Vision par Ordinateur

William Puech

Plan

I) Introduction

II) Acquisition et formation de l'image

III) Modèles géométriques de caméras

IV) Mathématiques pour l'image

V) Traitements numériques de l'image

VI) Vision 3D

II) Acquisition et formation de l'image

A) Formation de l'image

- Energie lumineuse, radiométrie, photométrie, système de prise de vue.

B) La vision humaine

- Capteur œil, vision achromatique, vision 3D, perception du mouvement.

C) Colorimétrie

III) Modèles géométriques de caméras

A) Capteurs de vision

- Classification, effets photo-électriques, caméra à tubes, caméra à cibles solides (CCD), autres capteurs (IR, vidéo rapide).

B) Modèle géométrique d'une caméra

- Modélisation du système d'acquisition, les différents repères utilisés, projection perspective, transformation caméra/image, les paramètres externes, calibration liée aux paramètres internes.

IV) Mathématiques pour l'image

A) Modélisation mathématique de l'image

- Espace de modélisation, processus stochastique, représentation déterministe : opérateurs et transformées classiques, transformées 2D, transformée de Fourier.

B) Numérisation et échantillonnage

- échantillonnage spatial, effet de Moiré, numérisation de l'amplitude.

V) Traitements numériques de l'image

A) Amélioration et correction d'images
- Filtrage, convolution.

B) Prétraitements
- détection de contours, seuillage, morphologie mathématique.

C) Segmentation de l'image
- en contours, en droite (Hough), en régions.

VI) Vision 3D

A) Paramétrisation et reconnaissance

B) Image en mouvement

C) Vision 3D

D) Multi-échelle

E) Vision industrielle

F) Vision et multimédia

I) Introduction

A) L'IMAGE

B) Que peut-on faire d'une image ?

C) Vue d'ensemble de la spécialité

D) Les différents niveaux de représentation de l'image

E) Historique et applications

I) A) L'IMAGE

Image : représentation d'une personne ou d'une chose par la peinture, la sculpture, le dessin, la photo, le film (LAROUSSE)

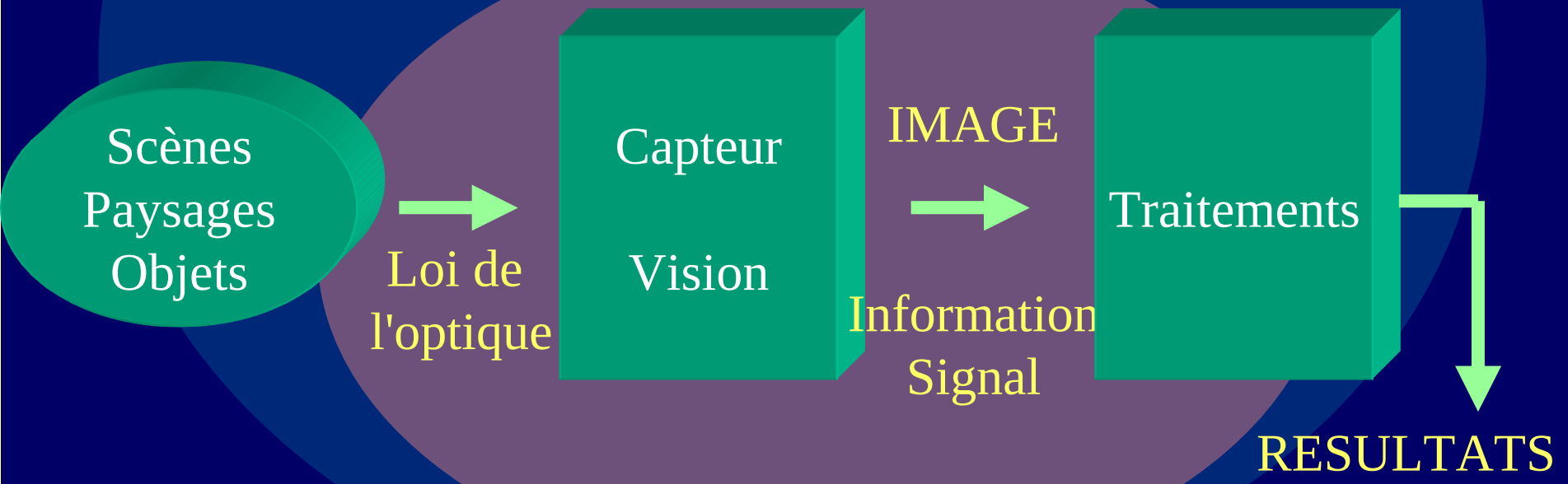
Vision : perception du monde extérieur par les organes de la vue.

I) A) L'IMAGE

- L'image est associée à la vision : représentation du monde extérieur.
- Composante subjective : chaque individu voit de manière différente la même scène.
- Les images : pas forcément un phénomène de vision précis (images mentales, rêves). Représentation d'objets immatériels ou concepts (signal électrique, oscillogramme)

I) A) L'IMAGE

- Dans ce cours : image = information issue d'un capteur de vision (œil, caméra).



I) A) L'IMAGE

Traitement : traitement + analyse + interprétation

- Traitement (prétraitement) opérations de manipulation de l'image pour améliorer la qualité.
  - la compression : réduction du volume de l'image.
  - la restauration : correction des défauts dus à une source de dégradation.
  - l'amélioration : modification de l'image dans le but de la rendre plus agréable à l'œil.
- Analyse : suite d'opérations pour l'extraction d'information contenue dans une image.
  - phase de segmentation avec partition de l'image.
  - techniques de description / modélisation pour obtenir la description structurelle de l'image.

I) A) L'IMAGE

- Interprétation : passage de la description structurelle à la description sémantique en regard à certains objectifs. (mesure de paramètres sur des formes, description du contenu de la scène en termes de concepts non mathématiques).
- En vision industrielle pour un objectif de qualité :
  - contrôle dimensionnel (mesures simples)
  - contrôle d'aspect (mesure de texture)
  - contrôle structurel (analyse des composants d'un objet)
  - tri (reconnaissance)

I) A) Mathématiquement

Image :

- forme discrète d'un phénomène continu.
- bidimensionnelle.
- L'information : caractéristique de l'intensité lumineuse (couleur ou niveaux de gris).

  $I : [0,L-1] \times [0,C-1] \Rightarrow [0,M]^p$: image de L lignes et C colonnes. Information dans un espace à p dimensions.

- image binaire $\Rightarrow (p,M) = (1,1)$
- image en niveaux de gris $\Rightarrow p = 1$ et $M = 255$
- image couleur $\Rightarrow p = 3$ et $M = 255$

I) A) Mathématiquement

Pixel : "picture element", unité de base de l'image correspondant à un pas de discrétisation.

- Position et valeur (niveaux de gris).
- Séquences vidéo du pixel :

I) A) Mathématiquement

Maillage : arrangement géométrique des pixels dans l'image.

3 types de tessélations du plan par des figures géométriques.

- Maillage carré : réalité physique du capteur CCD.
- Maillage hexagonal (référence en morphologie mathématique).
- Maillage triangulaire.

I) A) Mathématiquement

Distance : entre deux pixels P(xp,yp) et Q(xq,yq)

- distance de Manathan :

$$d_1(P,Q)=|xp - xq| + |yp - yq|$$

- distance euclidienne :

$$d_2(P,Q)=[(xp - xq)^2 + (yp - yq)^2]^{1/2}$$

- distance de l'échiquier :

$$d_{inf}(P,Q)=Max(|xp - xq| , |yp - yq|)$$

$$d_{inf}(P,Q) \leq d_2(P,Q) \leq d_1(P,Q)$$

I) A) Mathématiquement

Connexité : deux ordres de connexité : 4 et 8.

Un pixel a 4 voisins directs avec la distance d_1, et 8 avec la distance d_{inf}.

Histogramme : de l'image I, la fonction H définie sur l'ensemble des entiers naturels par:

$$H(x) = Card\{P : I(P) = x\}$$

- H(x) = nombre d'apparitions du niveau de gris x dans l'image I.
- outil privilégié en analyse d'images.

I) B) Que peut-on faire d'une image ?

1) Amélioration, Restauration, Correction

Traitement corrigeant des défauts de l'image, permettant un confort de visualisation.

- Augmentation de contraste,
- Correction des distorsions optiques,
- Filtrage du bruit

I) B) Que peut-on faire d'une image ?

2) Analyse

Conception d'une machine autonome pour conclure sur des critères visuels à partir d'images issues de capteurs.

- Reconnaissance d'objets,
- Localisation robot mobile,
- Mesures (spatiales, temporelles)

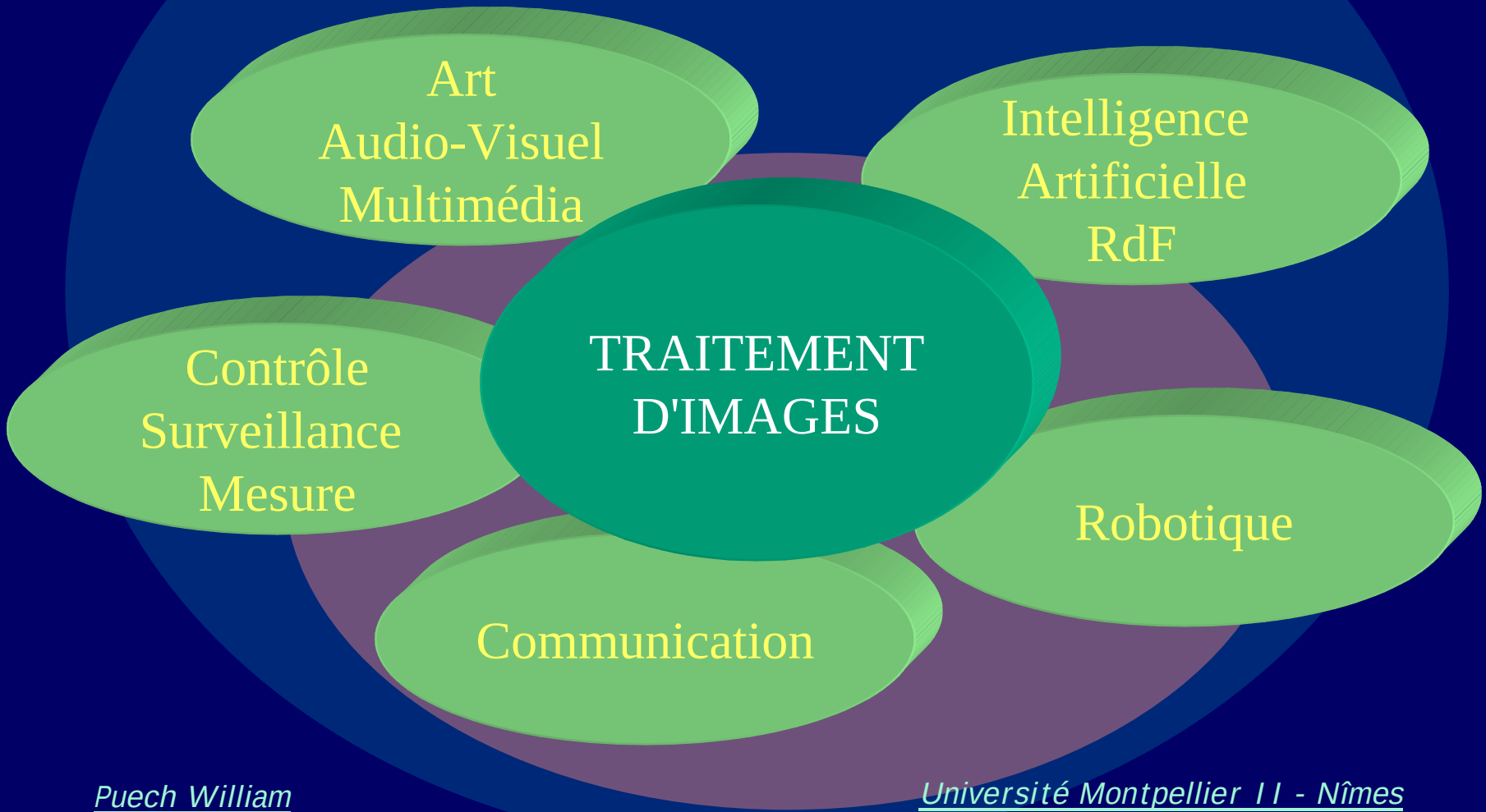
I) B) Que peut-on faire d'une image ?

3) Compression, Codage, Transmission

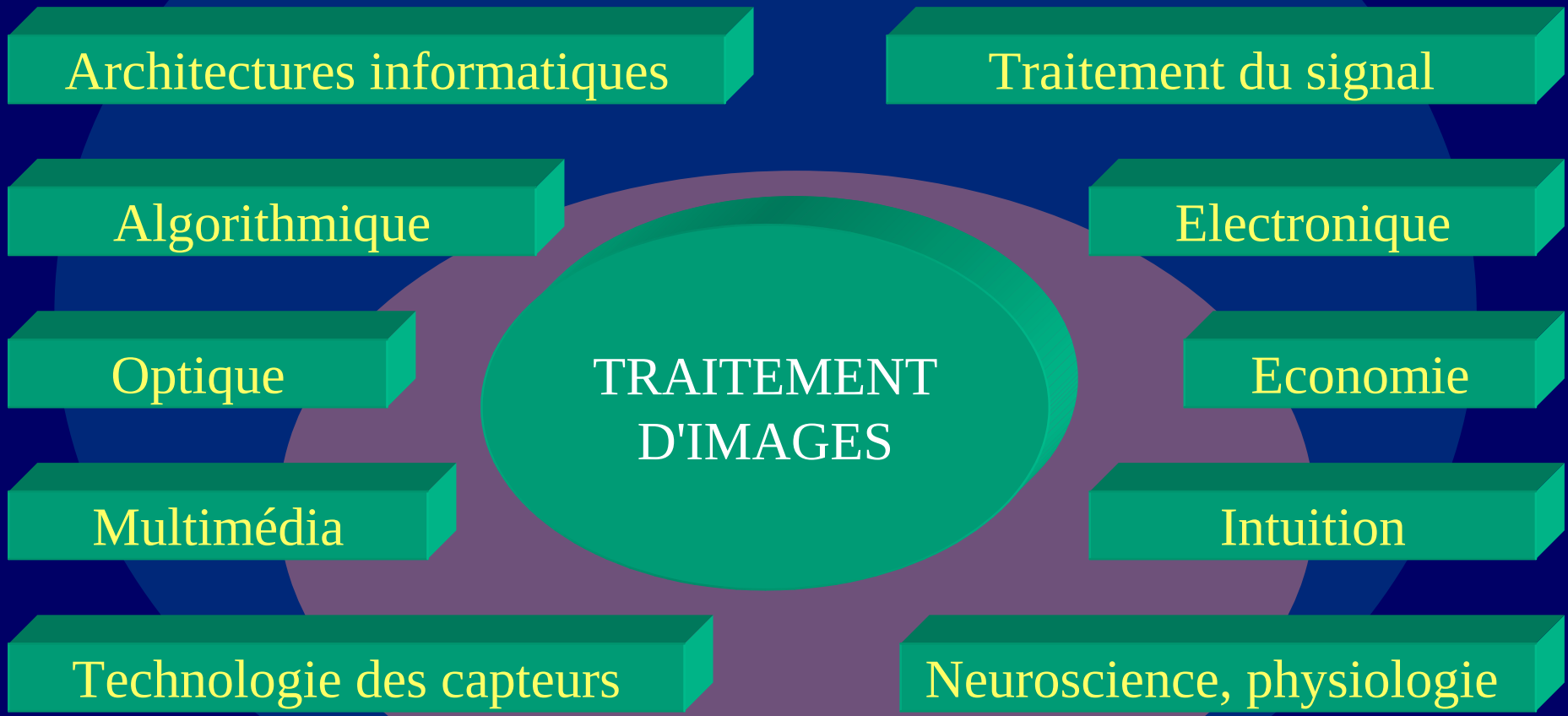
Transport des images par voies Hertziennes ou par réseau câblé sous forme numérique.

Codage nécessaire pour réduire l'information à transmettre sans en altérer la qualité.

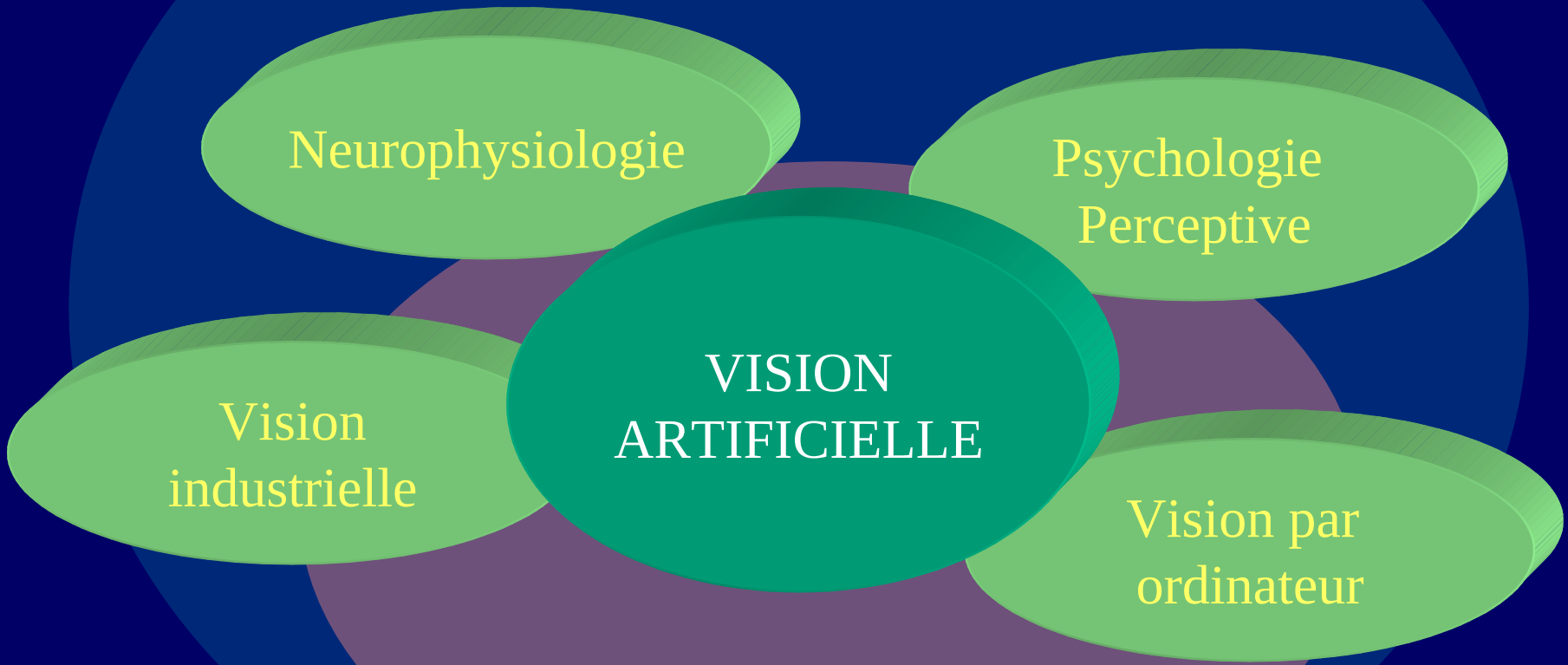
I) C) Vue d'ensemble de la spécialité



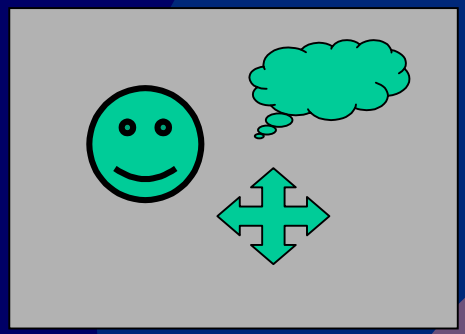
I) C) Vue d'ensemble de la spécialité

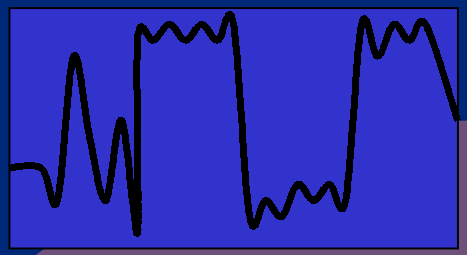


I) C) Vue d'ensemble de la spécialité



I) D) Les différents niveaux de représentation de l'image





45457878454545
45445454545478952322323245488765222

Bords,
Segments,
Lignes, ...

ANALOGIQUE OPTIQUE
10^{10} bits

ANALOGIQUE ELECTRIQUE
10^{8} bits

NUMERIQUE
10^{6} bits

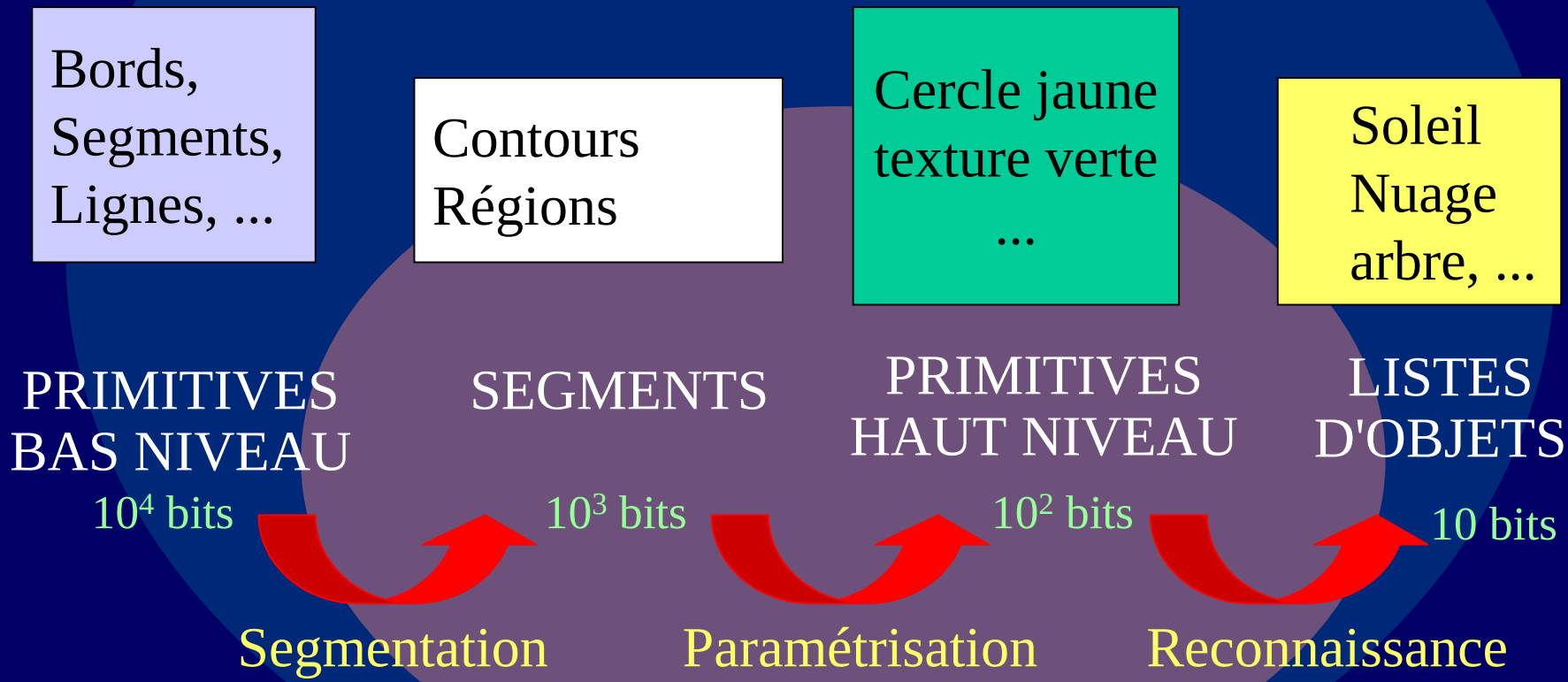
PRIMITIVES BAS NIVEAU
10^{4} bits

Trans. Optique
Capteurs

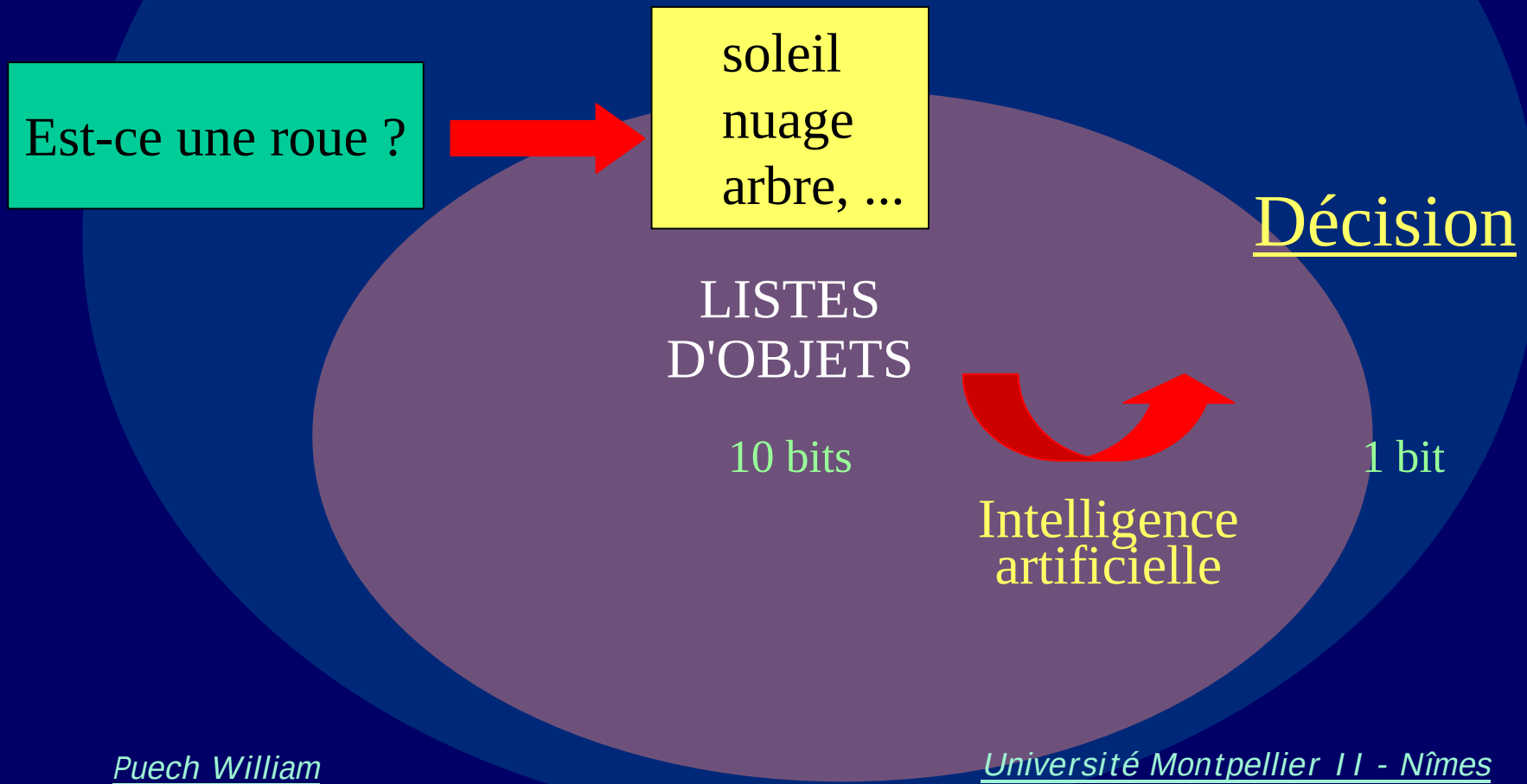
Electronique

Algorithmique
Trait. du signal

I) D) Les différents niveaux de représentation de l'image



I) D) Les différents niveaux de représentation de l'image



I) E) Historique et applications

- 1920 Transmission image par câble (New York - Londres) en quelques heures
- 1950 Origine du traitement d'images : analyses d'images dans les chambres à bulles (Rayons X, OCR, …)
- Images de mauvaise qualité et très volumineuses (700x500 pixels sur 8 bits par image)

I) E) Historique et applications

- 1960 Trois domaines dominants de traitements numériques d'images spatiales :
- Restauration (corriger les défauts liés à l'acquisition)
- Amélioration (rendre l'image "belle" pour l'affichage)
- Compression (réduire le volume de l'image)

I) E) Historique et applications

- 1970 extraction automatique d'informations.
- Apparition de la notion de description structurelle.
- Nouveaux thèmes : seuillage, segmentation, extraction de contours, morphologie mathématique.
- Interprétation d'images : engouement explosif avec l'apparition des systèmes experts. Puis déception car échec !

I) E) Historique et applications

- Les raisons de l'échec : pas d'expert, le savoir trop complexe pour être modélisé, et oubli du modèle perceptif
- 1980 Explosion du traitement d'images "industriel" - Micro-informatique + capteurs
- De l'image 2D aux modèles tri-dim.
- Analyse du mouvement, vision pour la robotique (mouvement, 3D, détection d'obstacle, trajectoire)

I) E) Historique et applications

- 1990 Explosion des transmissions de données avec Internet
- De la vision passive à la vision active (prise en compte de l'observateur dans l'analyse de la scène).
- Et maintenant...
- On ne sait pas faire grand chose, mais au moins maintenant, on le sait !!

I) E) Historique et applications

- les bibliothèques numériques : acquisition (du papier ou de la vidéo vers le numérique)
- Représentation (quel codage ?), Transmission (codage et réseaux), Accès (Indexation/Recherche)
- Une nouvelle tentative pour l'interprétation (il ne s'agit plus d'interpréter pour interpréter mais d'interpréter pour rechercher)

I) E) Historique et applications

- Imagerie aérienne et spatiale
  - Ressources naturelles et humaines,
  - Surveillance,
  - Météorologie.
- Industrie
  - Contrôle non destructif,
  - Inspection et mesures automatiques,
  - Vision robotique.

I) E) Historique et applications

- Médecine
  - Cytologie,
  - Tomographie,
  - Echographie.
- Sciences
  - Interventions en milieu confiné,
  - Astronomie, Robotique mobile,
  - Microscopie électronique, Biologie.

I) E) Historique et applications

- Art et communication
  - Télévision et vidéo,
  - Photographie, Edition,
  - Transport information visuelle, Archivage.
- Domaine militaire
  - Surveillance,
  - Guidage automatique et poursuite d'engins,
  - Topographie.

Les dix commandements de la vision par ordinateur

(J.P. Hermann, DTAA Renault)

1. Tu respecteras les **lois de la physique** (la reconnaissance des formes n'est pas un problème d'informatique mais d'optique).

2. Tu t'inquièteras de savoir ce que tu mesures avant de te demander ce que tu vas calculer.

3. Tu chercheras avant t

vitesse de calcul de ton ordinateur (importance de l

5. Tu honoreras l'**algori** nmation.

6. Tu te soucieras de l'**a** e des formes.

7. Tu effectueras les ess ur l'éclairage.

8. Tu mettras des **chiffr** définition des seuils de tolérance).

9. Tu auras la **simplicité** pour idéal.

10. Tu ne croiras pas seulement à la caméra vidéo (il existe d'**autres capteurs** optiques ... et non optiques).

et enfin, le onzième commandement : Tu ne te décourageras pas!

Puech William *Université Montpellier II - Nîmes*

II) Acquisition et formation de l'image

A) Formation de l'image

- Energie lumineuse, radiométrie, photométrie, système de prise de vue.

B) La vision humaine

- Capteur œil, vision achromatique, vision 3D, perception du mouvement.

C) Colorimétrie

I I) A) Formation de l'image

1) Energie lumineuse

En optique, une image = une quantité d'information véhiculée par des ondes électromagnétiques.

Longueur d'onde et énergie

Ondes lumineuses = émission d'énergie sous forme de photons due aux transitions atomiques de corps chauffés.

I I) A) Formation de l'image

1) Energie lumineuse- Loi de Planck

Un corps noir chauffé à une temp. T émet une puissance rayonnante P ($W.m^{-2}$) :

$$P(\lambda) = C_1 / \left(\lambda^5 \left(\exp(C_2 / \lambda T) - 1\right)\right)$$

avec $C_1 = 2c^2h$, $C_2 = c\,h/k$

c = vitesse de la lumière = $3\ 10^8$ $m.s^{-1}$, h = constante de Planck = $6.62\ 10^{-34}$ J.s

k = constante de Bolzmann = $1.38\ 10^{-23}$ $j.K^{-1}$

T = temp. en kelvin,

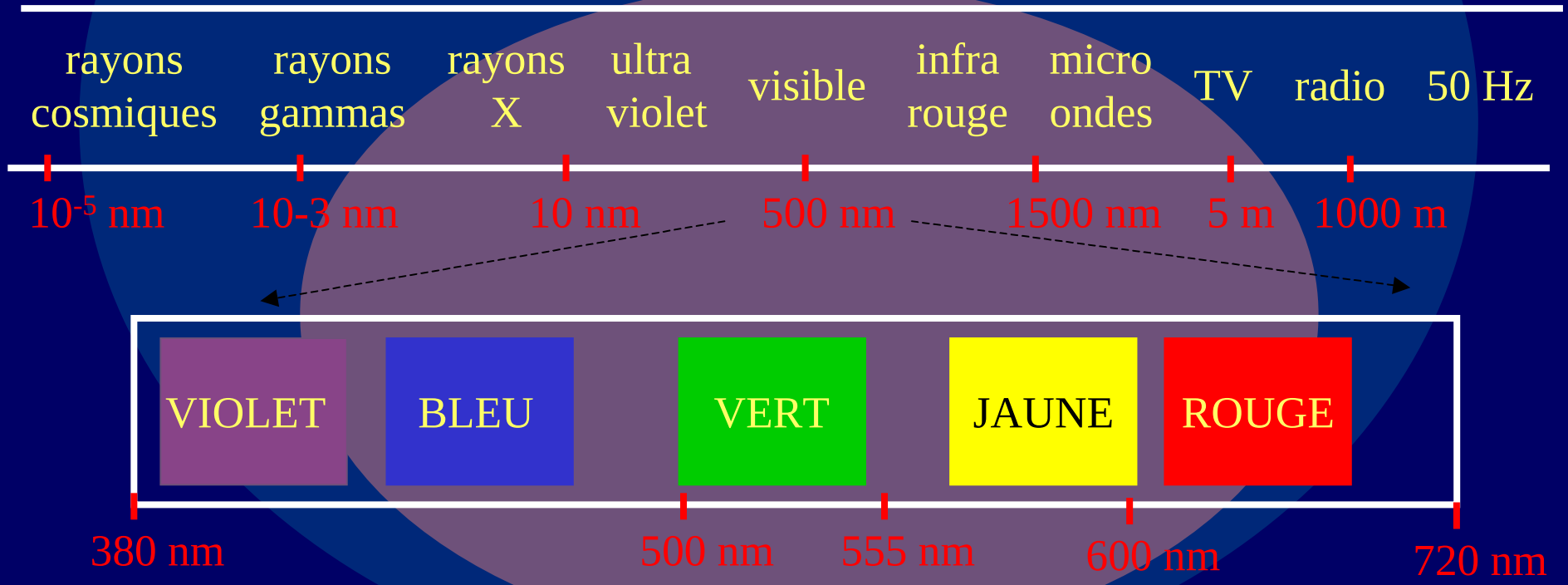
λ = longueur d'onde en m^{-1}

I I) A) Formation de l'image

1) Energie lumineuse- Classification fréquentielle des ondes lumineuses
   - Lumière visible : détectée par l'oeil
   - Lumière chromatique : composée de plusieurs longueurs d'onde
   - Lumière monochromatique : une seule longueur d'onde (LASER)
   - Lumière achromatique : seule l'énergie est prise en compte.

I I) A) Formation de l'image

1) Energie lumineuse- Classification fréquentielle des ondes lumineuses



I I) A) Formation de l'image

1) Energie lumineuse- Température de couleur

– Loi de Wien : un corps chauffé émet un spectre de lumière.

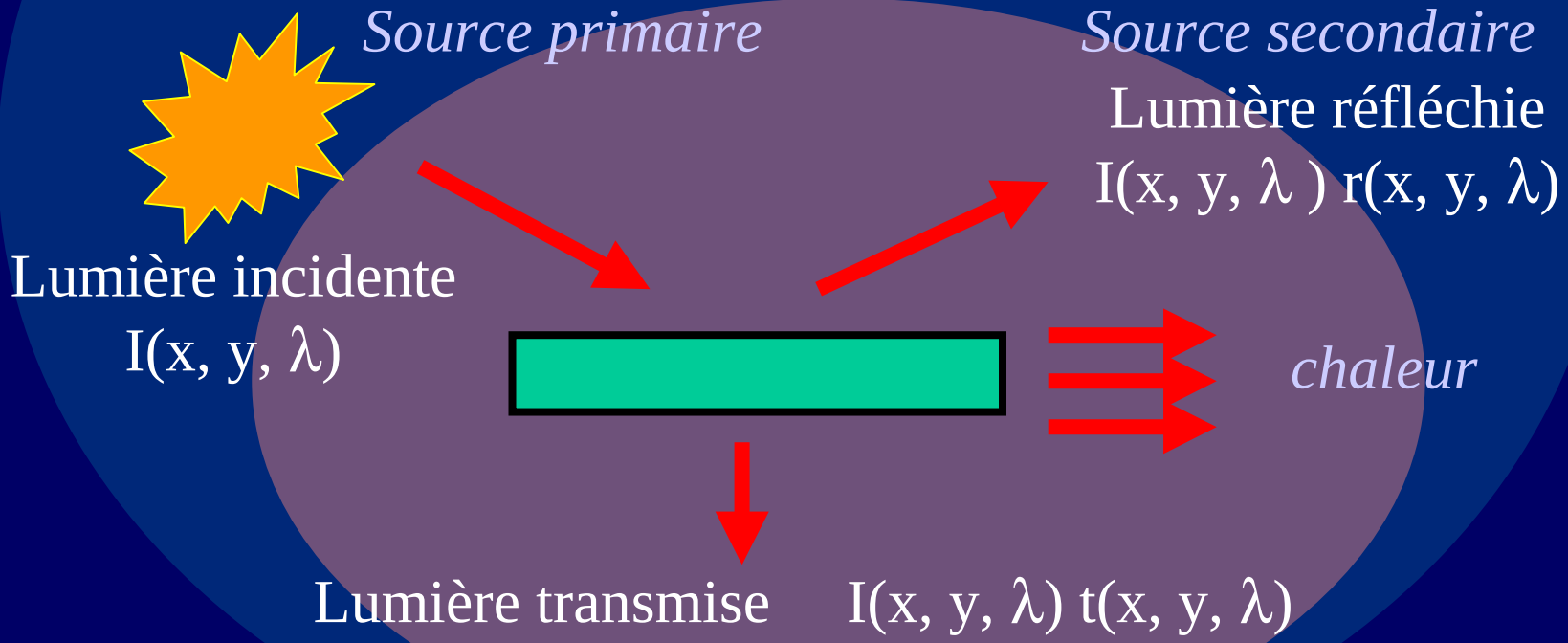
- Flamme bougie 1900 K
- Lampe à incandescence 2700 K
- Soleil 6000 K (blanc parfait)
- Tube cathodique 7000 K

– Si T ↗ le spectre → longueur d ’onde courte.

I I) A) Formation de l'image

2) Radiométrie

Vision d ’un objet = interaction avec une source lumineuse



I I) A) Formation de l'image

2) Radiométrie : réflexion et transmission

r(x,y) = 0.01	velours noir
r(x,y) = 0.8	blanc mat
r(x,y) = 0.93	neige fraîche
r(x,y) = 1	miroir

t(x,y) = 0	objet opaque
t(x,y) = 1	vitre

Réflexion maximale pour λ correspondante (vert, rouge)

I I) A) Formation de l'image

2) Radiométrie : loi de Lambert

"La quantité d'énergie émise à partir d'un élément de surface dans une direction déterminée est proportionnelle au cosinus de l'angle que fait cette direction avec la normale à la surface"

$$dW_\theta = dW_n \cos(\theta)$$

Vraie pour les corps noirs.

Pas valable pour les surfaces brillantes.

I I) A) Formation de l'image

2) Radiométrie : unités radiométriques

Source lumineuse ponctuelle

- Flux énergétique (W)

$$\Phi = \frac{dW}{dt}$$

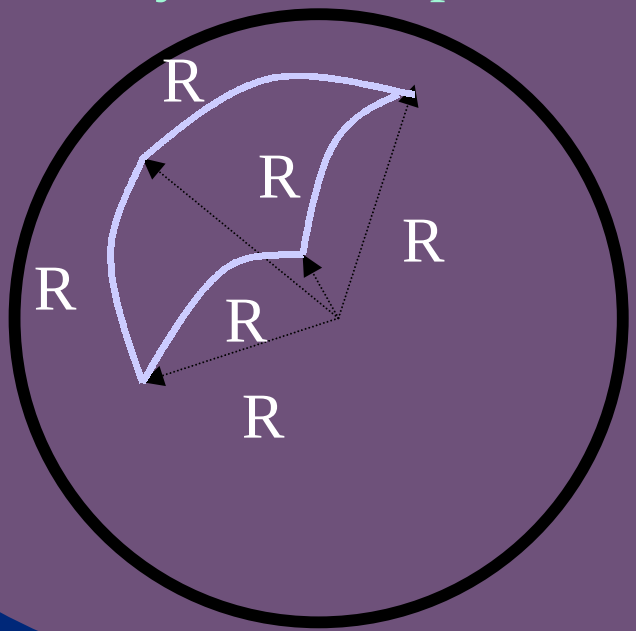
- Intensité énergétique ($W.sr^{-1}$)

$$I = \frac{d\Phi}{d\Omega}$$

I I) A) Formation de l'image

2) Radiométrie : le stéradian

"Angle solide qui ayant son sommet au centre d'une sphère, découpe sur la surface de cette sphère une aire égale à celle d'un carré ayant pour coté le rayon de la sphère."



I I) A) Formation de l'image

2) Unités radiométriques : Surface Lambertienne

- Eclairement énergétique

$(W.m^{-2})$

$$E = \frac{d\Phi}{dS}$$

- Emittance énergétique

$(W.m^{-2})$

$$M = \frac{d\Phi}{dS}$$

- Luminance énergétique

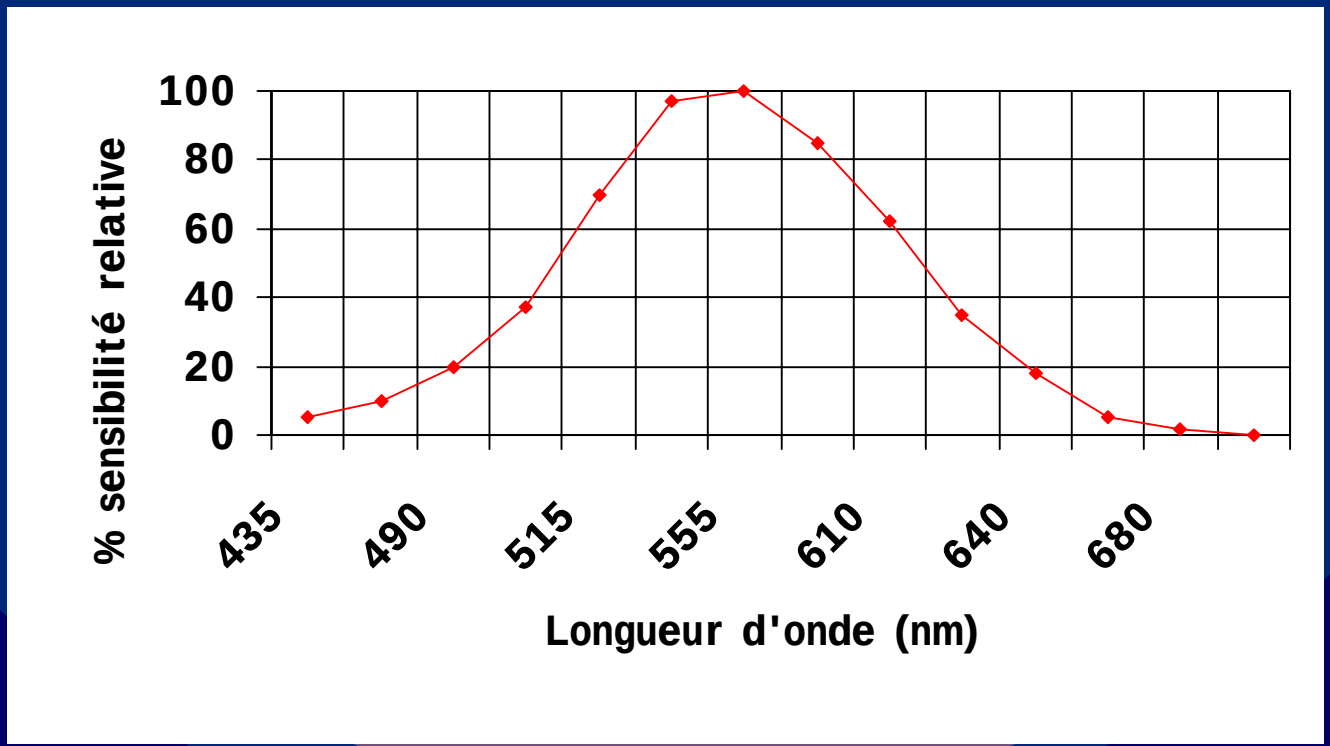
$(W.m^{-2}.sr^{-1})$

$$L = \frac{d^2\Phi_\theta}{d\Omega dS\cos(\theta)}$$

I I) A) Formation de l'image

3) Photométrie

– Sensibilité spectrale de l'œil humain, standard CIE (Commission Internationale de l'Eclairage)



I I) A) Formation de l'image

3) Unités photométriques

Le candela : "Intensité lumineuse dans une direction donnée d'une source qui émet un rayonnement monochromatique de 555 nm et dont l'IE dans cette direction est 1/683 $W.sr^{-1}$"

Radiométrie		Photométrie
Flux éner.	W	Flux lumineux Lumen (L)
Intensité éner.	$W.sr^{-1}$	Intensité lumineuse Candela (cd)
Eclairement éner.	$W.m^{-2}$	Eclairement Lux (=Lm. m^{-2})
Luminance éner.	$W .m^{-2}.sr^{-1}$	Luminance nit (=cd. m^{-2})

I I) A) Formation de l'image

3) Unités photométriques :

Efficacité lumineuse

$$k = \frac{\text{flux lumineux}}{\text{flux énergétique}}$$

à 555 nm, k = 683 pour le soleil, k =250

I I) A) Formation de l'image

3) Quelques valeurs

nuit sombre	10^{-4} Lux	minimum visible	10^{-5} nit
ciel étoilé	10^{-3} Lux	vert luisant	50 nits
pleine lune	10^{-1} Lux	flamme	15 10^{3} nits
norme couloir	>50 Lux	papier blanc soleil	30 10^{3} nits
norme salle de lecture	>300 Lux		
jour ciel ouvert	10^{3} Lux	arc électrique	1.5 10^{8} nits
table d'opération	10^{5} Lux	soleil	1.5 10^{9} nits

I I) A) Formation de l'image

3) Photométrie

La luminance est constante quelque soit l'angle d'observation.

L'émittance d'une surface lambertienne est égale au produit de sa luminance par π .

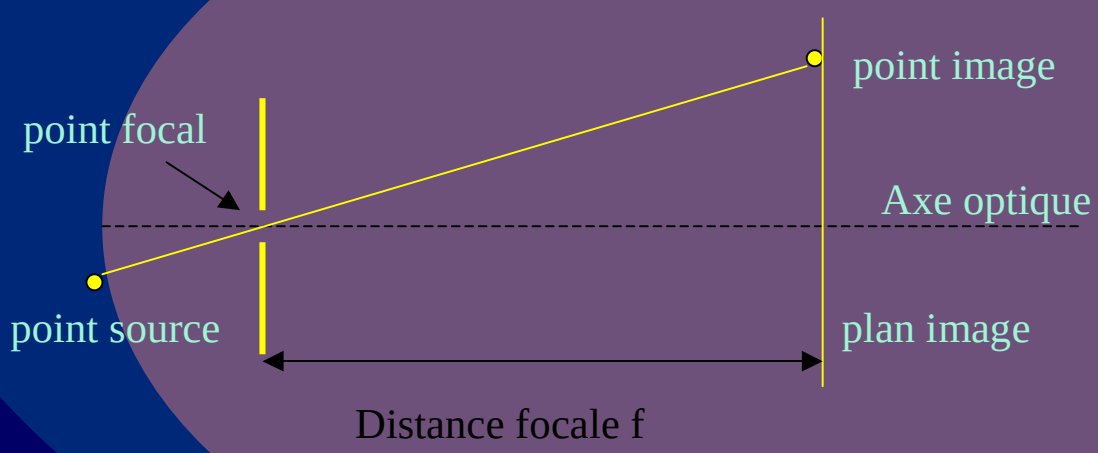
$$M = \pi L$$

I I) A) Formation de l'image

4) Systèmes de prise de vue

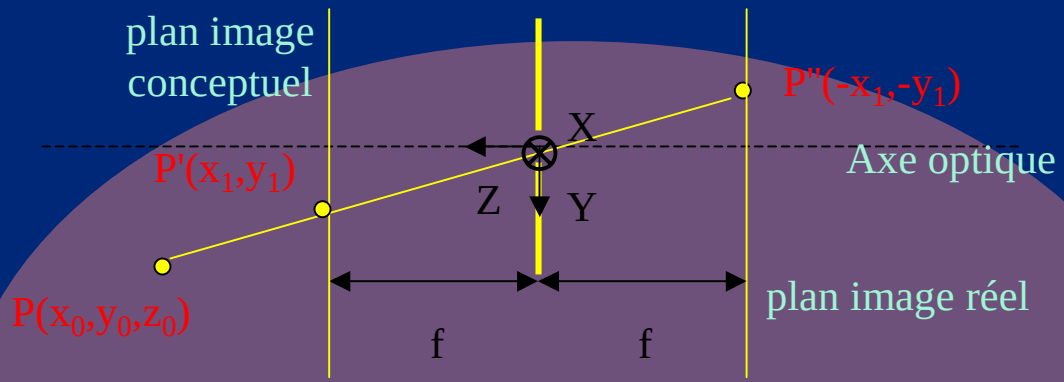
Une scène 3D doit être représentée sur un support 2D (film, CCD, …)

4.1 Appareil à sténopé (trou d'épingle)



I I) A) Formation de l'image

4.1 Appareil à sténopé : système réel et conceptuel

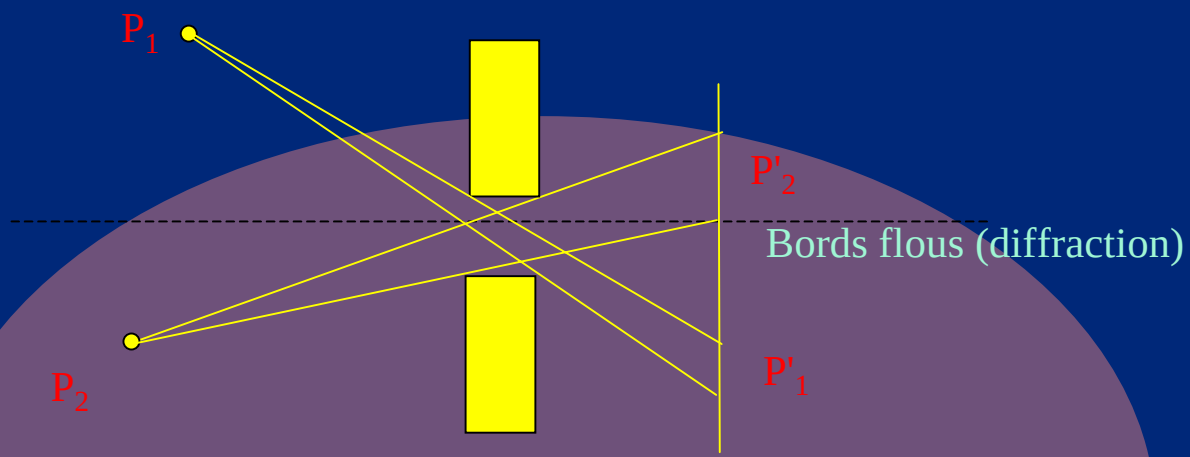


$$x_1 = f\frac{x_0}{z_0}$$

$$y_1 = f\frac{y_0}{z_0}$$

I I) A) Formation de l'image

4.1 Appareil à sténopé : inconvénients

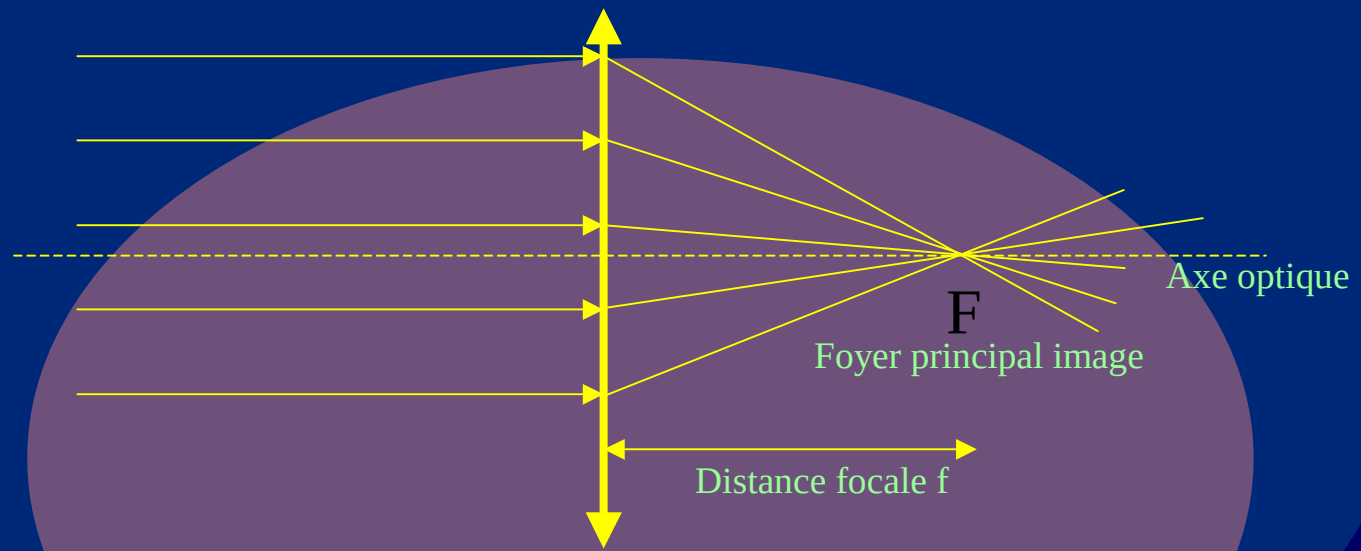


sensible : épaisseur de la plaque et aux positions des sources lumineuses.

I I) A) Formation de l'image

4.2 Systèmes optiques à lentilles

lentille convergente

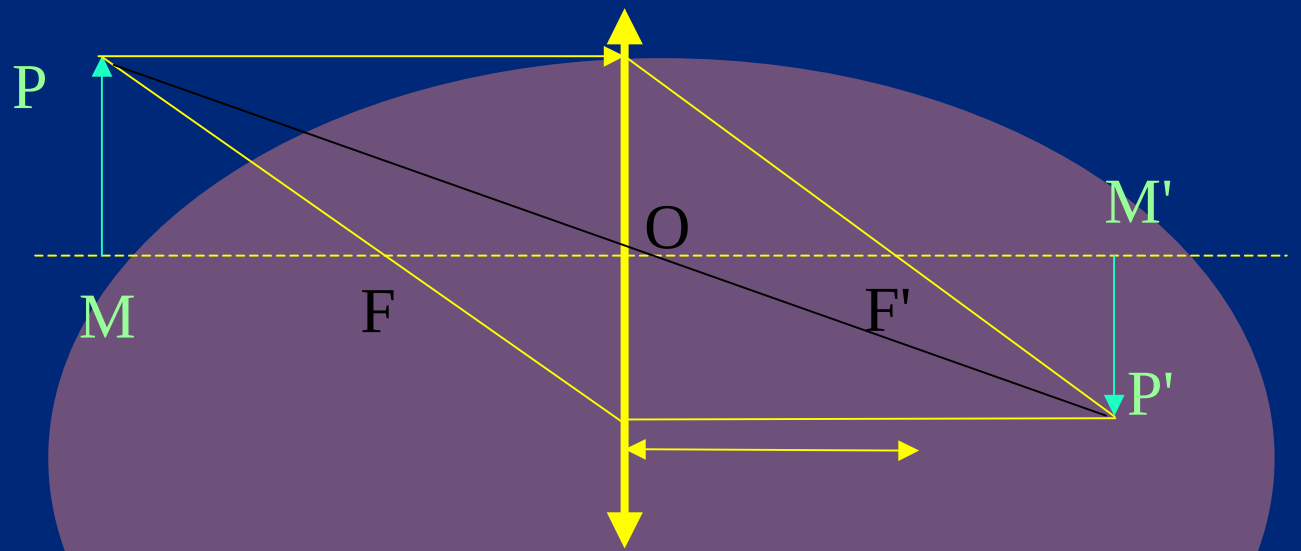


vergence = 1/f

I I) A) Formation de l'image

4.2 Formules de DESCARTES

lentille convergente



$$\frac{1}{OM}+\frac{1}{OM'}=\frac{1}{OF}=\frac{1}{OF'}$$

$$\frac{M'P'}{MP}=\frac{OM'}{OM}$$

I I) A) Formation de l'image

4.2 Aberrations optiques

• aberration sphérique : les rayons d'un point ne convergent pas tous en un seul point : **flou** .

Si taille lentille ↗ alors flou ↗ : diaphragme

• astigmatisme : trajet différent en fonction des axes vertical et horizontal de la lentille.

• Le défaut de coma (forme en goutte d'eau) : grossissement différent par rapport à l'écart à l'axe.

I I) A) Formation de l'image

4.2 distorsions géométriques

- liées à la qualité de l'optique
  - objectif grand angle
  - balayage d'une caméra
- distorsion tonneau
- distorsion coussin
- pour diminuer ces distorsions
  - ↗ le prix
  - correction par traitement d'images

I I)B) La vision humaine

1. Le capteur œil
2. La vision achromatique
3. Autres caractéristiques
4. Système de traitement visuel humain

I I)B) La vision humaine

1. Le capteur œil

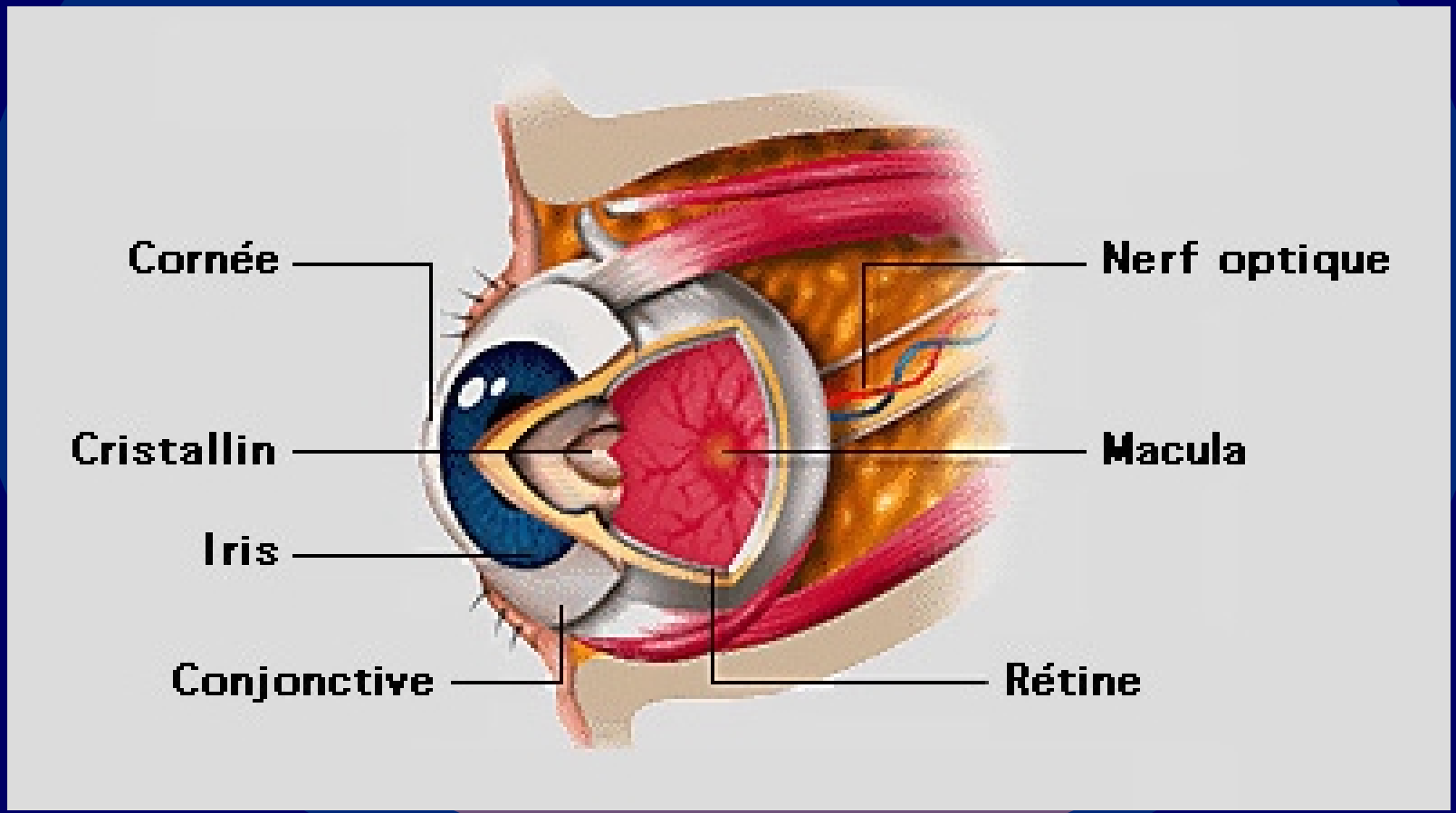
De forme approximativement sphérique, l'oeil est l'organe de base de la vision. Il comporte un ensemble d'éléments destinés à recevoir le rayonnement incident, former l'image des objets perçus et traiter les informations recueillies.

I I)B) La vision humaine

1. Le capteur œil : composants principaux :
   - cornée : protection filtre
   - Iris : diaphragme (variation d'un facteur 10 en surface). Son ouverture centrale est la pupille.
   - Cristallin : optique + focus (déformable, indice optique variable) : lentille à focale variable
   - Rétine : couche photo-sensible (120 millions de récepteurs : cônes et bâtonnets)
   - Macula : contient en son centre une petite dépression, la fovéa. zone d'acuité maximum de l'oeil.
   - Nerf optique : transport de l'information (100000 neurones)

I I)B) La vision humaine

1. Le capteur œil : composants principaux :



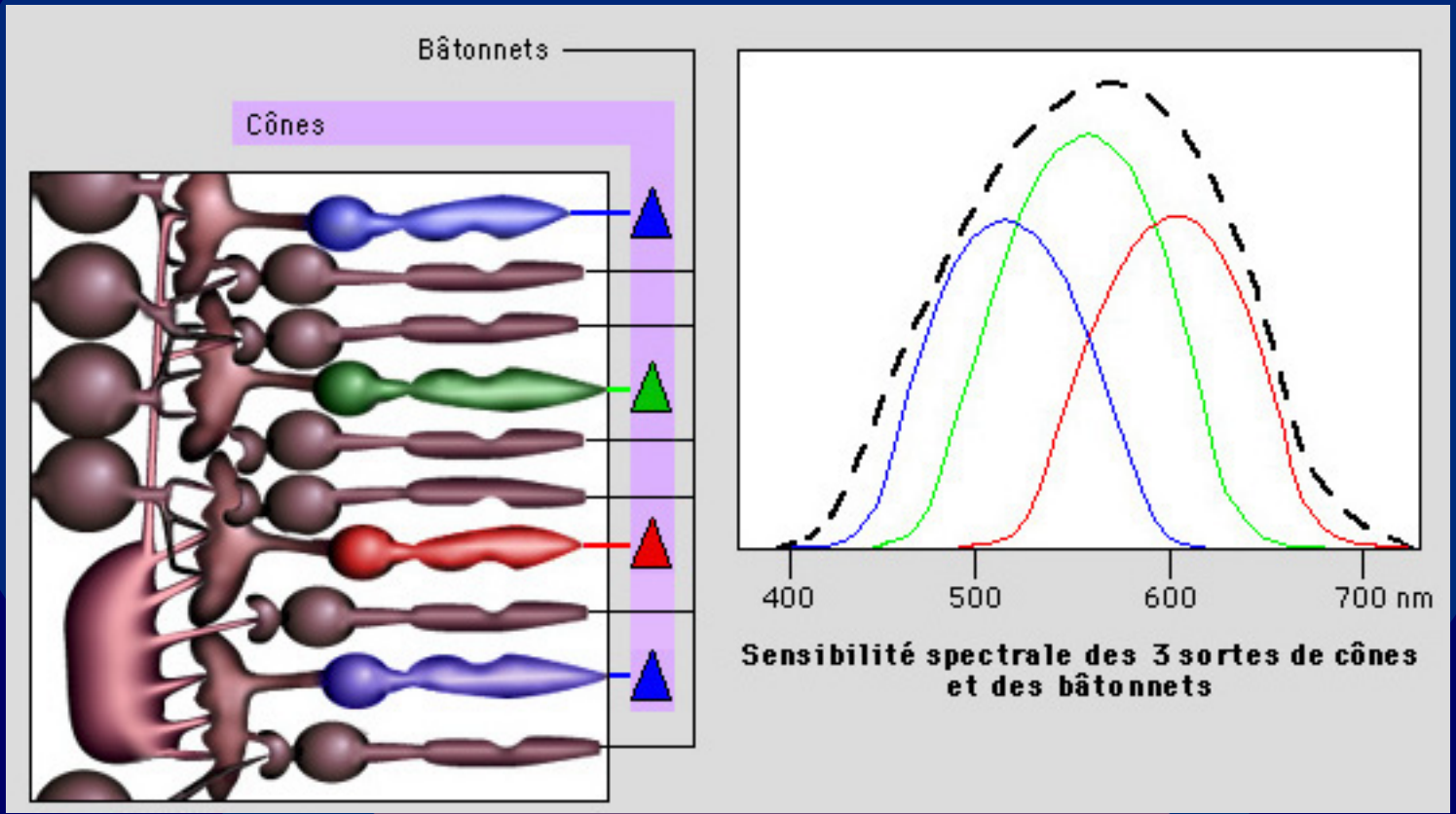
I I)B) La vision humaine

1. Les photos-récepteurs de la rétine

- Cônes :
  - vision photopique (diurne)
  - couleur
  - 6 à 7 millions
  - Zone fovéale
- Bâtonnets :
  - vision scotopique (nocturne)
  - Faible intensité, achromatique
  - 120 millions
  - Zone extra fovéale

I I)B) La vision humaine

1. Les photos-récepteurs de la rétine



I I)B) La vision humaine

1. La fovéa : région où la vision photopique est la plus précise et la plus sensible (angle de 20 minutes).

Le maximum de sensibilité en vision scotopique est a 20° de l'axe optique

La zone aveugle : rattachement du nerf optique sur la rétine.

Le capteur œil et le système à lentille

I I)B) La vision humaine

2 . La vision achromatique

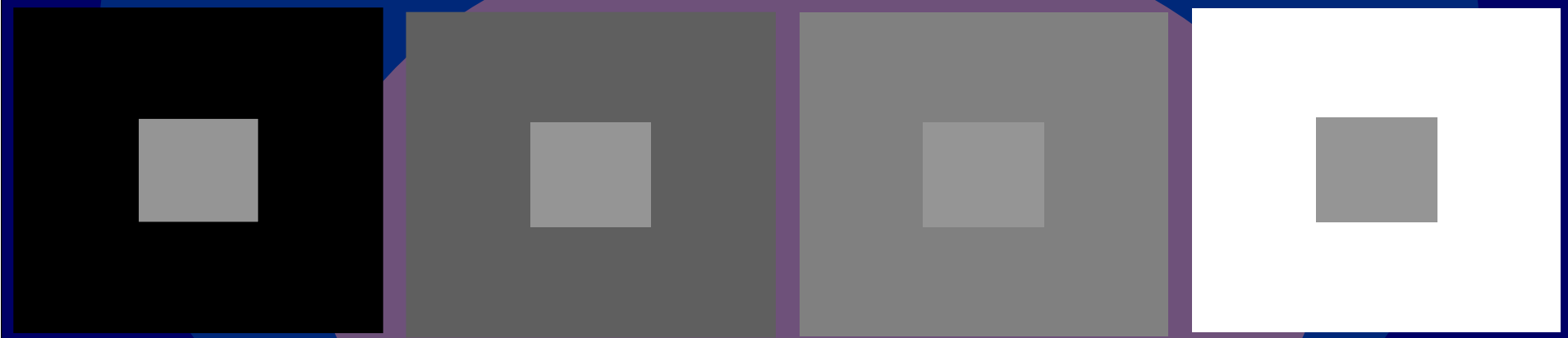
Caractéristiques statiques et dynamiques

- Non linéaire.
- Forte adaptation au niveau de luminance (échelle de 10^{10}).
- Transition scotopique-photopique graduelle
- Discrimination de luminance : de 50 à 100 niveaux
- Constante de temps d'adaptation pour les cônes et les bâtonnets.

I I)B) La vision humaine

2 . La vision achromatique

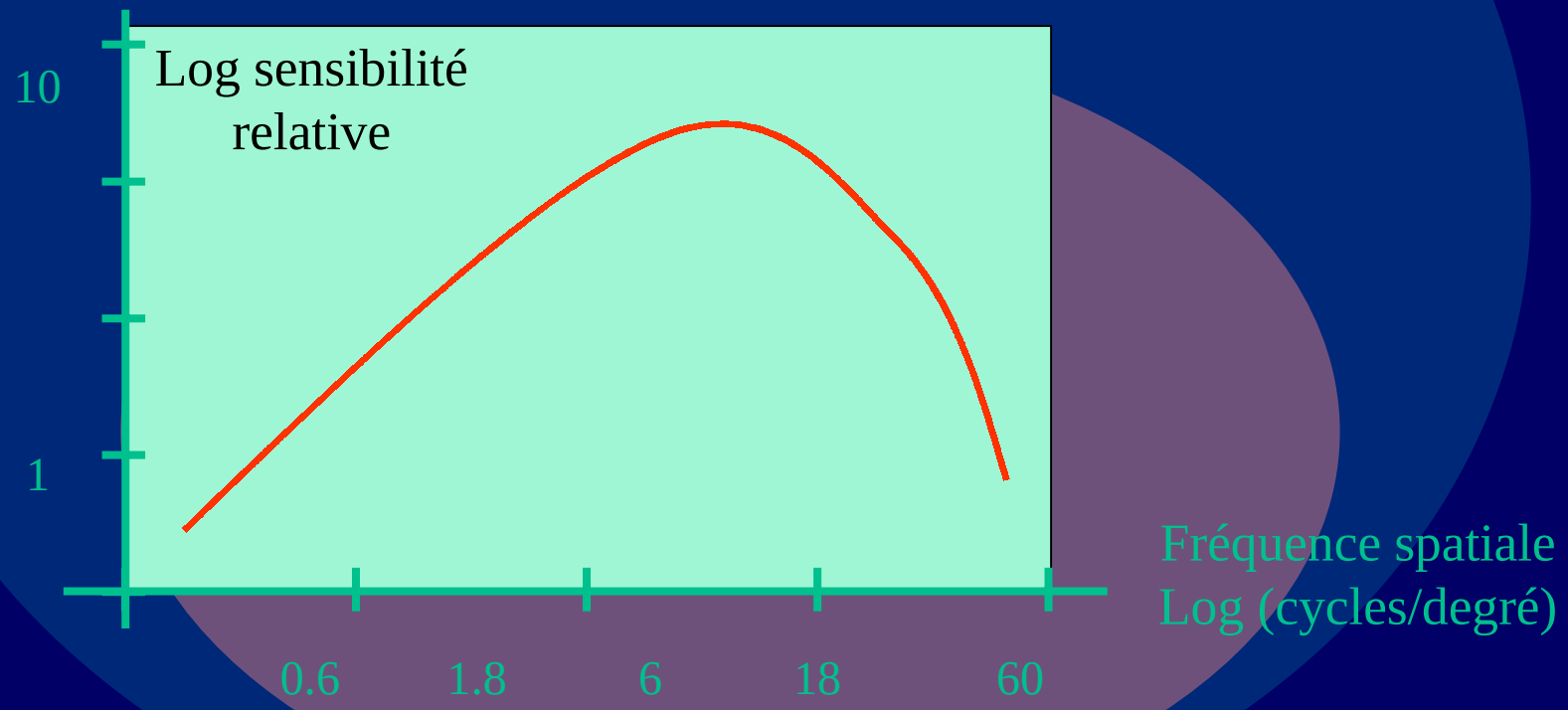
Seuil différentiel de luminance



I I)B) La vision humaine

2 . La vision achromatique

Réponse en fréquence spatiale

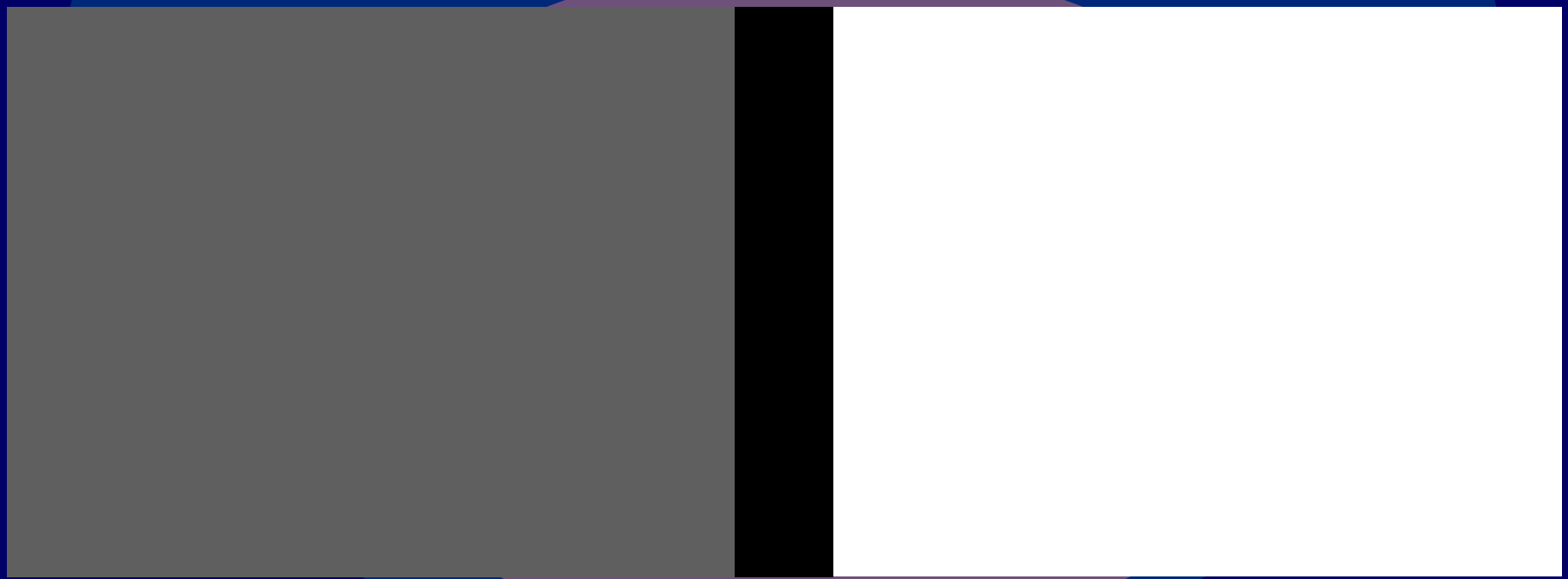


I I)B) La vision humaine

2 . La vision achromatique

Rehaussement physiologique des contours

transition en luminance : objectif ≠ subjectif



I I)B) La vision humaine

3 . Autres caractéristiques de la vision

Vision chromatique

Vision tridimensionnelle

Perception du mouvement

I I)B) La vision humaine

4 . Système de traitement visuel humain

Performance

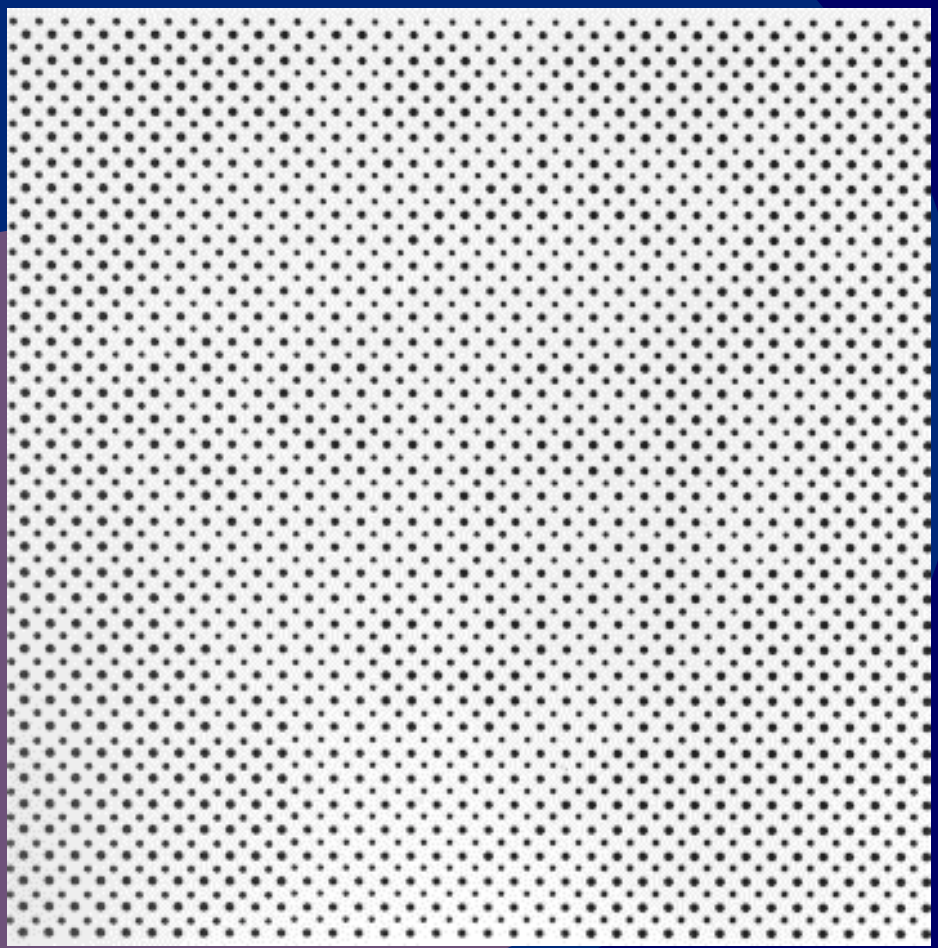
Architecture

Illusions optiques

I I)B) La vision humaine

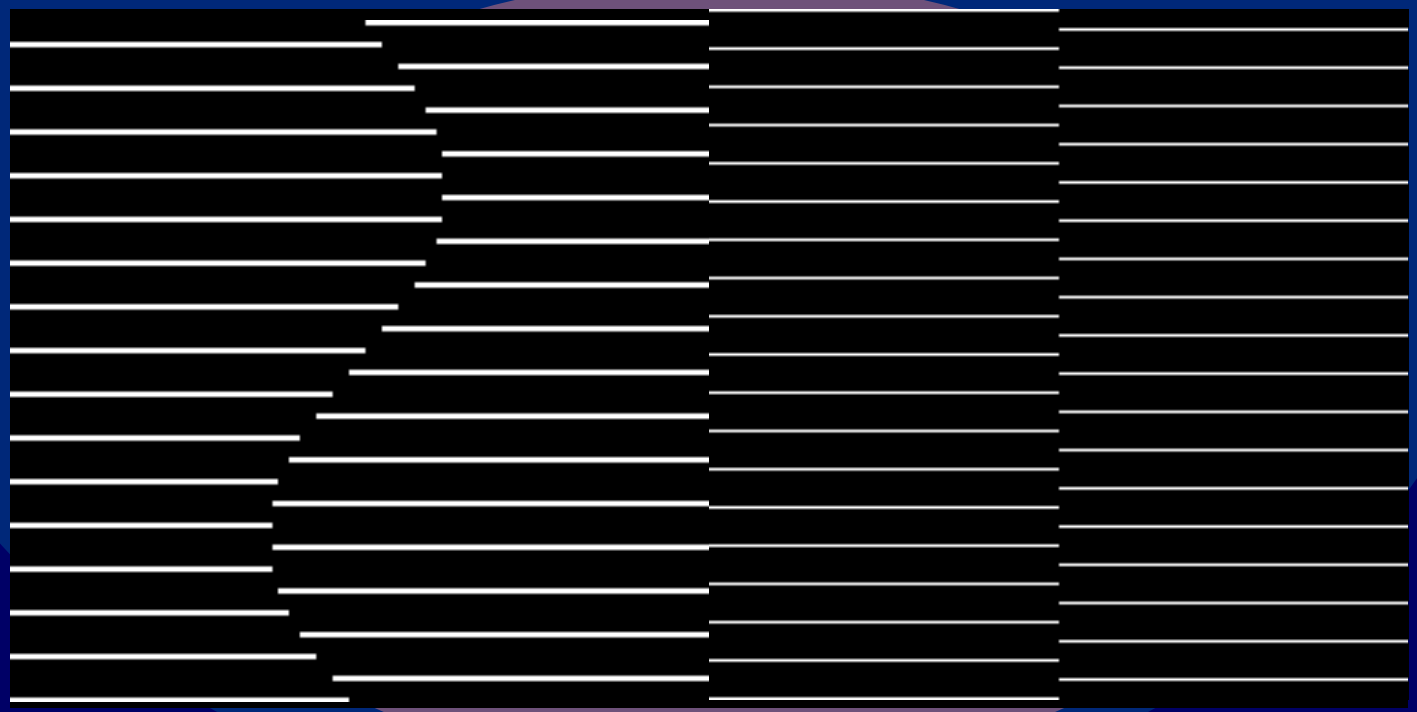
Illusions optiques

système visuel humain = référence. Pas un système parfait et piégeable : Que voyez-vous ? Rien, alors reculez-vous un peu ...



I I)B) La vision humaine

Les frontières qui n'existent pas...Le système visuel humain fait des "interpolations" à partir du stimulus perçu. L'exemple le plus marquant est celui des contours illusoires.



I I)B) La vision humaine

Les objets qui n'existent pas…

Le système visuel humain fait des "interpolations" symbolique car il préfère une information structurée à une information non structurée.

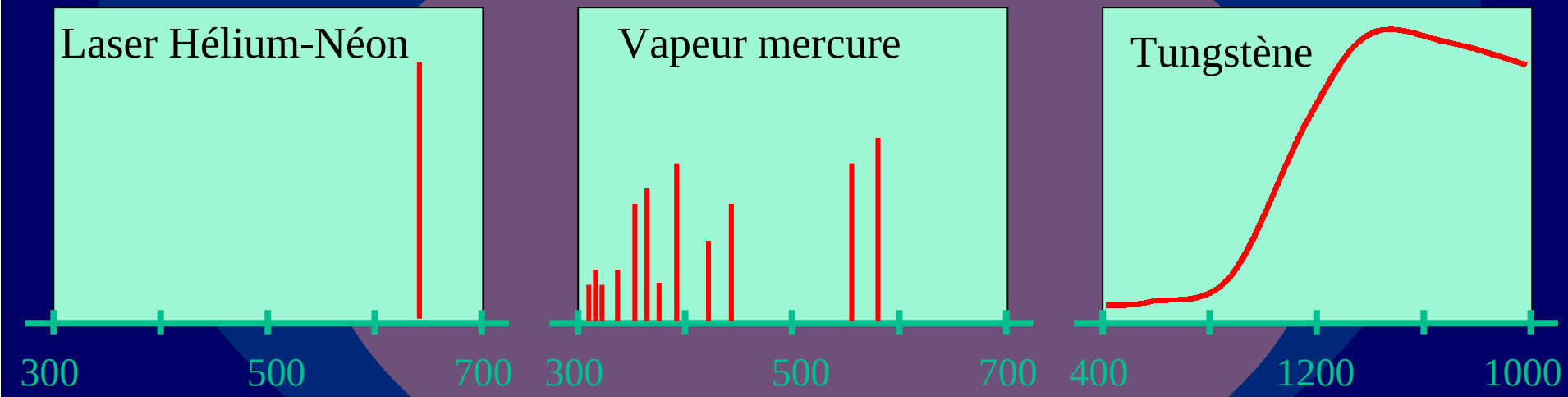
Que voyez-vous ? Rien, un mouton, un chien, un dalmatien peut-être ?

I I)C) Colorimétrie

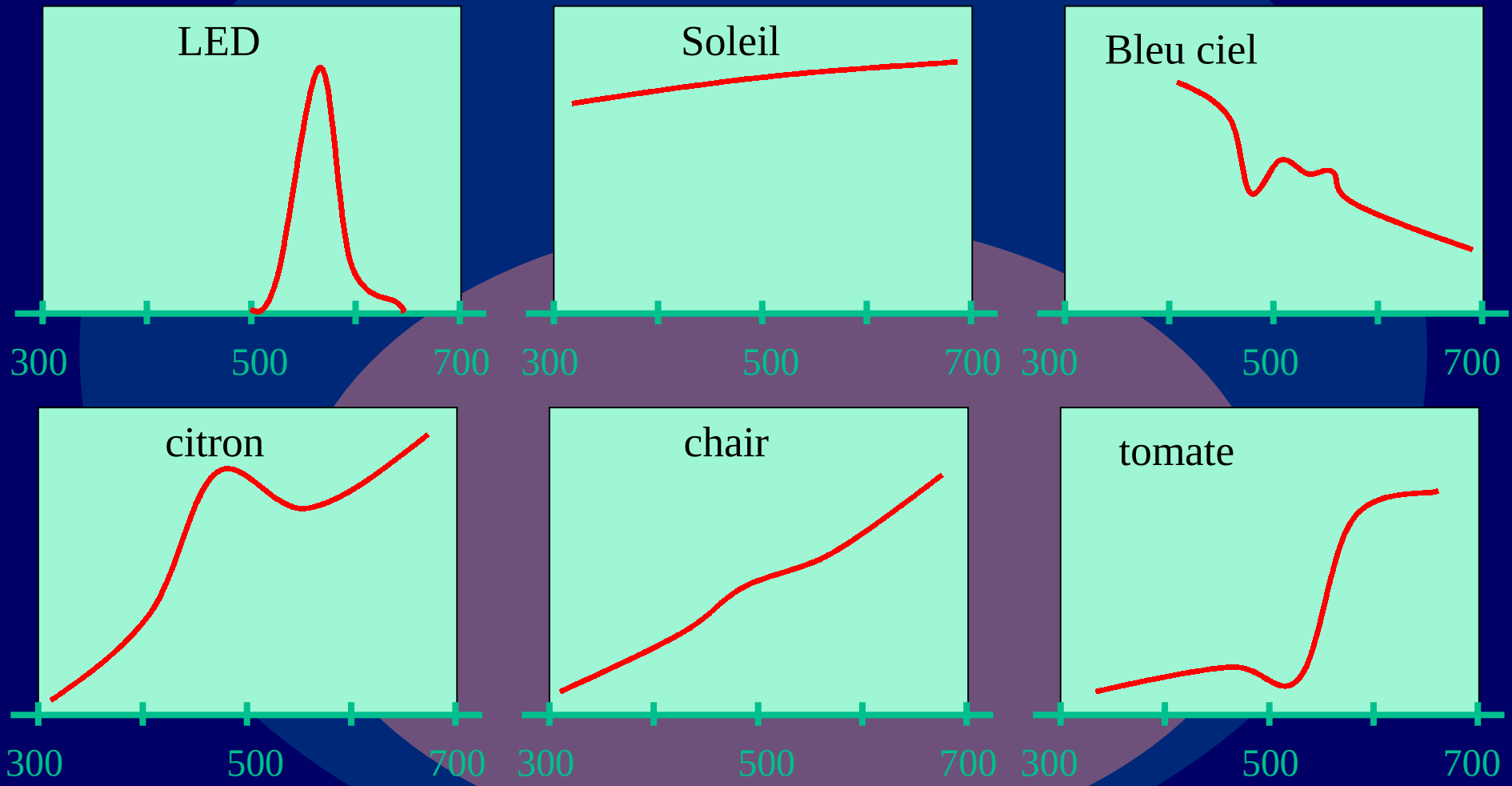
1. Principe
2. Trivariance et trichromie
3. Synthèse additive et soustractive
4. Modèles de représentation de la couleur

I I)C)1) Colorimétrie : Principe

Couleur : phénomène physique interprété par le système visuel humain $\Rightarrow$ objectif et subjectif.



I I)C)1) Colorimétrie : Principe



Puech William

Université Montpellier I I - Nîmes

II)C) Trichromie et trivariance

Trichromie : couleur perçue par un humain ⇒ décomposée dans un espace à 3 dimensions ⇒ 3 couleurs de base avec un spectre éloigné (RVB).

Trivariance : couleur Cx fonction de λ et des luminances de Cλ et de Cb.

- Cx : lumière colorée quelconque
- Cλ : lumière monochromatique
- Cb : lumière blanche

$$\mathbf{Cx = C\lambda + Cb}$$

II)C) Synthèse soustractive

soustraction à la lumière blanche de ses composantes bleues, vertes et rouges à l'aide de filtres jaunes, magenta et cyan.

Jaune + Magenta => Rouge

Jaune + Cyan => Vert

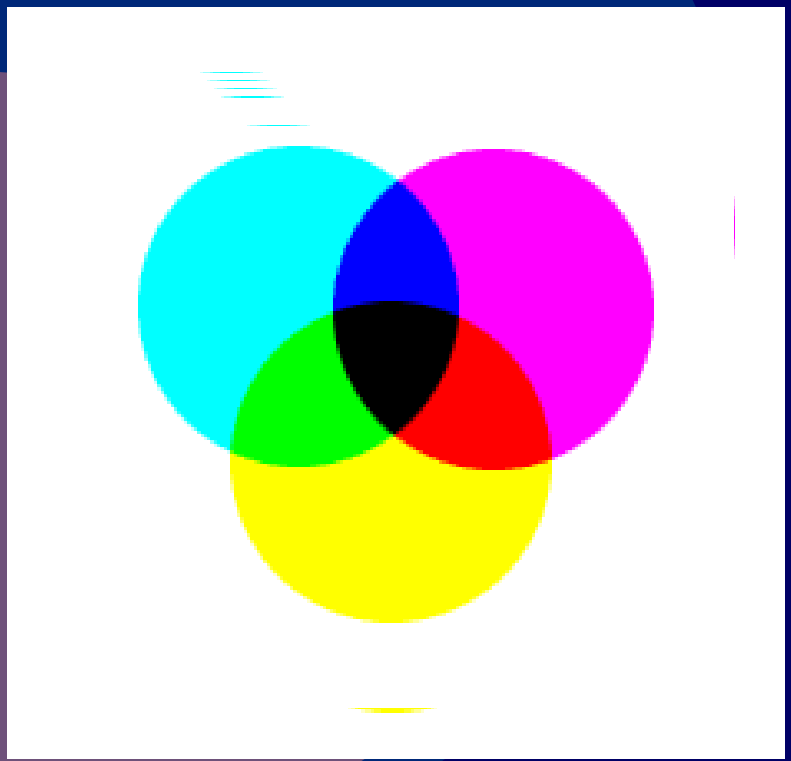
Magenta + Cyan => Bleu

Jaune + Magenta + Cyan => Noir

Les nuances intermédiaires par variation de l'absorption des filtres. Utilisé pour l'impression des couleurs.

II)C) Synthèse soustractive

⇒ lumière diffusée par des objets absorbants : peinture, filtrage.



I I)C) Synthèse additive

Création d'une couleur par addition de trois faisceaux lumineux de couleur rouge, verte et bleue.

⇒ presque toutes les couleurs visibles.

Rouge +Vert => Jaune

Rouge + Bleu => Magenta

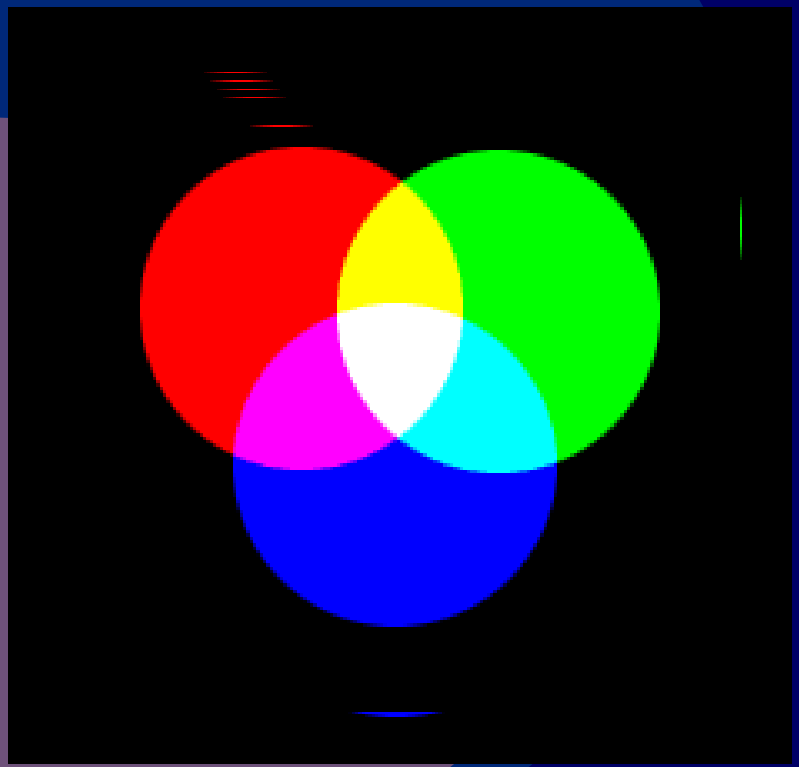
Bleu + Vert => Cyan

Rouge +Vert +Bleu => Blanc

Couleurs intermédiaires produites par variation de l'intensité des faisceaux lumineux.

I I)C) Synthèse additive

⇒ projection (lumière émise) : écrans de télévision ou d'ordinateur.



I I)C) Colorimétrie

4 Modèles de représentation de la couleur :

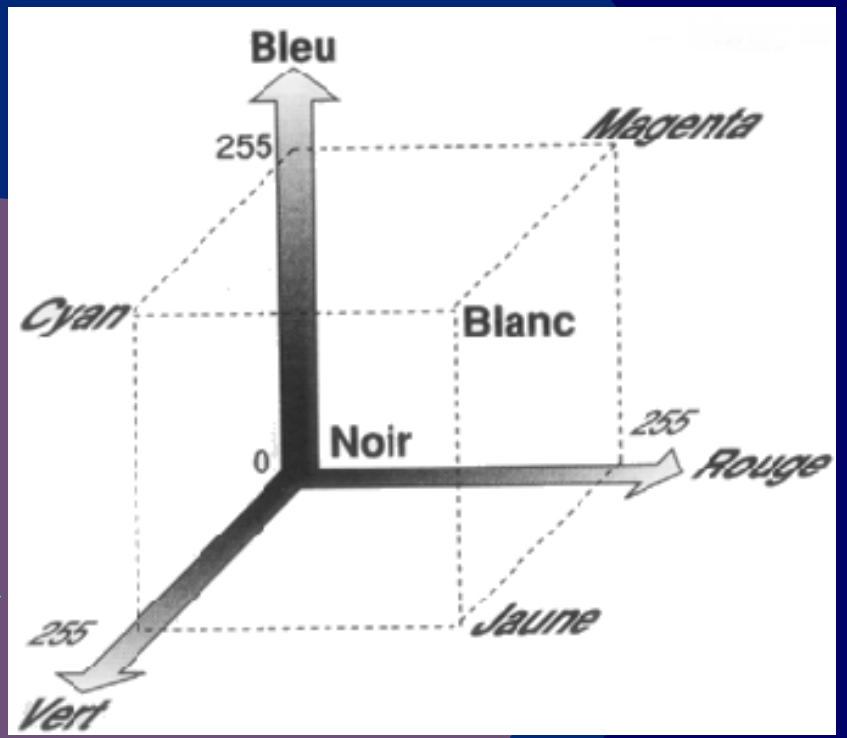
- RVB (RGB): Red Green Blue
- CMJN (CMYK): Cyan Magenta Yellow black (K pour Key black)
- Lab: luminance + chrominance (a et b)
- TSL (HLS) :Teinte (Hue), Saturation, et Luminosité.

I)C) RVB (RGB): Red Green Blue

Composition des couleurs basée sur le principe des couleurs additives : rouge, vert et bleu ⇒ trois primaires utilisées dans la constitution de couleurs à partir de sources lumineuses.

Une image RVB : composée de trois couches, codées chacune sur 8 bits. (256 niveaux de couleur par couche ⇒ 16 millions de couleurs).

RVB utilisé pour la reproduction de couleurs sur écran.

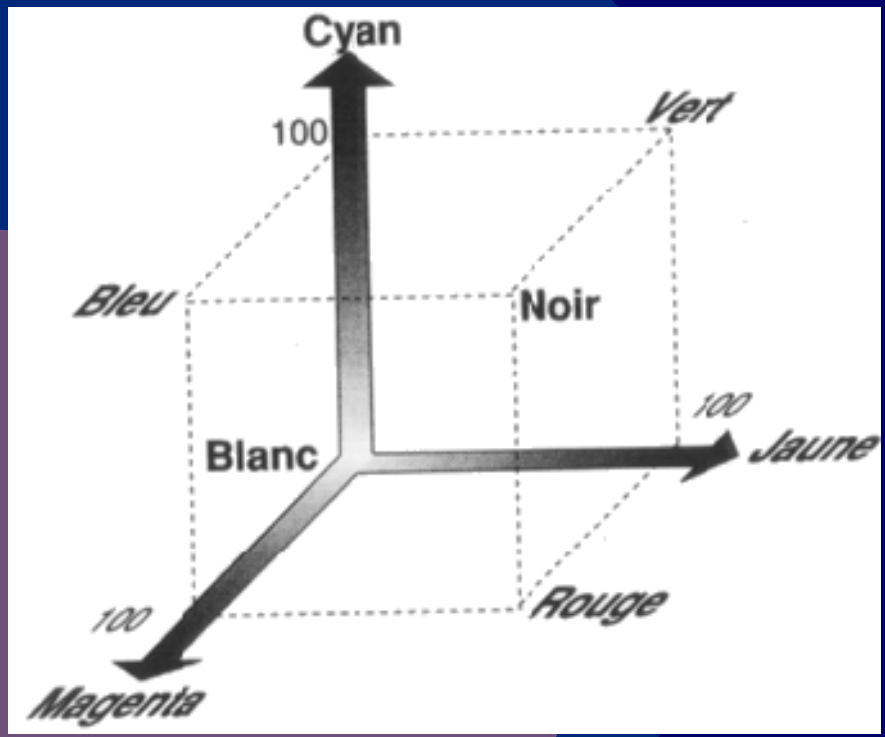


I)C) CMYK : Cyan Magenta Yellow black

Cyan, magenta, jaune et noir ⇒ quatre couleurs d'encre pour les impressions quadrichromiques et pour tout procédé de reproduction à base de pigments ou de colorants (principe des couleurs soustractives).

Sur chaque couche pixel avec un pourcentage d'une des couleurs.

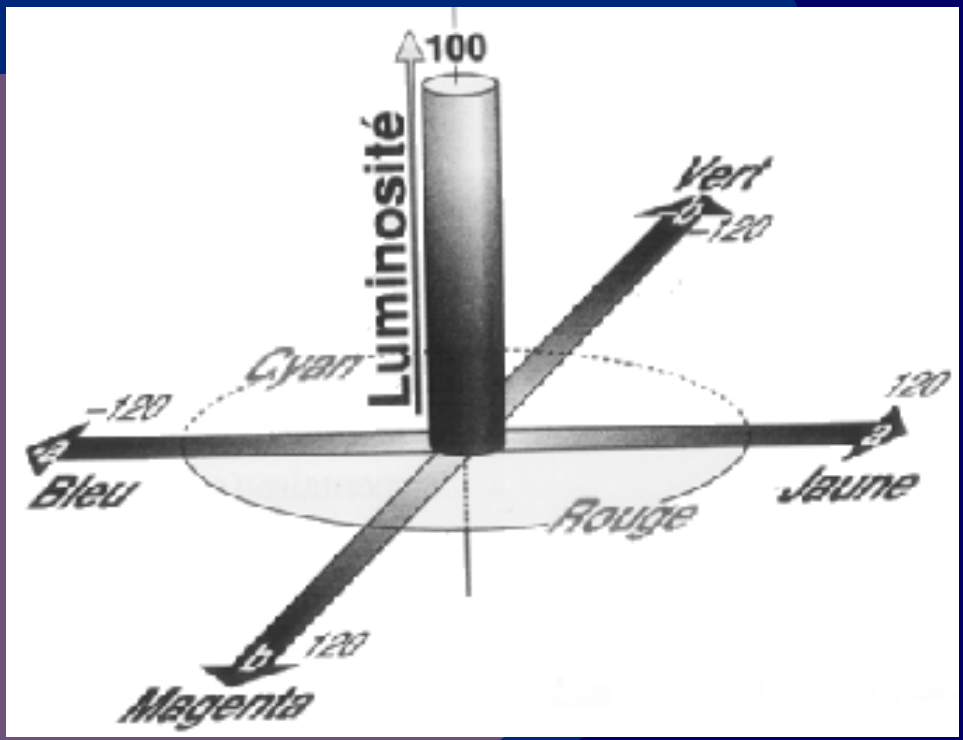
Norme pour l'imprimerie. La composante noire améliore le rendu des aplats noirs. Car le noir peut aussi être obtenu par mélange des trois autres composantes (par économie)



I)C) Lab : luminance + a et b

Couleurs définies par 3 valeurs : Luminosité (luminance) codée en pourcentages, a et b correspondent à l information colorée (chrominance) où la couleur est définie à partir d'un mélange de vert à magenta (a) et un mélange de bleu à jaune (b).

Valeurs comprises entre -120 et +120 pour a et b.



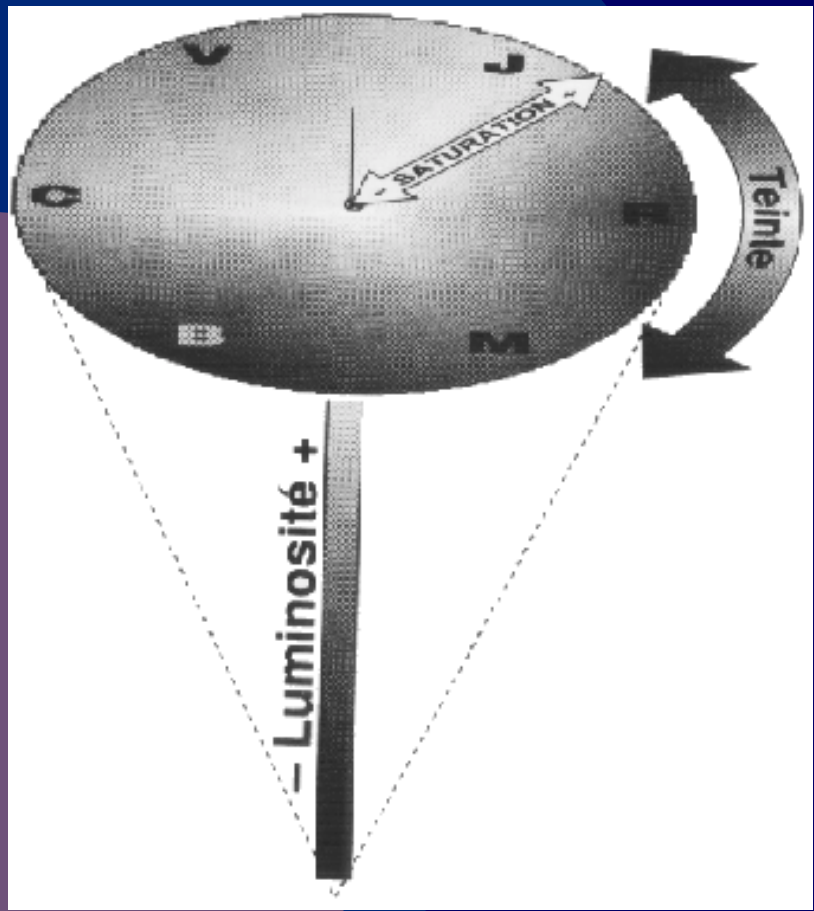
I)C) TSL (HLS) : Teinte (Hue), Saturation, et Luminosité.

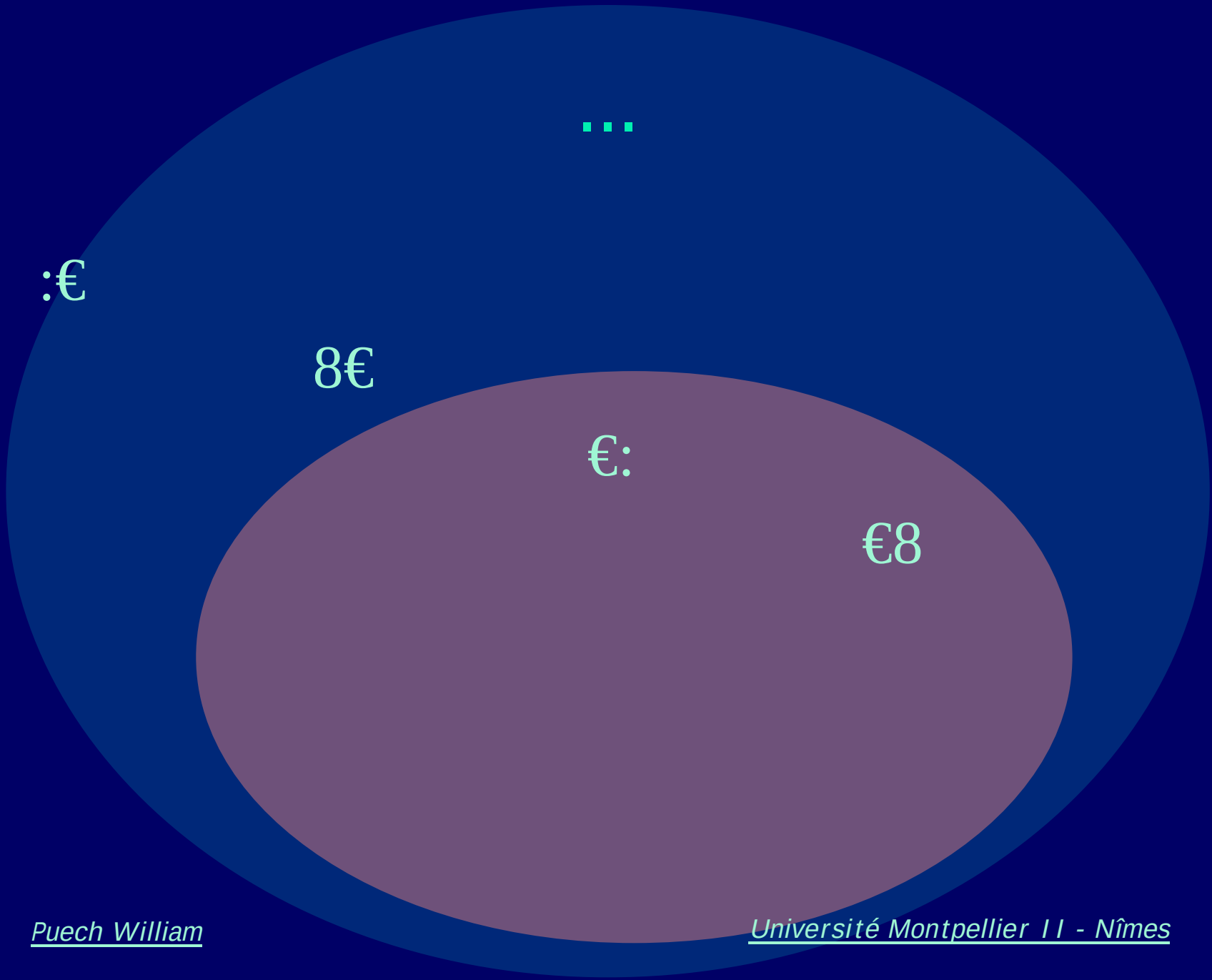
• Teinte : longueur d'onde de la lumière réfléchie, ou transmise par un objet ⇒ emplacement sur la roue chromatique, (0° < angle < 360°).

Le spectre circulaire part du rouge, passe par le vert et le bleu pour revenir au rouge.

• Saturation : pureté ou intensité de la couleur (des couleurs grisées 0 % aux couleurs vives, 100 %).

• Luminosité : variation d'intensité lumineuse d'une couleur, entre 0 %, noir et 100 %, blanc.




...
:€
8€
€:
€8

Les images binaires

Image binaire : image où chaque pixel ne peut avoir pour valeur que 0 ou 1.

Outils spécialisés et théories mathématiques pour la manipulation d'images binaires.

Début du traitement des images numériques : traitement d'images non complexes : problème de temps de calcul, d'espace mémoire disponible et qualité des périphériques de sortie).

Premières applications : reconnaissance de caractères, analyse de traces laissées dans les chambres à bulles par des particules (1950)

Les images binaires

Contexte simple pour la formalisation mathématique des problèmes par des outils tels que la topologie.

En vision industrielle : détection de défauts, contrôle qualité, mesure : on considère l'image binaire comme un passage obligé, suivant en général la phase de segmentation.

Deux catégories d'outils nécessaires pour :

- le codage efficace (et éventuellement la compression),
- et le traitement (analyse et description des formes).

Obtention d'une image binaire à partir d'une image en niveaux de gris par des techniques de seuillage.

Codage des images binaires

Codage dans le but de compression de données à partir de la matrice binaire.

Description des objets codés : codage faisant ressortir les caractéristiques et/ou facilitant les traitements.

Représentation de départ : matrice de pixels à 0 ou 1 dans laquelle apparaissent les objets.

4 types de codages:

- Le codage par segments.
- Le codage par contour.
- Le codage par région.
- Le codage par forme.

Codage par segments

Codage d'une image binaire par segments lignes ou colonnes.

Méthode : on ne conserve pour chaque ligne que la valeur du premier segment et la liste des longueurs des différents segments (dont la valeur change obligatoirement à chaque nouveau segment). Par exemple, la ligne de pixel 1000110 sera codée par (1,1,3,2,1) soit : première valeur 1, longueurs des segments successifs: 1,3,2,1 soit 7 pixels au total. Sur un autre exemple:

1 1 0 0 (1,2,2)
0 1 1 0 (0,1,2,1)
0 0 0 0 (0,4)
1 0 0 0 (1,1,3)

Plus les segments sont courts, moins le codage est efficace car les longueurs sont des entiers et non plus des valeurs binaires. Un codage équivalent est bien sûr possible avec une description en colonnes plutôt qu'en lignes.

Pour réaliser des manipulations de type algébrique (complémentation, ...) plutôt que pour décrire les objets.

Codage par contours

Points de contour : reconnaissables car au moins de leurs points voisins appartient au fond (pas d'ambiguïté de détection). Lors de la recherche du contour d'un objet : inutile de balayer toute l'image; il existe des algorithmes de suivi de contour fournissant un codage sous forme de Freeman de la suite des points du contour.

Soit Pn un point contour courant, le point contour suivant Pn+1 est un voisin de Pn. Le déplacement de Pn à Pn+1 ne peut se faire que dans une des 8 directions du voisinage de Pn selon le codage de Freeman:

3 2 1
4 * 0
5 6 7

Dans la majorité des cas le point Pn n'a que deux candidats potentiels pour Pn+1 dont un est le point Pn-1. Il n'y a donc pas d'ambiguïté. Cependant, il faut aussi prendre en compte les points anguleux n'ayant aucun suivant.

Codage par contours

Suppression de l'image de départ de toutes les configurations du type transition L et transition I :

0 0 0	0 0 0
0 1 0	0 1 0
0 1 1	0 1 0

8 configurations de chaque type par rotation de 45°.

Après extraction du contour : mémorisation des coordonnées du point de départ et la suite des directions d (d dans {0, ..., 8}).

La suite des directions traduit la forme et le point de départ de sa localisation spatiale. Très utile pour la reconnaissance de formes car signature invariante d'une forme. De très nombreux paramètres de forme peuvent être mesurés directement sur le codage de Freeman (périmètre, aire, centre de gravité, axes d'inertie, ...).

Le codage plus efficace (compression) si les objets sont gros. Soit O un objet comportant A pixels dont P sont des pixels de son contour.

Codage binaire classique : A bits dans l'image de départ

Codage de Freeman : 3P bits + les coordonnées d'un point de référence.

facteur de compression lié au ratio P/A.

Codage par régions

Il est bien sûr possible de généraliser le codage par segments au codage des entités surfaciques bidimensionnelles. Parmi les nombreuses techniques de codage par région, la plus connue fait appel aux quadtrees. Il s'agit d'un découpage récursif du support image jusqu'à obtention de blocs homogènes (ayant tous la même valeur). A la ième itération, on définit les carrés de côté 2n-i (où 2n est le côté de l'image initiale); le plus petit bloc possible est le pixel. Ces primitives sont ensuite organisées de façon arborescente. Exemple :

Codage par régions

Pour le codage, on utilise une technique de parenthésage (une parenthèse équivaut à un niveau de récursivité dans l'algorithme de recherche de blocs homogènes) et on mémorise uniquement la valeur du bloc. Pour lever toute ambiguité, on adopte un sens de balayage unique :

1 2
3 4

On obtient alors le codage suivant :

I = (((0001)(0)(0111)(1))((0010)(0)(1000)(0))(1)(0))

Cette représentation permet des opérations ensemblistes (intersection, union,...), la recherche de contours,...; mais elle n'est pas vraiment adaptée à la description de formes de par sa non invariance en translation.

D'autres types de pavages peuvent être définis, par exemple en polygones de Voronoï, à partir de germes prédéfinis dans l'image.

Codage par forme

Méthodes tenant compte de la morphologie de l'objet : codages par squelettes. Squelette : représentation filiforme centrée sur l'objet initial, obtenu par amincissements successifs. Cet amincissement se réalise sous la contrainte de préservation des points significatifs de son élongation, ou nécessaires à la connexité du squelette résultant (transformation homotope).

Recherche d'algorithmes efficaces (temps de calcul, qualité du résultat), pour la reconnaissance des caractères. Si le squelette est unique et sans problème dans un espace continu, il est beaucoup moins facile à trouver dans un espace discret qu'est une image numérique. Sa définition plus précise nécessite de redéfinir les concepts de base de la géométrie, ce qui a donné naissance à la géométrie discrète.

Transformation en squelette d'un objet irréversible : mais résultat significatif de l'allure de l'objet codé et particulièrement adapté aux objets minces.

Réversibilité du codage primordiale pour des objets à stocker. Notion d'axe médian intéressante: on recouvre l'objet par des boules de taille maximale incluse dans l'objet et centrées sur les points de l'objet. L'axe médian est ensuite formé des centres des boules qui ne sont pas incluses dans aucune autre et l'on associe aux centres conservés la taille de leur boule.

Codage par forme

Dans l'exemple suivant, on représente en chaque point la taille de la plus grande boule centrée sur le point et complètement contenue dans la forme (une taille de 1 équivaut à un carré 1 x 1, 2 à un carré 3x3 et 3 à un carré 5x5):

```
      1 1 1 1 1
  1 1 1 2 2 2 1
1 1 2 2 2 3 2 1
  1 1 1 2 2 2 1
      1 1 1 1 1
```

Les points en noirs sont les trois points de l'axe médian nécessaire pour reconstruire la forme initiale. la taille de la boule correspond en fait à la notion de distance au contour. Le résultat final dépend de la distance utilisée. Le problème de l'axe médian est qu'il n'est pas formé de points consécutifs. Pour le faire "ressembler" à un squelette, on reconnecte les poins de l'axe médian par un chemin de crête. On obtient alors la ligne médiane que l'on peut voir comme un squelette pondéré de l'objet initial. La ligne médiane et l'axe médian sont des codages réversibles.