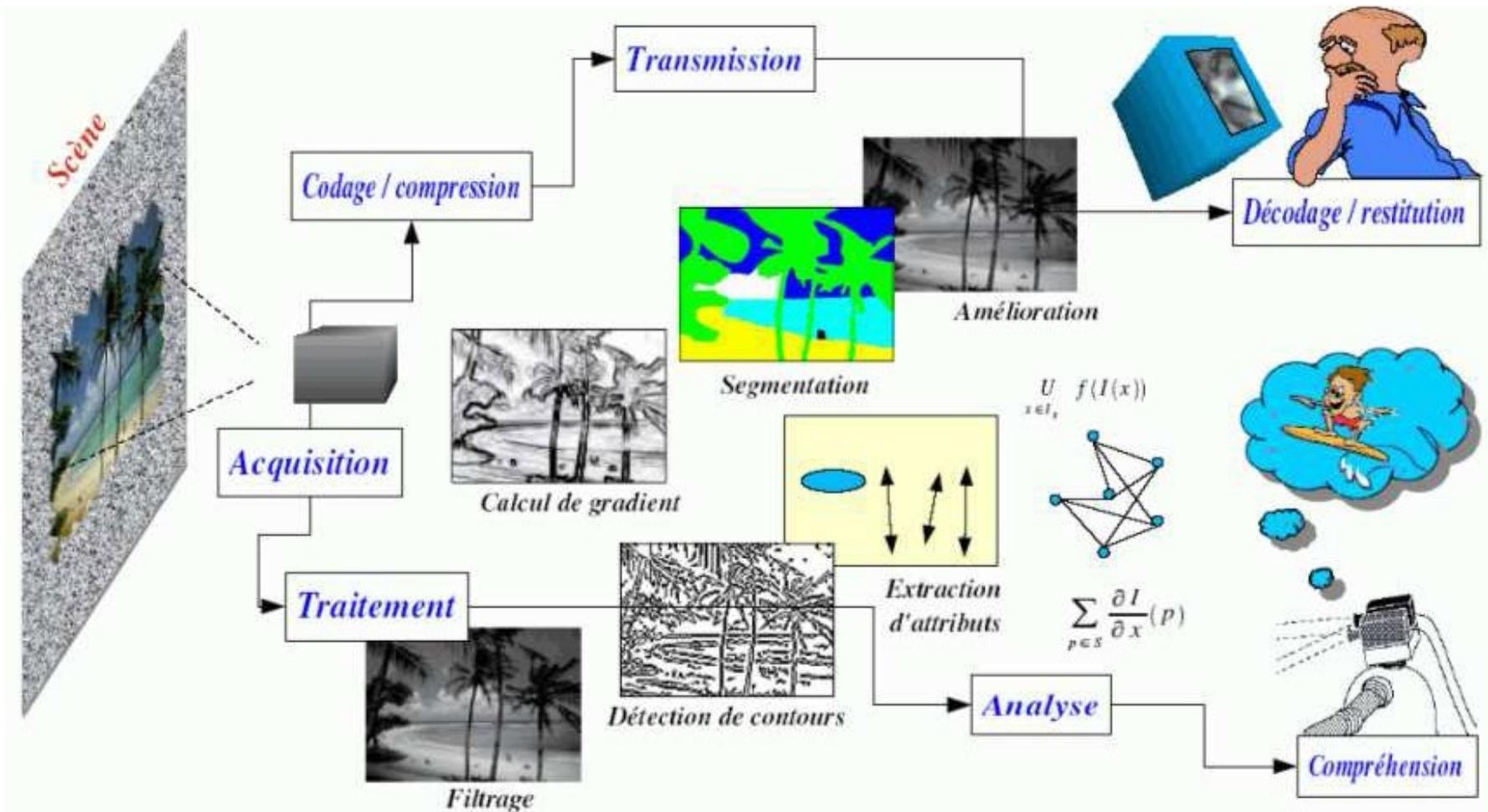


Introduction

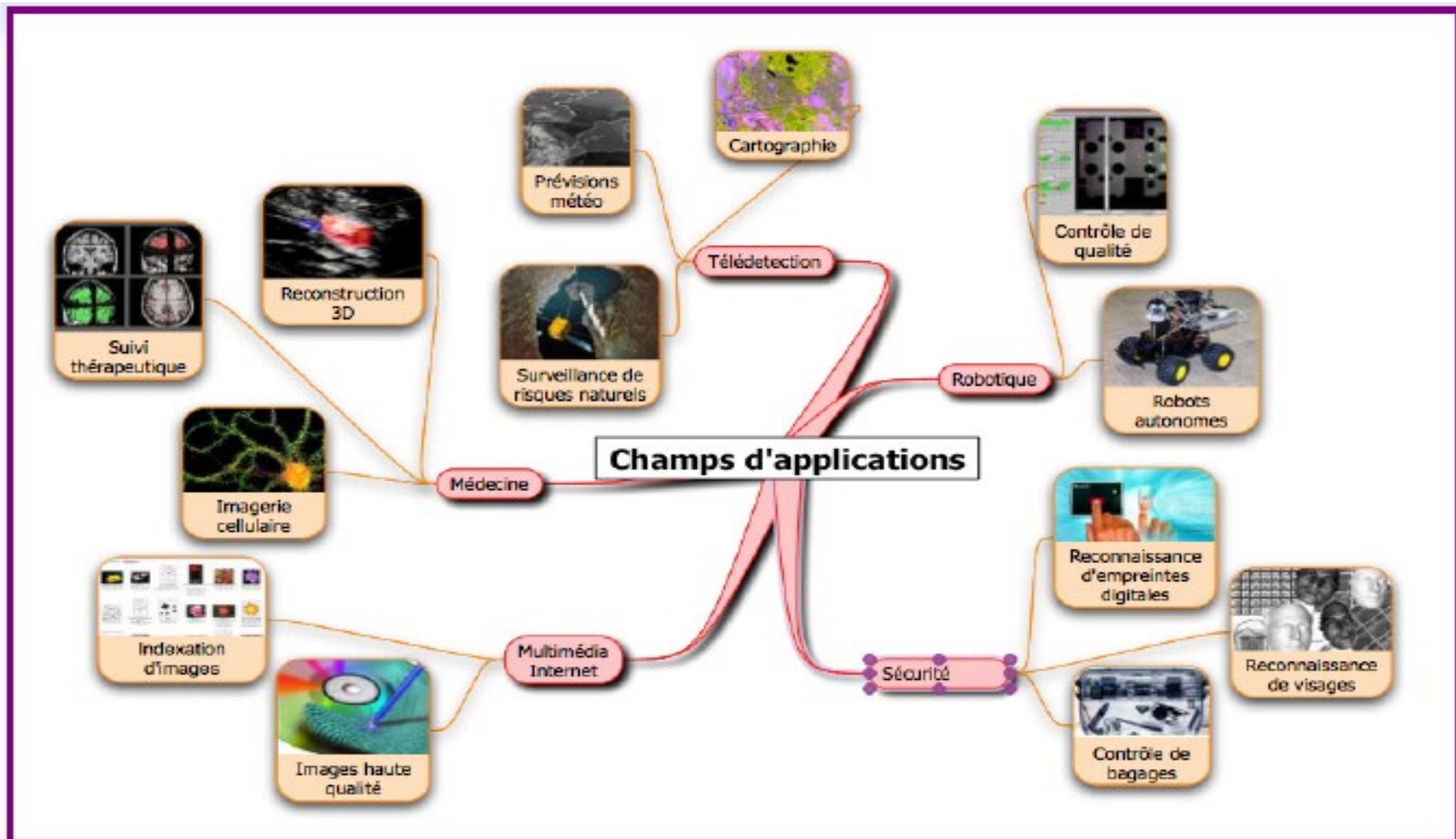
✍ Le **traitement d'images** est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'**améliorer** leur qualité ou d'**en extraire de l'information,**

✍ On désigne l'**ensemble des opérations sur les images numériques,** qui transforment une image en une autre image.

Vision globale de Systèmes à base de TI



Champs d'application



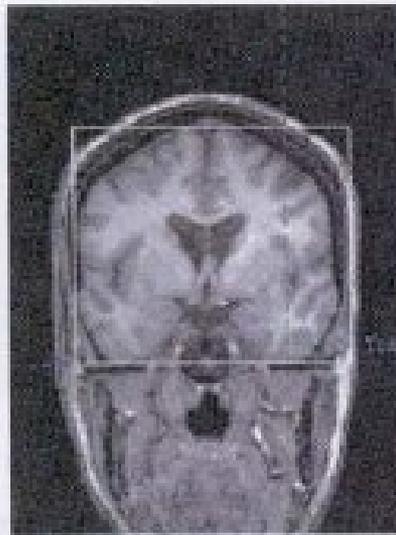
⇒ des milliers d'applications potentielles

Recherche de lésions ou d'anomalies dans le cerveau

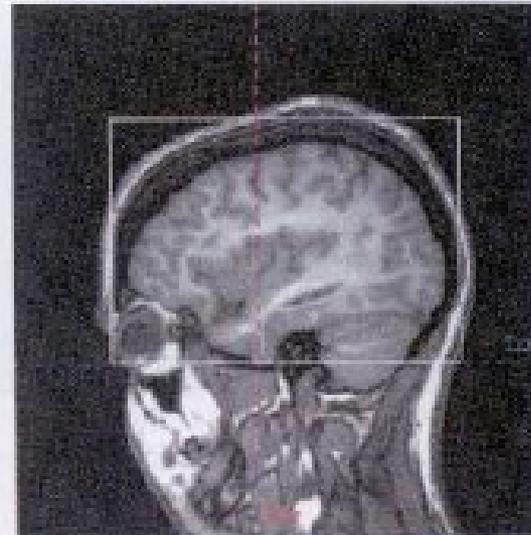
Axial View



Coronal View

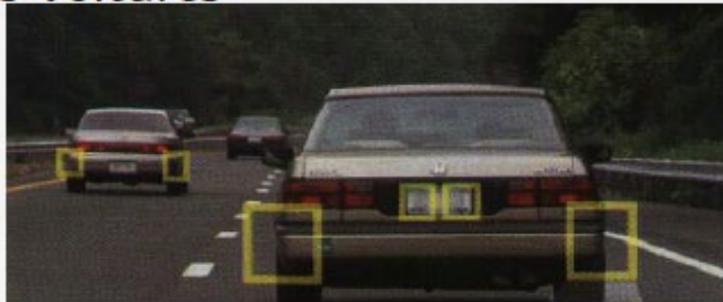


Sagittal View



Application : route intelligente

Localisation des voitures

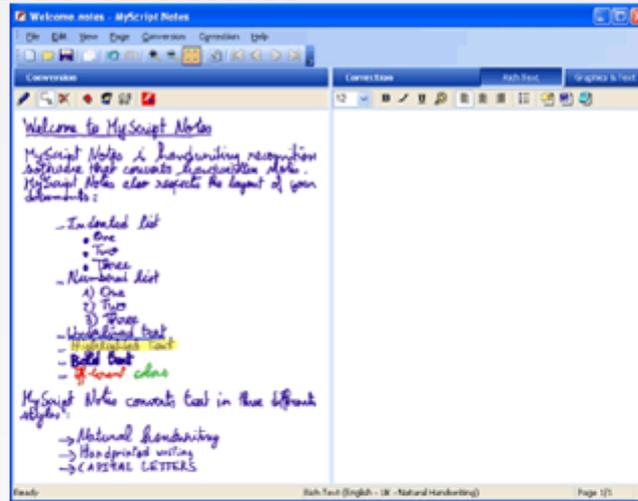


Radars automatiques

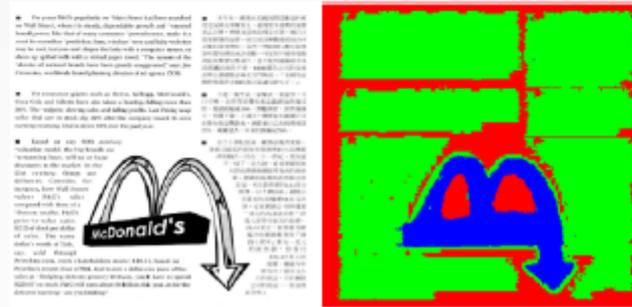


Application : analyse d'images de documents

Reconnaissance de caractères



Séparation texte/dessin



Télésurveillance

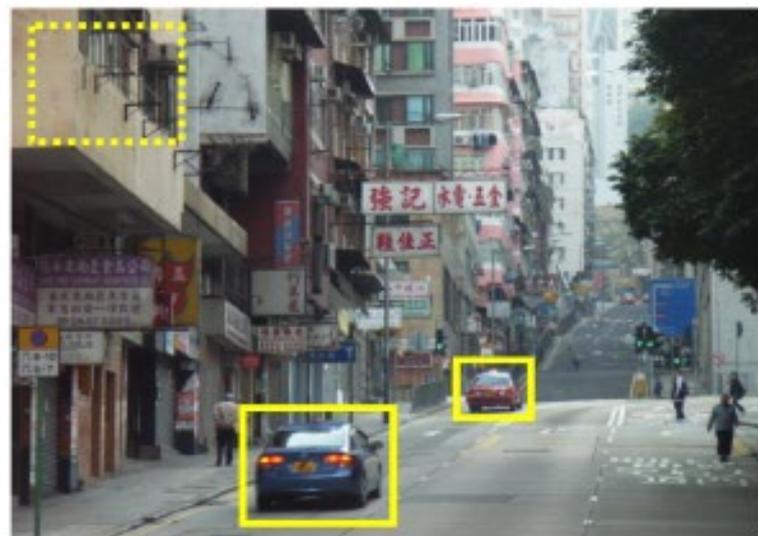
Intrusion
Trafic routier
Poursuite de cibles



Détection objets

Classification vs détection : où se situent les objets des catégorie(s) sémantique(s) présente(s) dans l'image ?

Principe : fenêtre glissante

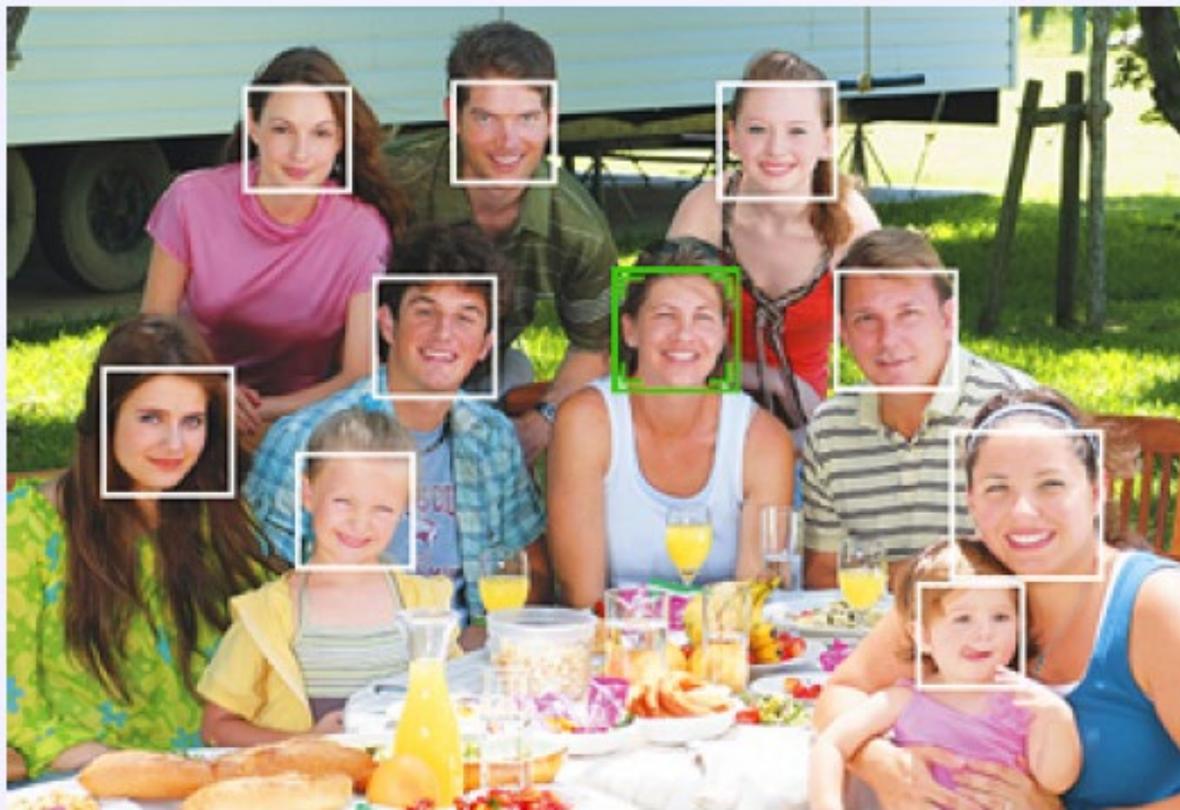


Car/non-car
Classifier

Détection d'objets

Détection visages : succès

- ▶ Dans la plupart des appareils photo

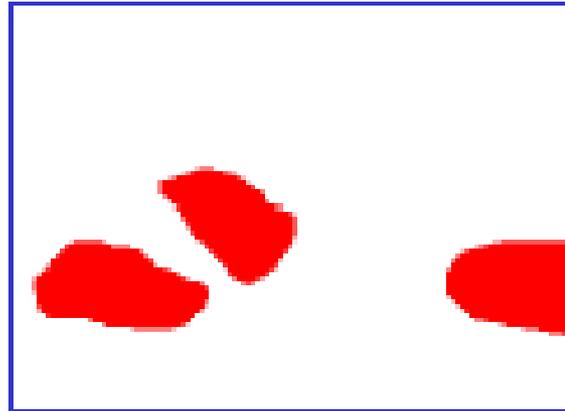


Robotique

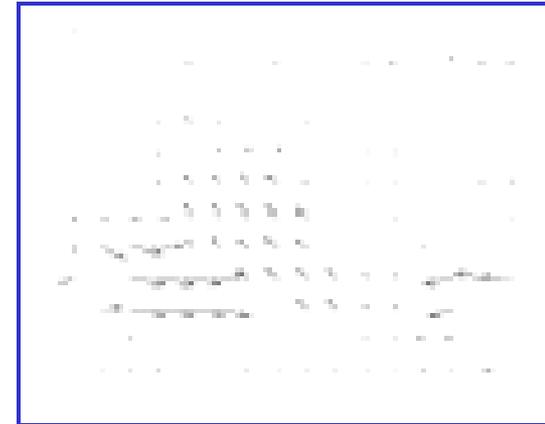
Détection d'obstacles
Ego-mouvement
Guidage de missiles



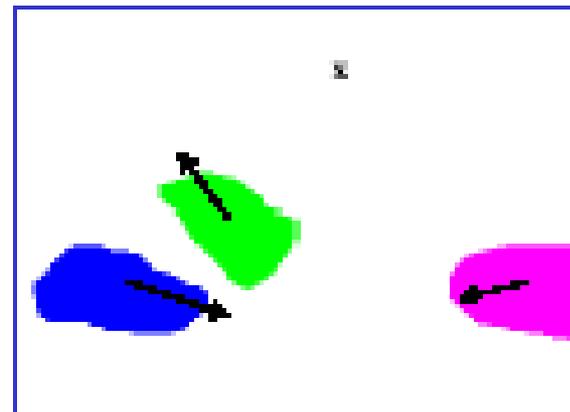
Analyse du mouvement



DETECTION



ESTIMATION

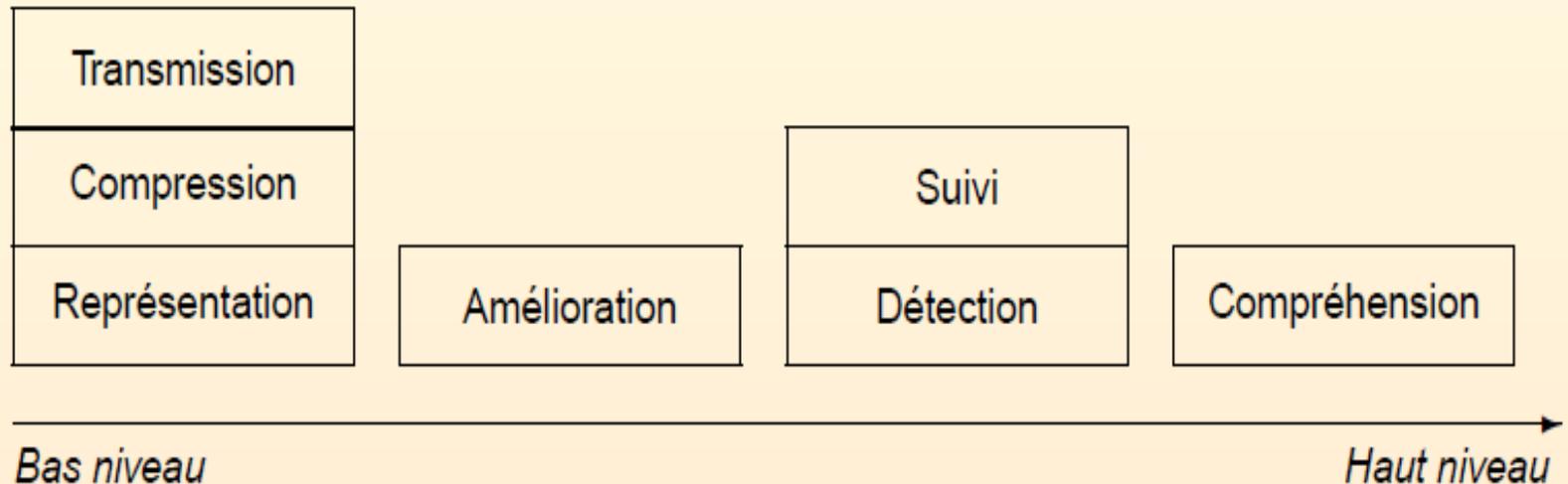


SEGMENTATION

Du traitement d'images à la vision par ordinateur

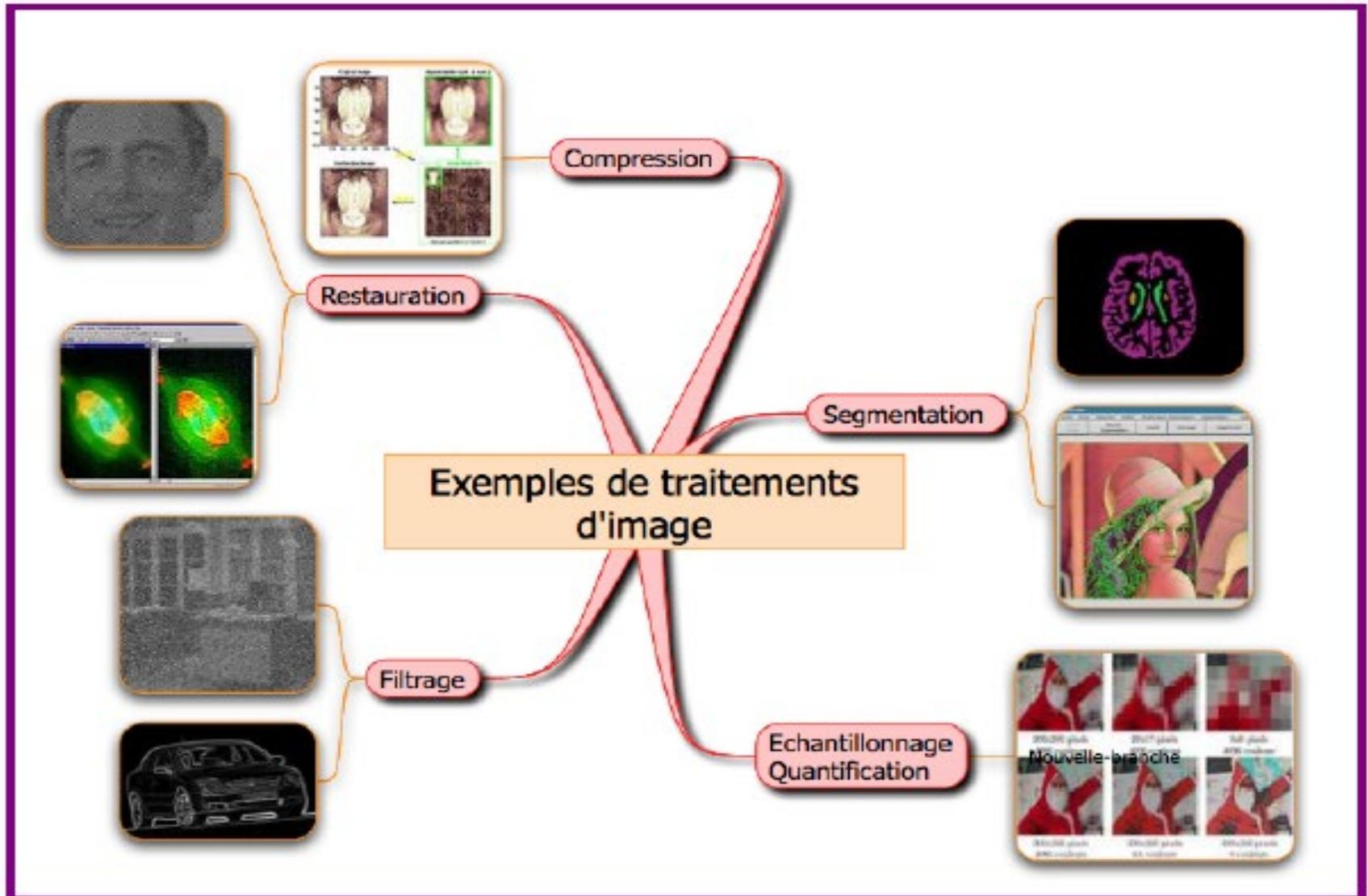
Traitement d'images

Vision par ordinateur



- Pourquoi le traitement d'images ?
 - Le futur est au multimedia : les images sont partout!
 - Les applications sont multiples

Exemples de traitements bas niveau des images



Exemples de traitements bas niveau des images

- **Amélioration d'images** : amélioration de l'aspect des images
 - ↳ restauration, débruitage, filtrage, déconvolution ...
- **Analyse fréquentielle** : changement d'espace de représentation
 - ↳ Transformée de Fourier , ondelette
- **Acquisition d'images** : disposer d'une image sous forme numérique
 - ↳ échantillonnage, quantification
- **Compression d'images** : réduction de la quantité d'informations nécessaires pour représenter une image
 - ↳ codage, transmission
- **Segmentation d'images** : division de l'image en parties, ou zones caractéristiques
 - ↳ détection de contours, partitionnement split & merge

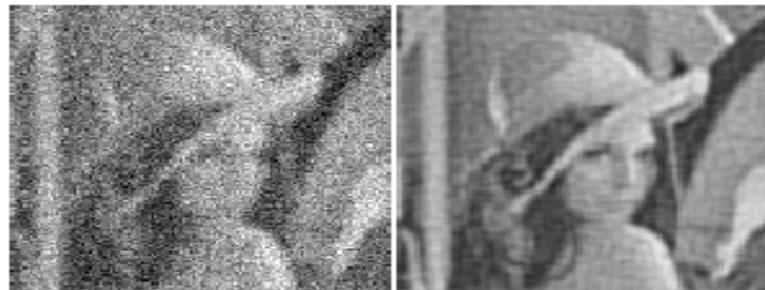
Exemples de traitements moyen est niveau des images

- **Représentation et description d'images** : transformation de l'image en un ensemble de "concepts" utilisables par l'ordinateur pour des algorithmes plus évolués
 - ↪ détection de caractéristiques, modèles (graphes, ...)
- **Indexation d'images** : tri d'images selon un ensemble de descripteurs
 - ↪ fouille dans les bases d'images
- **Reconnaissance/Compréhension** : association d'une étiquette à un objet à partir de ses descripteurs
 - ↪ Reconnaissance de visages...
- **Analyse** : convertir en informations
- **Reconstruction 3D**: obtenir un volume à partir de plans (images)

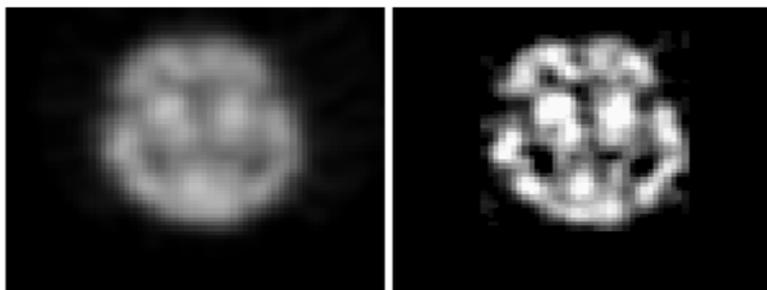
Exemples de traitement d'images



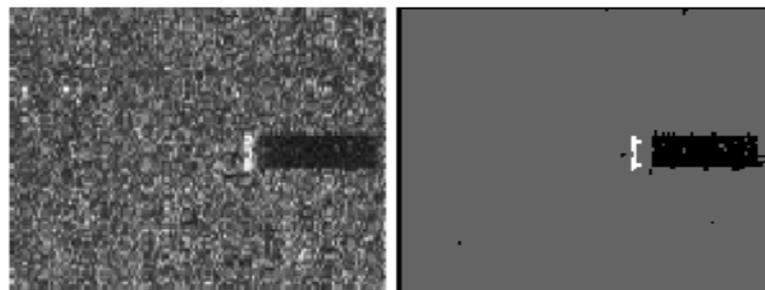
Compression



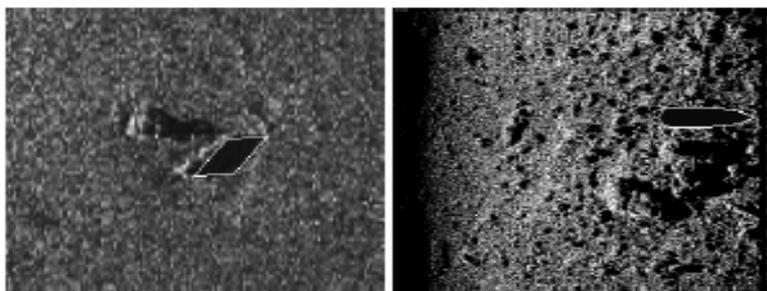
Restauration



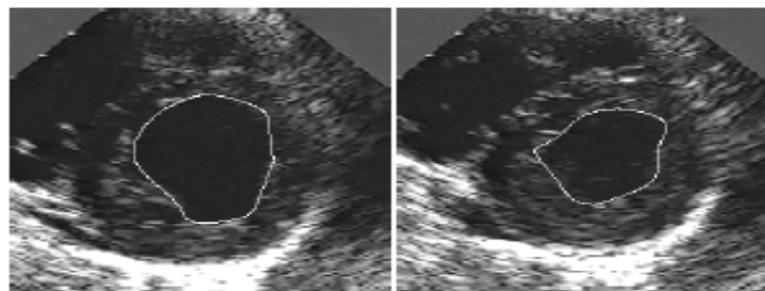
Déconvolution



Segmentation



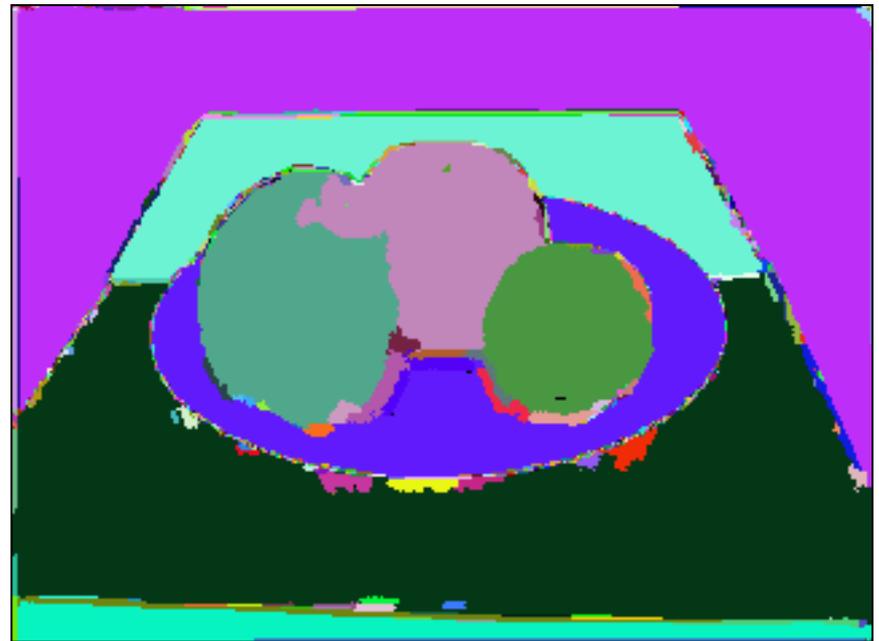
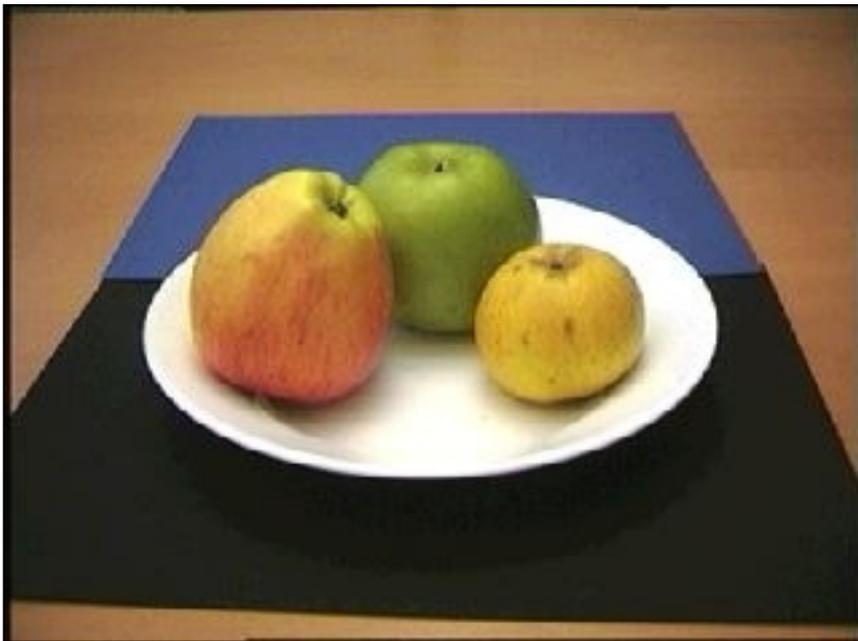
Reconnaissance



Suivi de formes

Segmenter une image?

*Partitionner l'image en zone homogènes selon un critère déterminé:
couleur, texture, niveau de gris, indice,...*



Exemples de traitement d'images

Segmentation : Objet ou région ?

Deux concepts de base:



Objets

Avec une signification sémantique:
Leur sélection dépend de l'application.
Composés de régions.

dépendant de
l'application

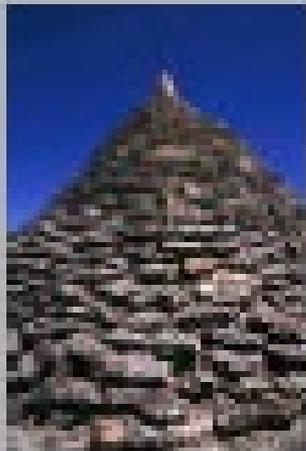


Régions

Homogènes d'après un critère donné
(couleur, mouvement, texture...).

Exemples de traitement d'images

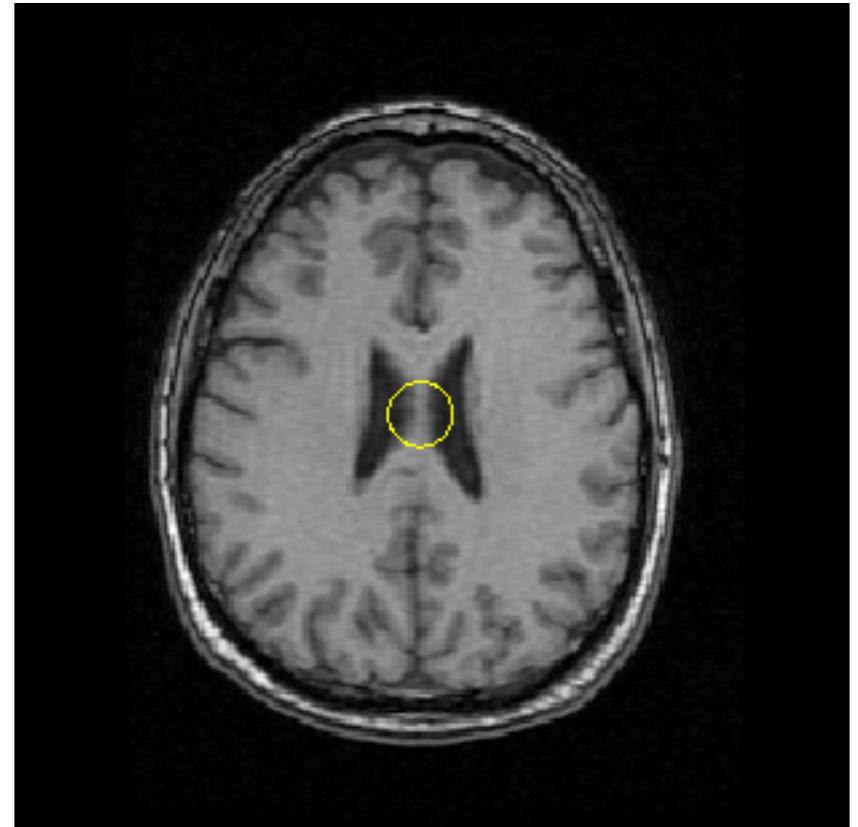
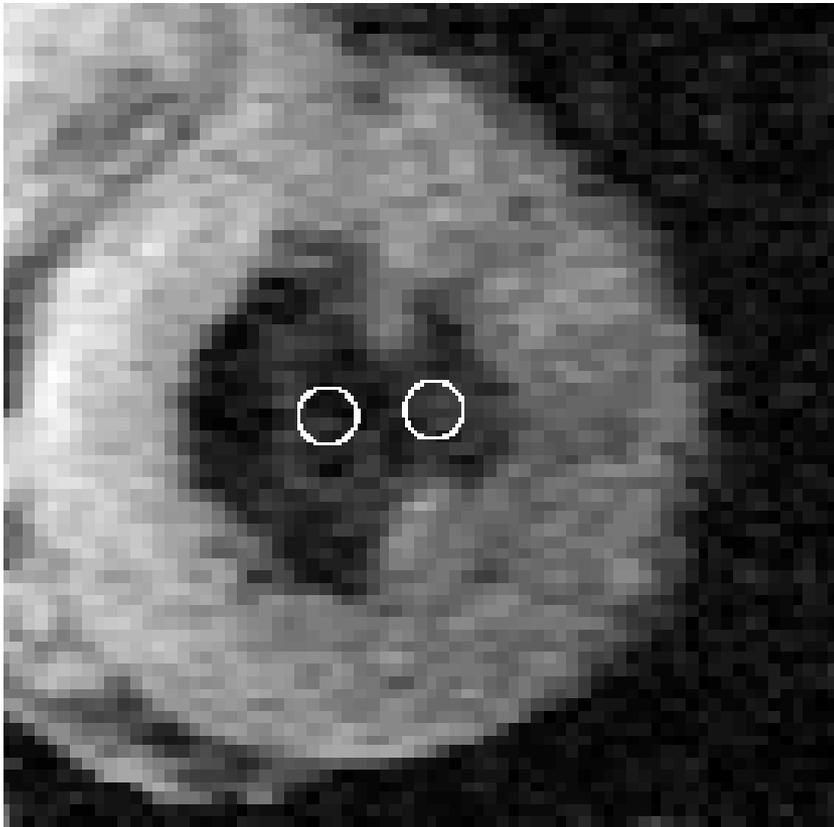
- Segmentation par classification (suite) :



Images originales

Images segmentées

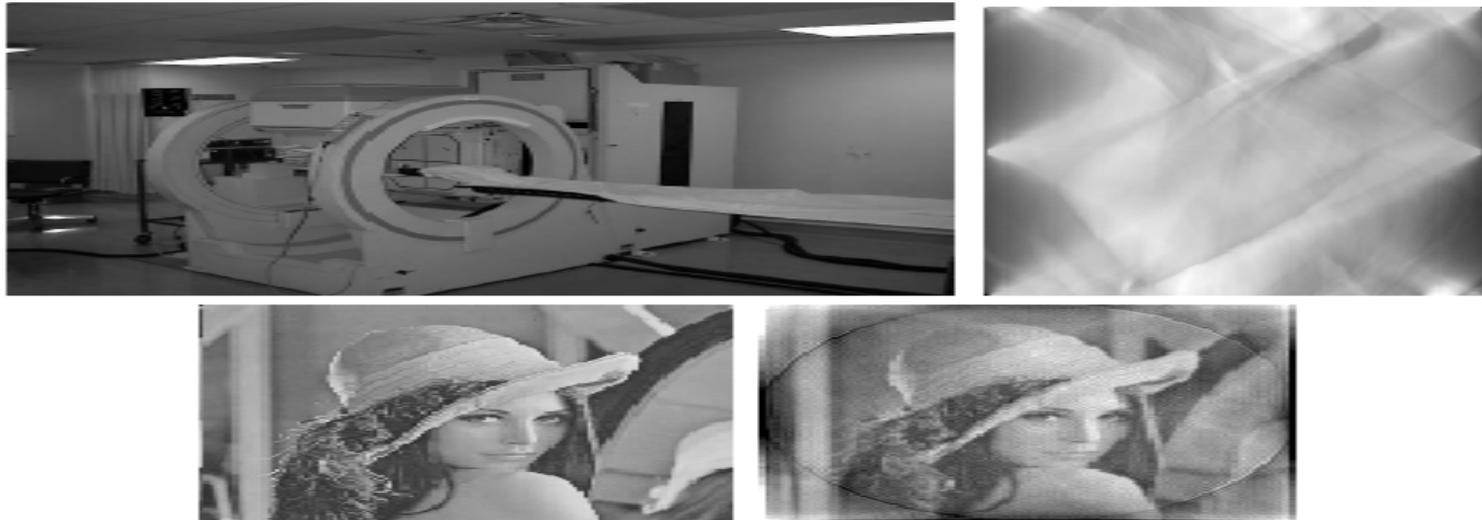
Segmentation



Exemples de traitement d'images



Rehaussement & déconvolution



Reconstruction tomographique

Qu'est ce qu'une image ?

- Une image est avant tout un **signal 2D** (x,y)
- Souvent, cette image représente une **réalité 3D** (x,y,z)

- D'un point de vue **mathématique** :
 - Une image est un matrice de nombres représentant un signal
 - Plusieurs outils permettent de manipuler ce signal

- D'un point de vue **humain** :
 - Une image contient plusieurs informations sémantiques
 - Il faut interpréter le contenu au-delà de la valeur des nombres

Définition d'une image réelle

Définition

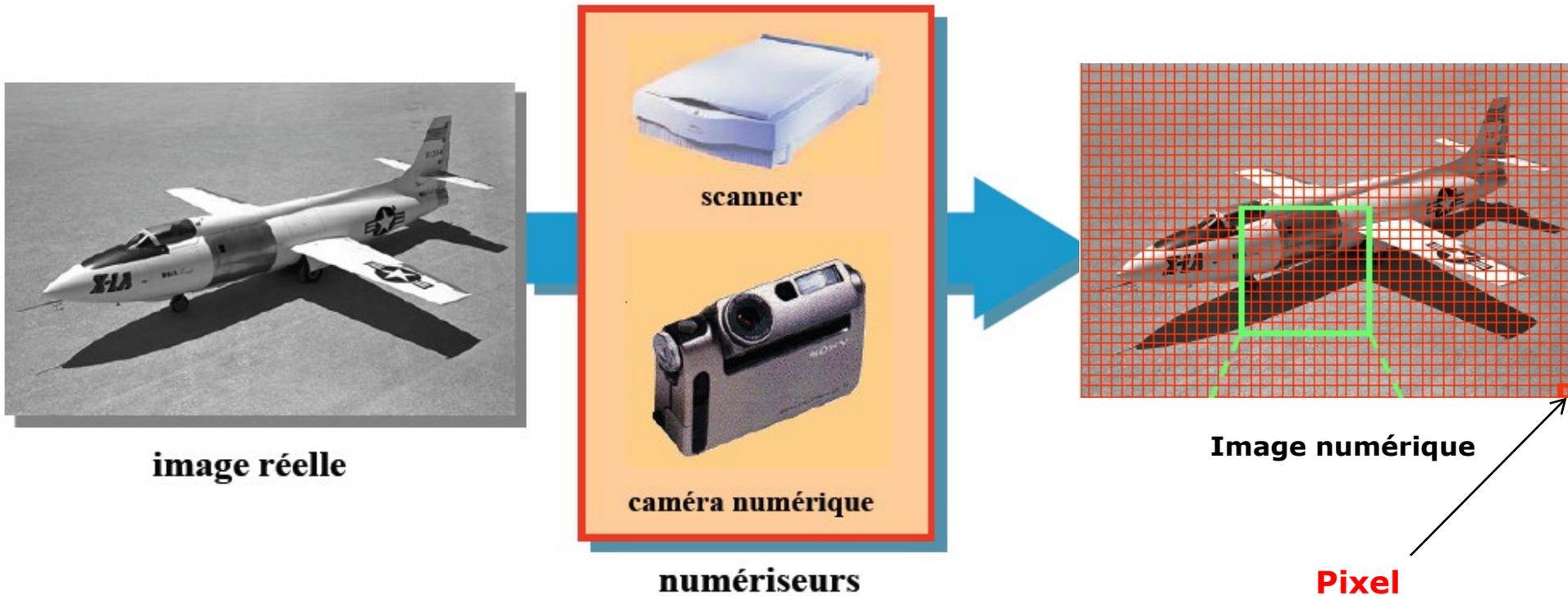
- Une image est la **projection** sur un plan d'une scène 3D
- Elle peut être définie comme une fonction à deux variables $f(x, y)$
 - (x, y) est la **position** d'un point de l'espace sur le plan de projection
 - $f(x, y)$ est l'**intensité** (ou **brillance**) au point de coordonnées (x, y)
- Une image est un plan analogique dans lequel les intensités sont réelles

Définition d'une image numérique

Définitions et notations

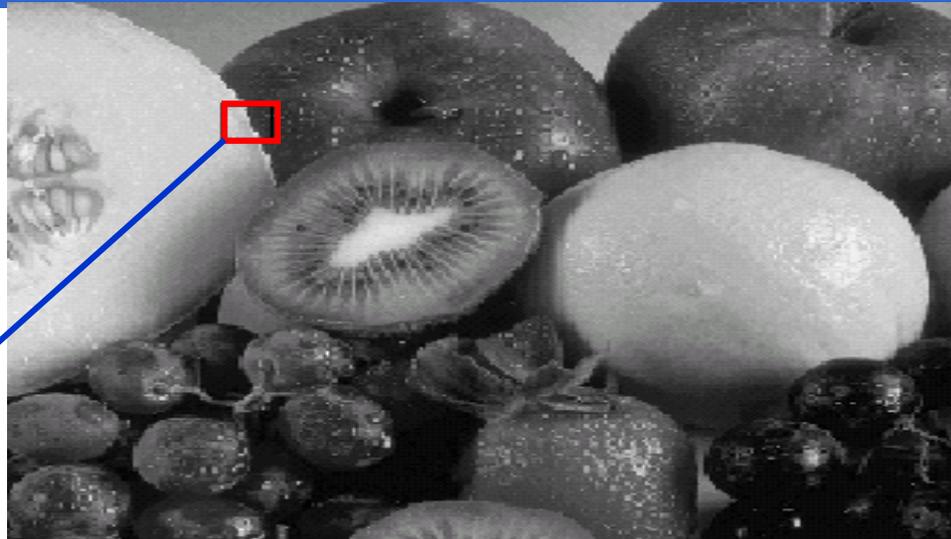
- ▶ Une image est représentée :
 - par fonction continue $f(x, y)$, $x, y \in \mathbb{R}$;
 - par une fonction numérique $f(i, j)$ (ou $f(n, m)$), $i, j \in \mathbb{N}$ ($n, m \in \mathbb{N}$) et $f \in \mathbb{N}^+$, après numérisation.
- ▶ Image analogique \rightarrow image numérique : numérisation en deux étapes :
 - ① échantillonnage spatial : discrétisation des coordonnées de l'image réelle ;
 - ② quantification des luminances : discrétisation des intensités de l'image réelle.
- ▶ Une image numérique est composée d'un ensemble fini d'éléments, appelés *picture element*, ou pixels (voxels en 3D)

Numérisation



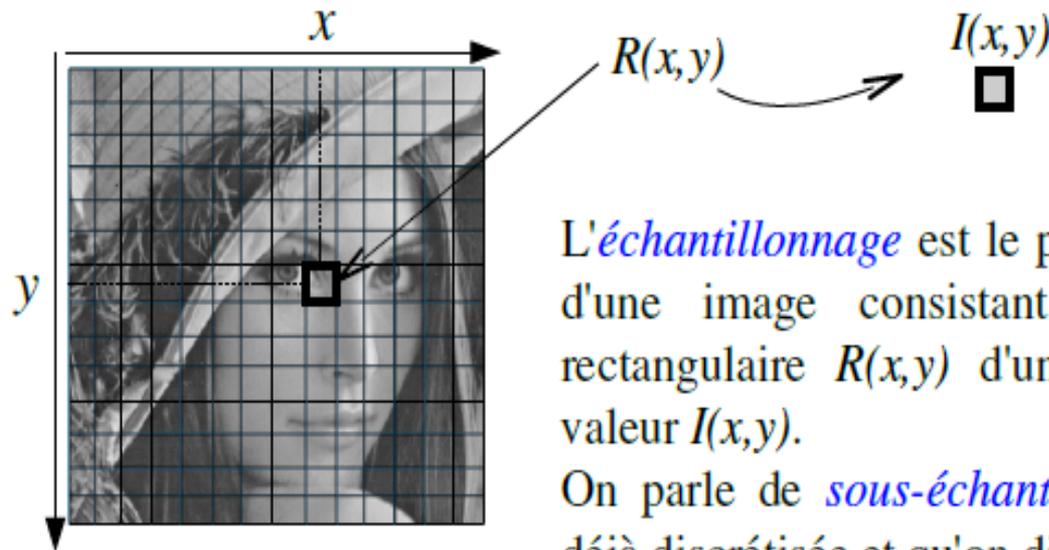
Numérisation = Échantillonnage + Quantification

Exemple : Images numériques



x =	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
y =															
41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
42	206	196	203	197	195	210	207	56	63	58	53	53	61	62	51
43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	50
44	216	206	211	193	202	207	208	57	69	60	55	77	49	62	61
45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	108
46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48
47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55
48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56
49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48
50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56
51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52
52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56
53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60
54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66
55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63

Images numériques



L'*échantillonnage* est le procédé de discrétisation spatiale d'une image consistant à associer à chaque zone rectangulaire $R(x,y)$ d'une image continue une unique valeur $I(x,y)$.

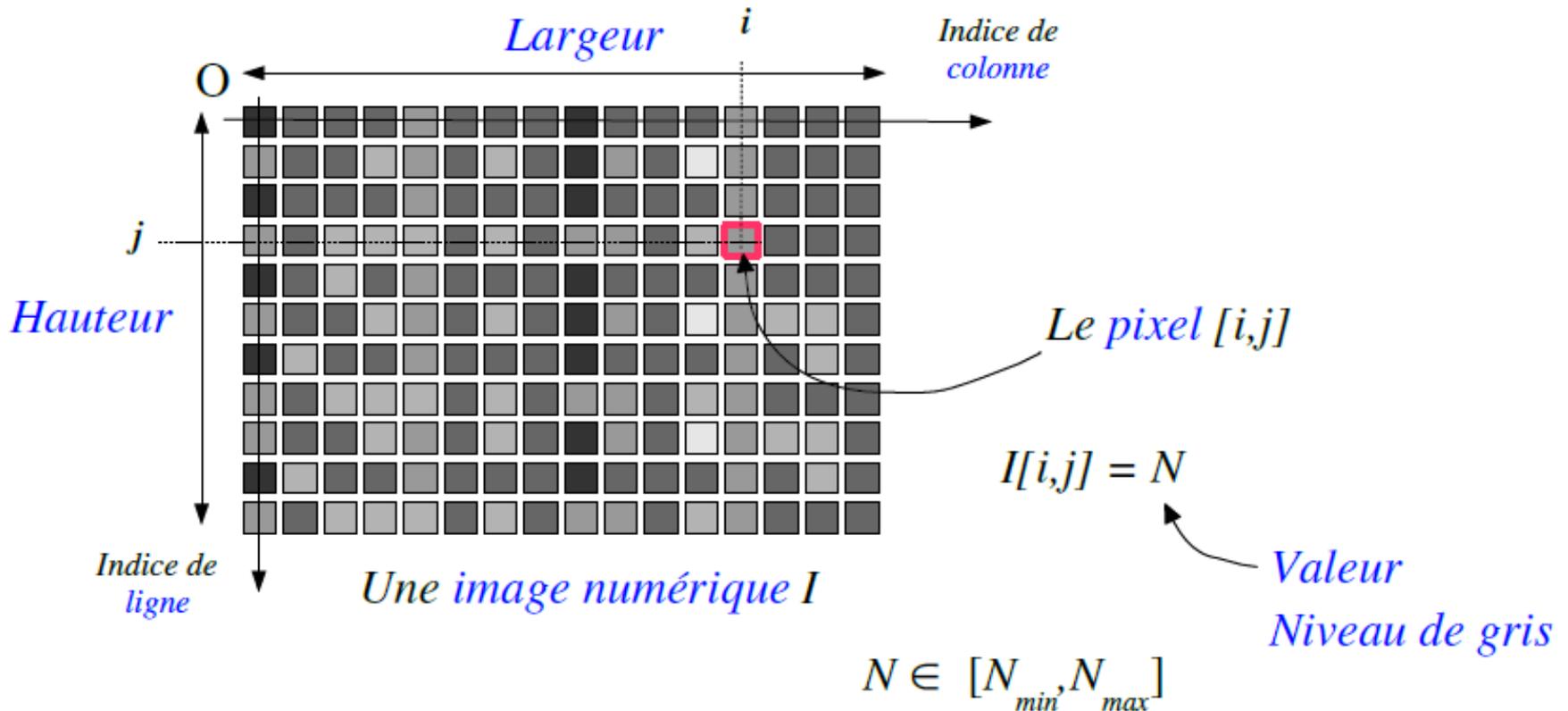
On parle de *sous-échantillonnage* lorsque l'image est déjà discrétisée et qu'on diminue le nombre d'échantillons.

La *quantification* désigne la limitation du nombre de valeurs différentes que peut prendre $I(x,y)$.



Une *image numérique* est une image *échantillonnée et quantifiée*.

Pixels et niveaux de gris



Images numériques

Une image numérique $s(k,l)$ de taille $K \times L$ est représentée par une matrice de même taille:

$$s(k,l) = \begin{bmatrix} s(0,0) & s(0,1) & \cdots & s(0,L-2) & s(0,L-1) \\ s(1,0) & s(1,1) & \cdots & s(1,L-2) & s(1,L-1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s(K-2,0) & s(K-2,1) & \cdots & s(K-2,L-2) & s(K-2,L-1) \\ s(K-1,0) & s(K-1,0) & \cdots & s(K-1,L-2) & s(K-1,L-1) \end{bmatrix}$$

Chaque élément est représenté par un entier, i.e. par un nombre fini de bits.

Échantillonnage

Échantillonnage - Définition

- Prélèvement périodique des échantillons d'un signal analogique
- Exemple signal 2-D $f(x, y)$

$$s(k, l) \equiv s_e(x, y) \Big|_{x=k\Delta x, y=l\Delta y}$$
$$s_e(x, y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(x, y) \delta(x - k\Delta x, y - l\Delta y)$$

Δx Pas d'échantillonnage pour dimension X

Δy Pas d'échantillonnage pour dimension Y

Quantification des luminances

Définition

- L'intensité I est quantifié sur m bits et peut prendre $L = 2^m$ valeurs : $I \in [0, \dots, 2^m - 1]$
- Plus petit changement d'intensité discernable dans l'image

Attention

Un quantification trop faible provoque des "faux contours"

Exemple

- $m = 1$: 2 valeurs possibles (images **binaires**)
- $m = 8$: 256 valeurs possibles (images en **niveaux de gris**)
- $m = 16$: 65535 valeurs possibles (images en **couleurs**)

Exemple : Échantillonnage et Quantification

Résolution...

...spatiale :

Échantillonnage



256x256



128x128



64x64



32x32

...tonale :

Quantification



6 bits



4 bits



3 bits



2 bits



1 bit

Exemple : Quantification

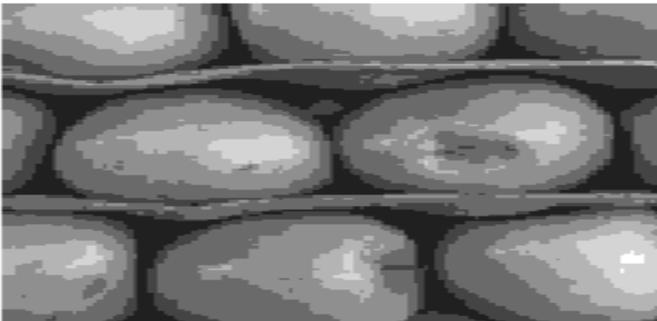
bits=1



bits=2



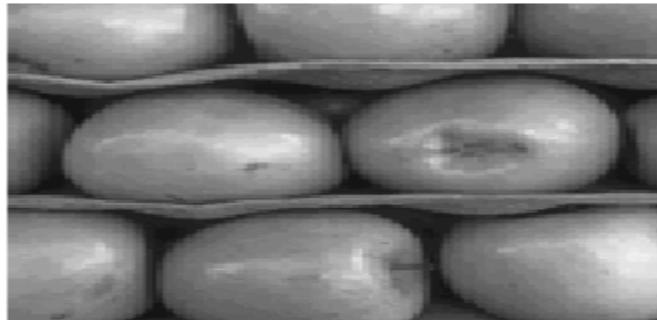
bits=3



bits=4



bits=8



Types d'images numériques

Définition

Niveau de gris : valeur de l'intensité lumineuse $f(x, y)$ au pixel de coordonnées (x, y)

Exemple

- **Image binaire** : deux valeurs possibles d'intensité (0 ou 1) pour les pixels
- **Image en niveaux de gris** :
 - Quantification des luminances sur l'intervalle $[0, 255]$
 - Codage sur 8 bits (1 octet) : $2^0 - 1 \leq k \leq 2^8 - 1$
 - Convention : **noir=0**, **blanc=255**



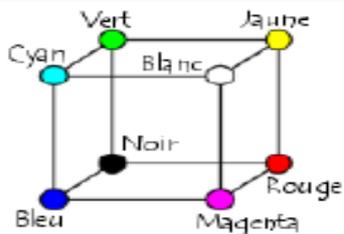
Codage des valeurs des pixels

En niveau de gris

- ▶ la valeur $I(i, j)$ d'un pixel $s = (i, j)$ représente son intensité lumineuse
- ▶ binaire : $I(i, j) = 0$ noir ou $I(i, j) = 1$ blanc
- ▶ codage 8 bits (le plus classique) : $I(i, j) = 0, \dots, 255$ du plus foncé au plus clair

En couleur

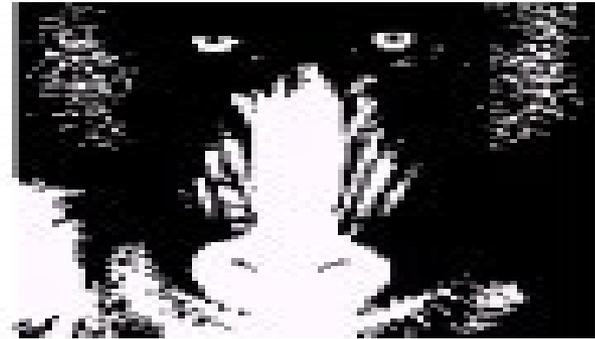
- ▶ codage dans l'espace RGB : trois intensités lumineuses rouge, vert, bleu.
- ▶ codage 24 bits : $R(i, j) = 0, \dots, 255$; $V(i, j) = 0, \dots, 255$; $B(i, j) = 0, \dots, 255$



Trois principaux types d'images



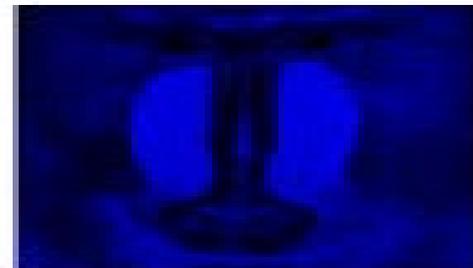
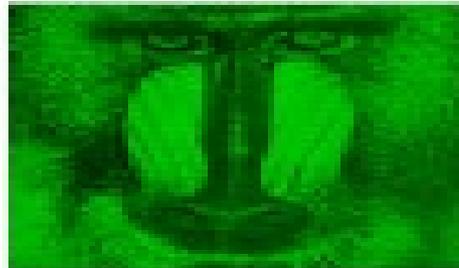
Images en niveaux de gris
 $I(x,y) \in [0..255]$



Images binaires
 $I(x,y) \in \{0, 1\}$



Images couleurs
 $I_R(x,y) I_G(x,y) I_B(x,y)$

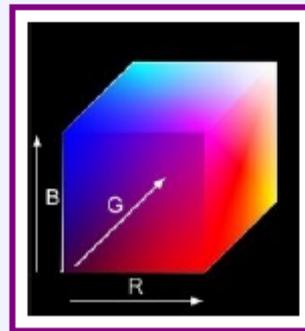


Espace RVB / RGB

- ❖ Espace le plus couramment utilisé, notamment lors de l'acquisition par scanner et de l'affichage sur moniteur d'ordinateur.

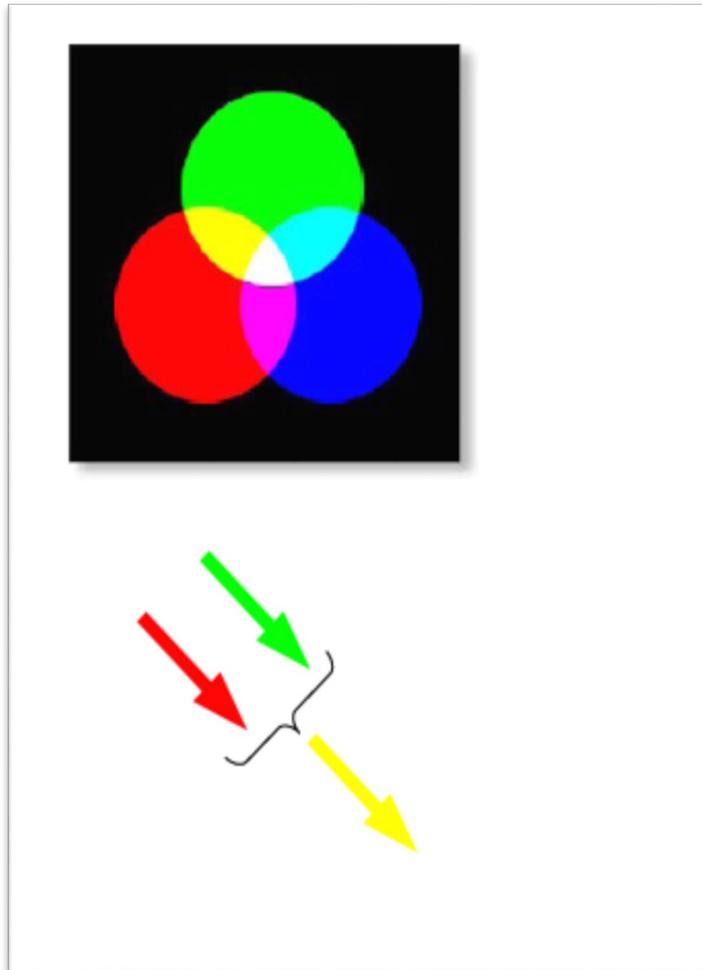
Définition

- ▶ Trois couleurs primaires : rouge, vert et bleu
- ▶ Synthèse additive
- ▶ Utilisé en éclairage afin d'obtenir toutes les couleurs visibles par l'homme.
- ▶ Difficile de se faire une idée de la luminosité dans ce mode colorimétrique : $L = 0.3R + 0.59G + 0.11B$



Synthèses additives

Synthèses additives des couleurs



- basé sur la synthèse additive des couleurs → mélange de 3 composantes
- Souvent, chaque couleur sur 8bits donc $I = \{0, \dots, 255\}^3$
- Utilisé par les écrans d'ordinateurs

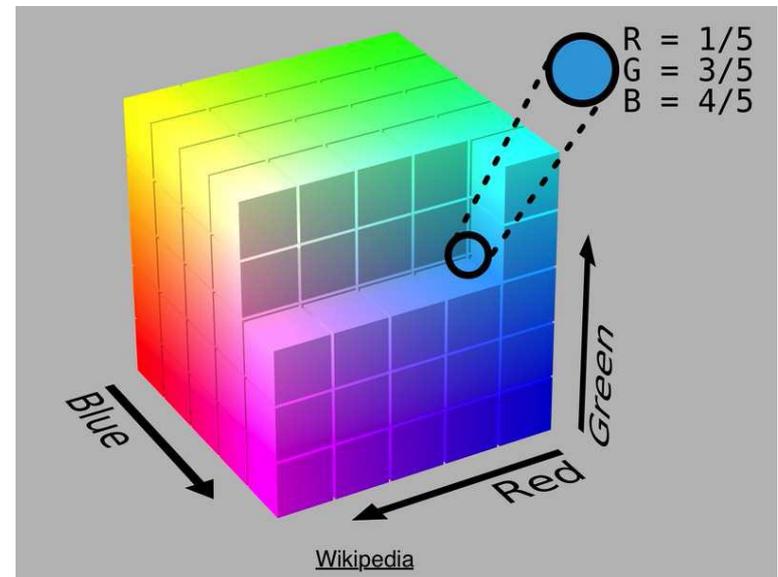
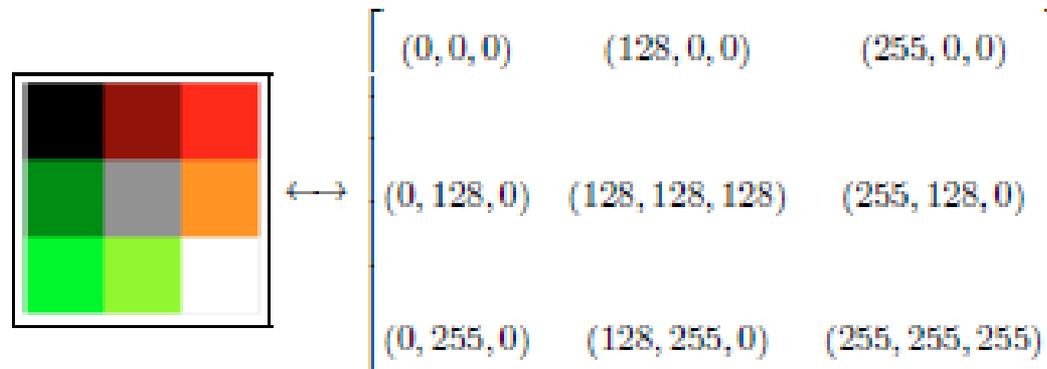


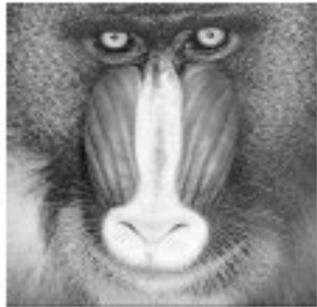
Image couleurs

exemple

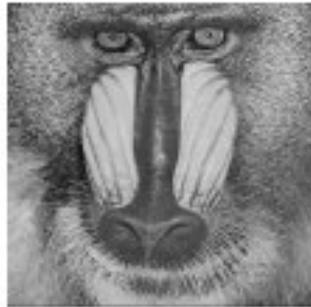


ou encore → $\left(\begin{bmatrix} 0 & 128 & 255 \\ 0 & 128 & 255 \\ 0 & 128 & 255 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 128 & 128 & 128 \\ 255 & 255 & 255 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 128 & 0 \\ 0 & 0 & 255 \end{bmatrix} \right)$

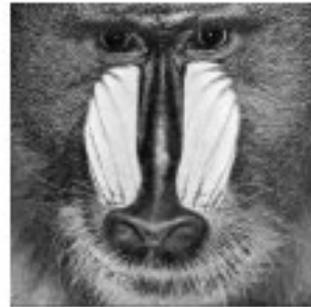
Espace RVB



R



G



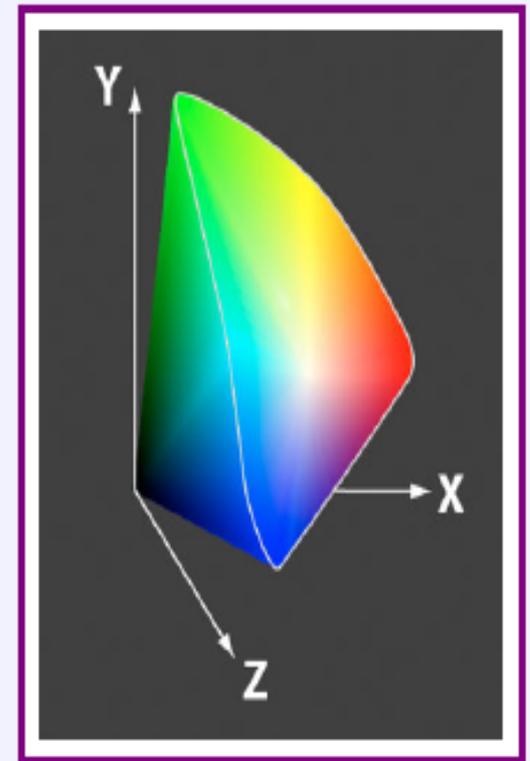
B

Espace XYZ

Définition

- ▶ Défini à partir d'une transformation linéaire sur l'espace (rgb) telle que toutes les couleurs du spectre visible soient contenues dans le triangle (xyz) par :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2.76 & 1.75 & 1.13 \\ 1.00 & 4.59 & 0.06 \\ 0.00 & 0.05 & 5.59 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$



Espace HSV

Espace qui permet de décomposer une couleur en trois composantes plus intuitives qui sont la teinte, la saturation et la luminance.

Définition

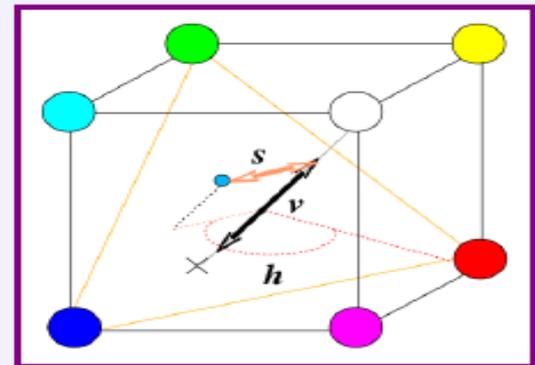
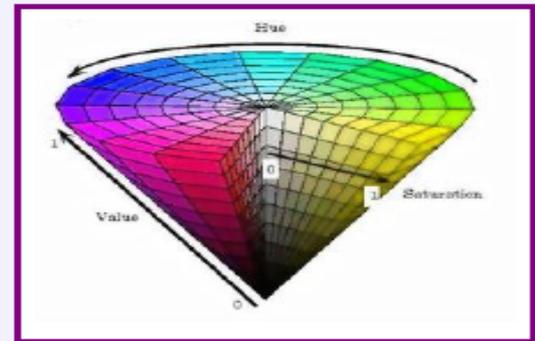
- ▶ Trois composantes : Hue (teinte), Saturation, Value (luminance)
- ▶ Passage RGB \leftrightarrow HSV :

$$V = \frac{R + G + B}{3}$$

$$H = \begin{cases} \theta & \text{si } B \leq G \\ 2\pi - \theta, & \text{sinon} \end{cases}$$

$$\theta = \arccos \left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} \right)$$

$$S = 1 - \frac{3\min(R, G, B)}{R + G + B}$$



Espace HSV

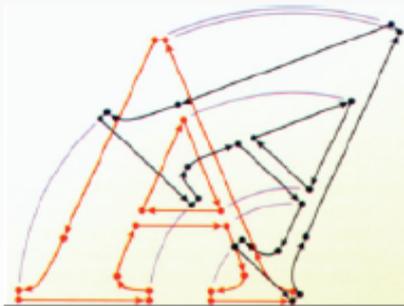
Exemple



Représentation matricielle/vectorielle

Image vectorielle

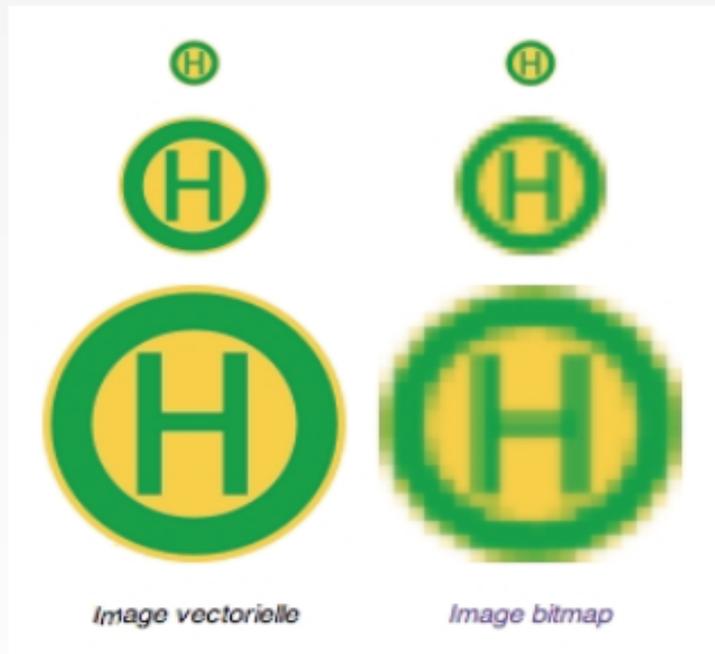
- ▶ Codage par des formes géométriques simples :
 - ▶ Polygone : coordonnées des sommets, couleur du contour, couleur de remplissage
 - ▶ Arc de cercle ...
- ▶ Exemples de formats : DXF, SVG, Illustrator
- ▶ Domaines : PAO (publication assistée par ordinateur), DAO (dessin assisté par ordinateur)...
- ▶ Avantages : peu de place en mémoire, transformations géométriques sans aliasing
- ▶ Inconvénients : l'affichage nécessite des calculs, la taille de l'image est liée à sa complexité



Représentation matricielle/vectorielle

Image matricielle

- ▶ Codage de l'image par une matrice de pixels
- ▶ Exemples de formats : BMP, PGM, JPEG
- ▶ Avantages : affichage direct
- ▶ Inconvénients : taille mémoire, transformations géométriques avec aliasing / perte d'information



Formats d'images matricielles

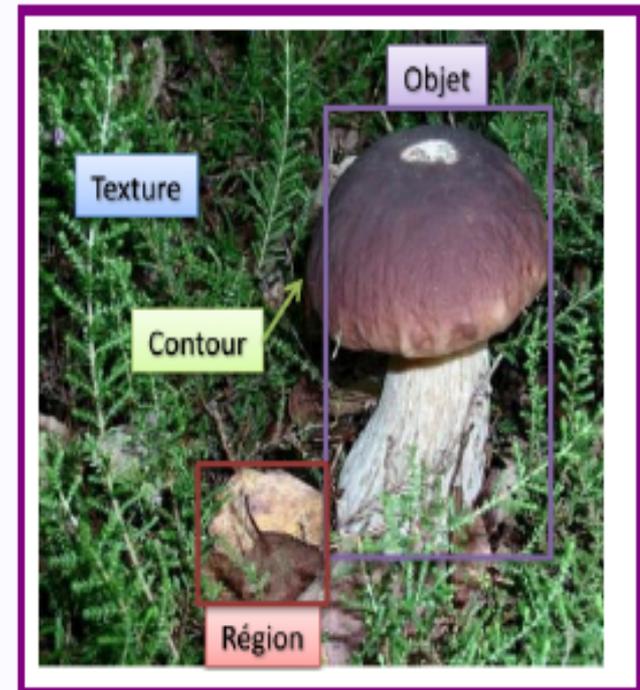
Comment stocker dans un fichier toutes les informations de l'image : taille de l'image, format des pixels, valeurs des pixels ?

Il existe une multitude de formats :

- ▶ **Les formats simples** : fichiers ASCII ou binaires comportant un entête contenant les dimensions de l'image et le format des pixels puis la liste des valeurs des pixels.
Exemple : les formats PNM (PBM - Portable BitMap, PGM - Portable Grayscale Map, PPM - Portable PixMap).
- ▶ **Les formats compressés** : l'information est codée de manière à réduire la taille du fichier image.
Exemple : les formats image du web (GIF , PNG, JPEG)

Quelques éléments simple d'une image

- ▶ **Texture** : répartition statistique ou géométrique des intensités dans l'image
- ▶ **Contour** : limite entre deux (ou un groupe de) pixels dont la différence de niveaux de gris (ou de couleurs) est significative
- ▶ **Région** : groupe de pixels présentant des caractéristiques similaires (intensité, mouvement, *etc.*)
- ▶ **Objet** : région (groupe de régions) entièrement délimitée par un contour, possédant une indépendance dans l'image
↳ description sémantique : on peut donner un nom à un objet



Propriétés des images : Qualité

Définition

- **Lignage** : phénomène d'alternance des lignes claires (ou sombres) de même direction et qui tranchent avec le reste de l'image
- **Contraste** : qualité de la dynamique des intensités de l'image
- **Bruit** : signal "parasite" dont la distribution dans l'image est aléatoire et la plupart du temps inconnue
- **Déformations géométriques** : défauts dus à la différence d'axe entre le capteur d'acquisition et le centre de la scène observée

Propriétés de l'image

Brillance

- Moyenne des niveaux de gris de l'image, ou intensité moyenne

$$B = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i,j)$$



Propriétés de l'image

Contraste

- ▶ Définition 1 : variation maximale entre valeurs de niveaux de gris min et max dans l'image :

$$C = \frac{\max_{i,j}[f(i,j)] - \min_{i,j}[f(i,j)]}{\max_{i,j}[f(i,j)] + \min_{i,j}[f(i,j)]}$$

- ▶ Définition 2 : écart-type des variations de niveaux de gris dans l'image :

$$C = \sqrt{\frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (f(i,j) - B)^2}$$

Attention, l'écart type traduit une moyenne de variation, ce qui est un peu différent

- ▶ Deux images totalement différentes peuvent avoir le même contraste

Exemples de calcul de contraste

