

Décision probabiliste

M1 AIV1 - Cours 9

2 mars 2022

- 1 Rappels de probabilités
- 2 Probabilités conditionnelles
- 3 Décision probabiliste - apprentissage supervisé
- 4 Décision probabiliste - apprentissage non supervisé

Motivations et objectifs du cours

Décision
probabiliste

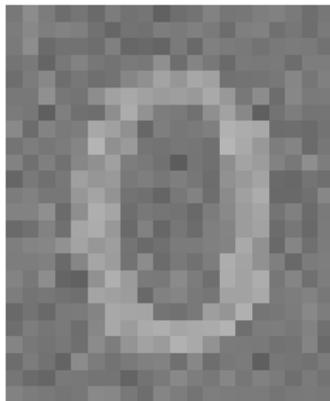
M1 AIV1 -
Cours 9

Rappels de
probabilités

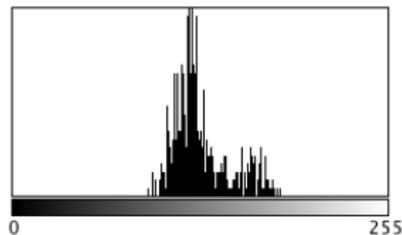
Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

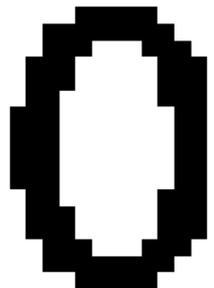
Décision
probabiliste -
apprentissage
non supervisé



(a) Image originale



(b) histogramme



(c) Image binaire

- Seuiller les images pour en extraire les moments caractéristiques (cf. Cours de Mr Cabestaing).
- Construire d'une fonction de densité de probabilité
- Mettre en oeuvre des fonctions de décision probabiliste

Éléments d'un modèle probabiliste

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- L'univers Ω est l'ensemble de tous les résultats possibles que l'on peut obtenir.
- Exemple 1 : On lance une pièce une fois, $\Omega = \{face, pile\}$.
- Exemple 2 : On lance une pièce deux fois,
 $\Omega = \{(face, face), (face, pile), (pile, face), (pile, pile)\}$.

Eléments d'un modèle probabiliste

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

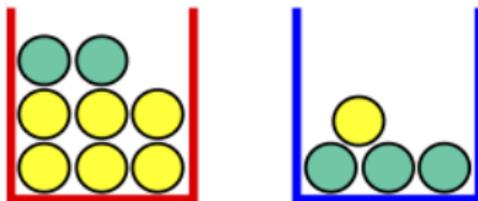
Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

Un évènement A est un sous-ensemble des résultats possibles
 $A \subset \Omega$ est l'ensemble de tous les résultats possibles que l'on
peut obtenir.

- Exemple 1 : On tire une seule carte. Tirer une carte avec le roi est un évènement, tirer un chiffre est un évènement
- Exemple 2 On choisit une urne au hasard et on tire une boule de l'urne choisie.
 - Tirer une boule verte est un évènement
 - Tirer une boule verte de l'urne bleue est un évènement



Eléments d'un modèle probabiliste

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

La **loi de probabilité** est une fonction sur l'ensemble des évènements Ω qui fournit une valeur réelle, positive à tout évènement $A \subset \Omega$, respectant l'axiome suivant :

- $P(A) \geq 0$
- $P(\Omega) = 1$
- Pour les évènements indépendants A_1, \dots, A_n
 $\forall i \neq j \in 1, \dots, n$, tels que $P(A_i \cap A_j) = \prod_{i=1}^n P(A_i)$
- $P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$

Estimation fréquentielle d'une probabilité

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

La probabilité d'un évènement A_i est la limite de la proportion d'occurrences que se produit A_i lorsque le nombre N de fois que l'expérience se répète tend vers l'infini.

- Soit N_{A_i} le nombre de fois que se produit A_i (par exemple "la pièce lancée est face").
- La probabilité de réalisation de A_i est $\lim_{N \rightarrow \infty} \frac{N_{A_i}}{N}$
- En pratique on pose $P(A_i) \simeq \frac{N_{A_i}}{N}$

Exemple : histogramme d'une image

Décision
probabiliste

M1 AIV1 -
Cours 9

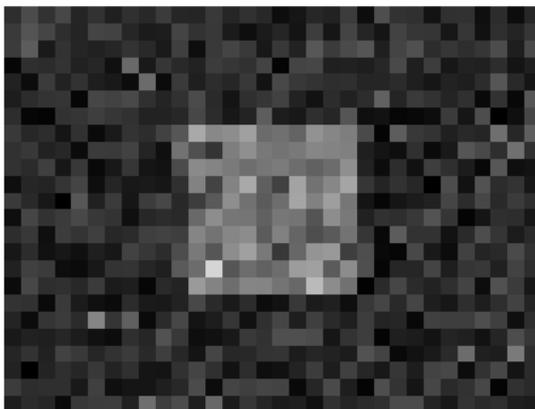
Rappels de
probabilités

Probabilités
conditionnelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Soit une image en niveaux de gris de taille 36×24 et considérons l'évènement que le niveau de gris I d'un pixel est égal à i
- Soit $N_{I=i}$ le nombre de fois que le niveau de gris d'un pixel est égal à i .



Exemple : histogramme d'une image

Décision
probabiliste

M1 AIV1 -
Cours 9

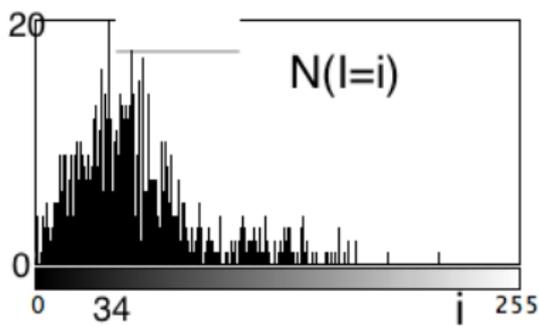
Rappels de
probabilités

Probabilités
conditionnelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Soit une image en niveaux de gris I de taille 36×24 .
- Soit $N(I = i)$ le nombre de fois que le niveau de gris d'un pixel est égal à i .



Exemple : histogramme d'une image

Décision
probabiliste

M1 AIV1 -
Cours 9

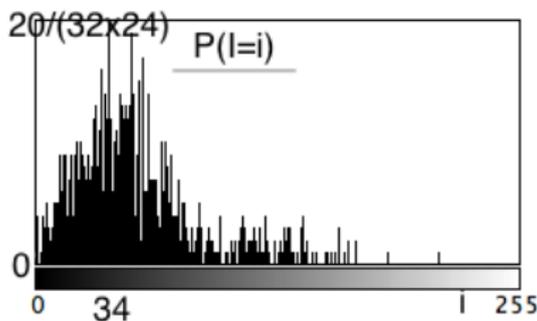
Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Soit une image en niveaux de gris I de taille $N = 36 \times 24$.
- Soit $P(I = i)$ la probabilité que le niveau de gris I d'un pixel est égal à i .



Notion d'indépendance

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

Définition

- Deux évènements A et B sont indépendants si

$$P(A \cap B) = P(A).P(B)$$

- Les n évènements A_i sont indépendants si

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = \prod_{i=1}^n P(A_i)$$

Exemple : On lance un dé deux fois

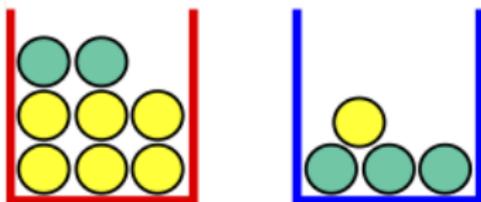
- Les évènements $A :=$ "6 sort au premier lancer" et $B :=$ "6 sort au deuxième lancer" sont **indépendants**

Probabilité conditionnelle

La probabilité conditionnelle d'un évènement A sachant un autre évènement B (tel que $P(B) \neq 0$) est définie par

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Exemple : On choisit une urne au hasard et on tire une boule de l'urne choisie. Quelle est la probabilité d'avoir une boule jaune issue de l'urne bleue ?



Théorème de Bayes

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

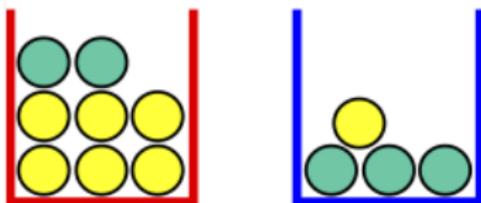
Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

Soient les évènements A_1, \dots, A_n disjoints tels que $P(A_i) > 0, i = 1..n$, la probabilité conditionnelle

$$P(A_i/B) = \frac{P(B/A_i).P(A_i)}{P(B)} = \frac{P(B/A_i).P(A_i)}{\sum_{i=1}^n P(B/A_i).P(A_i)}$$

Exemple : Sachant qu'une boule jaune est tirée, quelle est la probabilité qu'elle provienne de l'urne bleue ?



Probabilité a posteriori

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Soit K classes $\omega_k, k = 1, \dots, K$.
- Chaque classe ω_k contient N_k .
- La probabilité a priori de la classe ω_k

$$P(\omega_k) = \frac{N_k}{N}.$$

- La somme des observations $N = \sum_{k=1}^K N_k$.
- On peut calculer une table d'occurrences $h_k(X)$ pour les observations X de la classe ω_k .
- Probabilité conditionnelle :

$$P(X/\omega_k) \simeq \frac{h_k(X)}{N_k}.$$

- Probabilité d'une variable X :

$$P(X) = \sum_{k=1}^K P(\omega_k) \cdot P(X/\omega_k)$$

Probabilité a posteriori

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Probabilité d'une variable X :

$$P(X) = \sum_{k=1}^K P(\omega_k) \cdot P(X/\omega_k)$$

$$P(X) \simeq \frac{1}{N} \sum_{k=1}^K N_k \cdot h_k(X) = h(X)$$

- Probabilité a posteriori (selon th. De Bayes) :

$$P(\omega_k/X) = \frac{P(X/\omega_k) \cdot P(\omega_k)}{P(X)}$$

$$P(\omega_k/X) = \frac{h_k(X)}{h(X)}$$

Classification à 2 classes

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

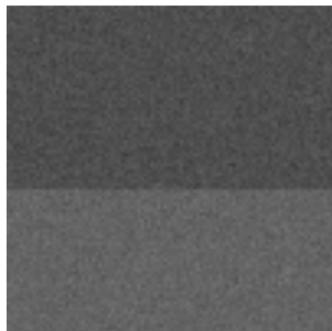
Décision
probabiliste -
apprentissage
non supervisé

- Pour 2 classes ω_1 et ω_2 ,
- Probabilité a posteriori (selon th. De Bayes) :

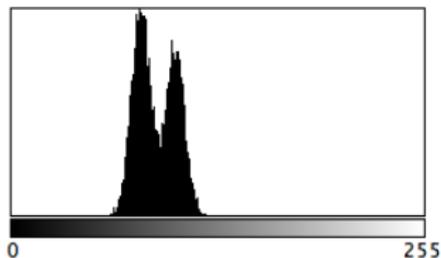
$$P(\omega_1/X) = \frac{P(X/\omega_1).P(\omega_1)}{P(X)}$$

avec

$$P(X) = P(X/\omega_1).P(\omega_1) + P(X/\omega_2).P(\omega_2)$$



(d) Image



Count: 10100 Min: 57
Mean: 89.259 Max: 123
StdDev: 12.221 Mode: 79 (366)

(e) Histogramme

Classification à 2 classes

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Comment décider l'assignation de X à ω_1 ou ω_2
- Minimiser la probabilité $P(\varepsilon/X)$ d'erreur ε d'assignation :
 - $P(\varepsilon/X) = P(\omega_1/X)$ si $X \in \omega_2$.
 - $P(\varepsilon/X) = P(\omega_2/X)$ si $X \in \omega_1$.
- Règle : si $P(\omega_1/X) > P(\omega_2/X)$ alors assigner X à ω_2 sinon assigner X à ω_1 .
- Règle (de Bayes) : si $P(X/\omega_1).P(\omega_1) > P(X/\omega_2).P(\omega_2)$ alors assigner X à ω_2 sinon assigner X à ω_1 .
- les X tels que $P(X/\omega_1).P(\omega_1) = P(X/\omega_2).P(\omega_2)$ définissent la frontière de décision entre les 2 classes.

Classification supervisée des pixels 2 classes

Décision
probabiliste

M1 AIV1 -
Cours 9

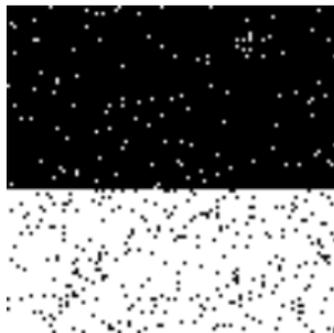
Rappels de
probabilités

Probabilités
conditionnelles

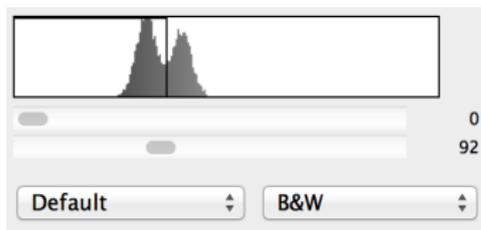
Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Frontière quand le niveau de gris = 93



(f) Image



(g) Histogramme

Assignation à 2 classes (Bayes)

Décision probabiliste

M1 AIV1 -
Cours 9

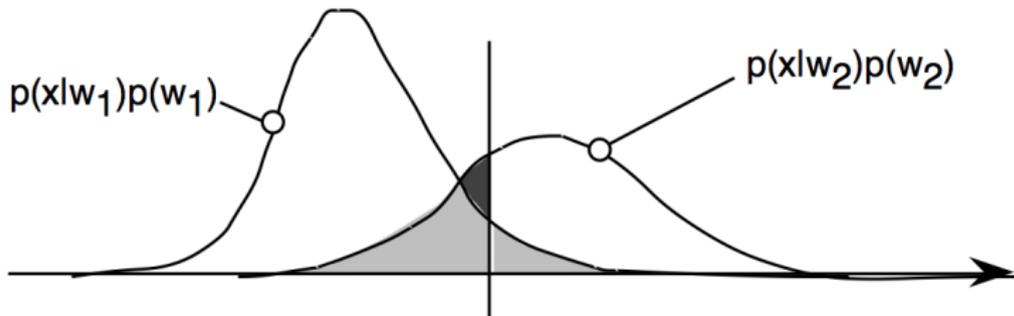
Rappels de
probabilités

Probabilités
conditionnelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- On connaît à priori ω_1 et ω_2 .
- Une valeur \hat{X} de décision définit 2 classes $\hat{\omega}_1$ et $\hat{\omega}_2$.
- La probabilité d'erreur est donc
$$P(\varepsilon/\hat{X}) = P(X \in \hat{\omega}_2, \omega_1) + P(X \in \hat{\omega}_1, \omega_2)$$
- $P(\varepsilon/\hat{X}) = P(X \in \hat{\omega}_2/\omega_1) \cdot P(\omega_1) + P(X \in \hat{\omega}_1/\omega_2) \cdot P(\omega_2)$
- $P(\varepsilon/\hat{X}) = \int_{X \in \hat{\omega}_2} P(X/\omega_1) \cdot P(\omega_1) dX + \int_{X \in \hat{\omega}_1} P(X/\omega_2) \cdot P(\omega_2) dX$



Assignment à 2 classes (Bayes)

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Soit la valeur optimale \hat{X}^0 qui minimise l'erreur

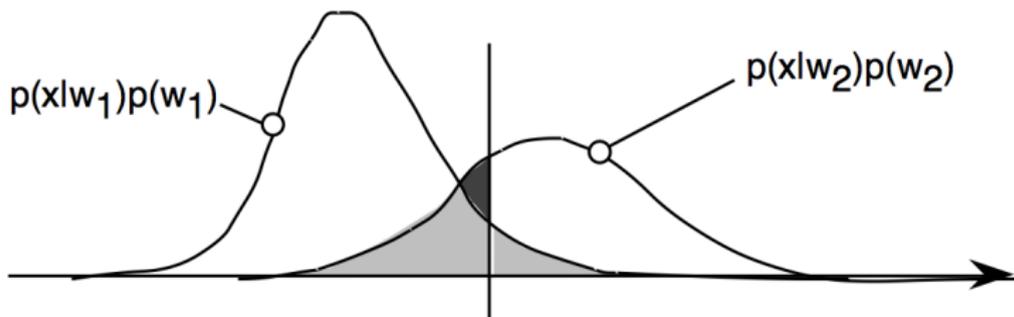
- $\hat{X}^0 = \arg \min (P(\varepsilon/\hat{X}))$



$$\frac{dP(\varepsilon/\hat{X})}{d\hat{X}} \Big|_{\hat{X}=\hat{X}^0} = 0$$

- soit

$$P(\hat{X}^0/\omega_1) \cdot P(\omega_1) = P(\hat{X}^0/\omega_2) \cdot P(\omega_2)$$



Segmentation en 2 classes à partir d'un histogramme

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
conditionnelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- La probabilité a priori de la classe ω_k

$$P(\omega_k) = \frac{N_k}{N}.$$

- On peut calculer une table d'occurrences $h_k(X)$ pour les observations X de la classe ω_k .
- Probabilité conditionnelle :

$$P(X/\omega_k) \simeq \frac{h_k(X)}{N_k}.$$

- Pour t compris entre 0 et 255, calculer
 - $\varepsilon(t) = \sum_{i=0}^t \frac{h_2(i)}{h(i)} \cdot P(\omega_2) + \sum_{i=t+1}^{255} \frac{h_1(i)}{h(i)} \cdot P(\omega_1)$
- le seuil optimal $\hat{t} = \arg \min (\varepsilon(t))$

Assignation à K classes (Bayes)

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- On connaît à priori ω_k .
- La probabilité de bonne assignation à maximiser est donc
$$P(\bar{\varepsilon}/\hat{X}) = \sum_{k=1}^K P(X \in \hat{\omega}_k, \omega_k)$$
- $$P(\bar{\varepsilon}/\hat{X}) = \sum_{k=1}^K P(X \in \hat{\omega}_k / \omega_k) \cdot P(\omega_k)$$

$$P(\bar{\varepsilon}/\hat{X}) = \sum_{k=1}^K \int_{X \in \hat{\omega}_k} P(X/\omega_k) \cdot P(\omega_k) dX$$

- Soit en pratique

$$P(\bar{\varepsilon}/\hat{X}) \simeq \sum_{k=1}^K \sum_{\hat{X} \in \hat{\omega}_k} \frac{h_k(\hat{X})}{N_k} \cdot \frac{N_k}{N} = \sum_{k=1}^K \sum_{\hat{X} \in \hat{\omega}_k} \frac{h_k(\hat{X})}{N}$$

Segmentation en 3 classes à partir d'un histogramme

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
conditionnelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- La probabilité a priori de la classe ω_k

$$P(\omega_k) = \frac{N_k}{N}.$$

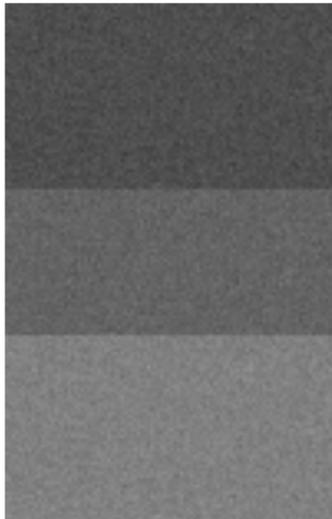
- On peut calculer une table d'occurrences $h_k(X)$ pour les observations X de la classe ω_k .
- Probabilité conditionnelle :

$$P(X/\omega_k) \simeq \frac{h_k(X)}{N_k}.$$

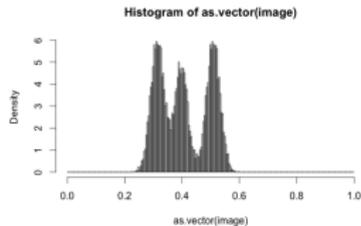
- Pour t_1 et t_2 compris entre 0 et 255, calculer
 - $B(t_1, t_2) = \sum_{i=0}^{t_1} \frac{h_1(i)}{h(i)} \cdot P(\omega_1) + \sum_{i=t_1+1}^{t_2} \frac{h_2(i)}{h(i)} \cdot P(\omega_2) + \sum_{i=t_2+1}^{255} \frac{h_3(i)}{h(i)} \cdot P(\omega_3)$
- le seuil optimal $(\hat{t}_1, \hat{t}_2) = \arg \max (B(t_1, t_2))$

Classification supervisée des pixels en 3 classes

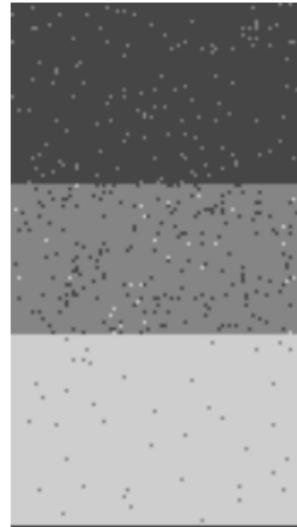
- Seuils déterminés automatiquement



(h) Image



(i) Histogramme



(j) Image segment

Décision probabiliste

M1 AIV1 - Cours 9

Rappels de probabilités

Probabilités conditionnelles

Décision probabiliste - apprentissage supervisé

Décision probabiliste - apprentissage non supervisé

Segmentation non supervisée en 2 classes - recherche du max de la variance inter-classe (Otsu)

Décision probabiliste

M1 AIV1 - Cours 9

Rappels de probabilités

Probabilités conditionnelles

Décision probabiliste - apprentissage supervisé

Décision probabiliste - apprentissage non supervisé

- Méthode d'Otsu cherche à maximiser la dispersion entre classes
- La probabilité a priori de la classe $\omega_k(t)$

$$P(\omega_1(t)) = \sum_{i=0}^t \frac{h(i)}{N} \text{ et } N(\omega_1(t)) = \sum_{i=0}^t h(i)$$

$$P(\omega_2(t)) = 1 - P(\omega_1(t)) \text{ et } N(\omega_2(t)) = N - N(\omega_1(t))$$

- La moyenne de la classe $\omega_k(t)$

$$\mu(\omega_1(t)) = \frac{\sum_{i=0}^t i \cdot h(i)}{N(\omega_1(t))}$$

$$\mu(\omega_2(t)) = \frac{\sum_{i=t+1}^{255} i \cdot h(i)}{N(\omega_2(t))}$$

Segmentation non supervisée en 2 classes - recherche du max de la variance inter-classe (Otsu)

Décision
probabiliste

M1 AIV1 -
Cours 9

Rappels de
probabilités

Probabilités
condition-
nelles

Décision
probabiliste -
apprentissage
supervisé

Décision
probabiliste -
apprentissage
non supervisé

- Méthode d'Otsu cherche à maximiser la dispersion entre classes
- Pour t compris entre 0 et 255, calculer
 - $\sigma(t) = P(\omega_1(t)).P(\omega_2(t)).(\mu(\omega_1(t)) - \mu(\omega_2(t)))^2$
- le seuil optimal $\hat{t} = \arg \max (\sigma(t))$

Segmentation non supervisée en 3 classes - recherche du min de la variance intra-classe (Otsu)

- Pour t_1 et t_2 compris entre 0 et 255, calculer

- $$P(\omega_1(t_1)) = \sum_{i=0}^{t_1} \frac{h(i)}{N} \text{ et } N(\omega_1(t_1)) = \sum_{i=0}^{t_1} h(i)$$

- $$P(\omega_2(t_1, t_2)) = \sum_{i=t_1+1}^{t_2} \frac{h(i)}{N} \text{ et } N(\omega_2(t_1, t_2)) = \sum_{i=t_1+1}^{t_2} h(i)$$

- $$P(\omega_3(t_2)) = \sum_{i=t_2+1}^{255} \frac{h(i)}{N} \text{ et } N(\omega_3(t_2)) = \sum_{i=t_2+1}^{255} h(i)$$

- $$\mu_1(t_1) = \sum_{i=0}^{t_1} \frac{i * h(i)}{N(\omega_1(t_1))}$$

Segmentation non supervisée en 3 classes - recherche du min de la variance intra-classe (Otsu)

- Pour $t1$ et $t2$ compris entre 0 et 255, calculer

- $$\mu_2(t1, t2) = \sum_{i=t1+1}^{t2} \frac{i * h(i)}{N(\omega_2(t1, t2))}$$

- $$\mu_3(t2) = \sum_{i=t2+1}^{255} \frac{i * h(i)}{N(\omega_3(t2))}$$

- $$\sigma(t1, t2) = P(\omega_1(t1)) * \sum_{i=0}^{t1} \frac{(i - \mu_1(t1))^2 * h(i)}{N(\omega_1(t1))} +$$

$$P(\omega_2(t1, t2)) * \sum_{i=t1+1}^{t2} \frac{(i - \mu_2(t1, t2))^2 * h(i)}{N(\omega_2(t1, t2))} +$$

$$P(\omega_3(t2)) * \sum_{i=t2+1}^{255} \frac{(i - \mu_3(t2))^2 * h(i)}{N(\omega_3(t2))}$$

- le seuil optimal $(\hat{t1}, \hat{t2}) = \arg \min (\sigma(t1, t2))$

Décision probabiliste

M1 AIV1 - Cours 9

Rappels de probabilités

Probabilités conditionnelles

Décision probabiliste - apprentissage supervisé

Décision probabiliste - apprentissage non supervisé

Ludovic MACAIRE

Laboratory CRISTAL UMR 9189 - University of Lille

Work of the color team : <http://color.univ-lille.fr/presentation>

PhD students

Sofiane Mihoubi

Anis Amziame

Arezki Aberkane

Color team members

Dr. Olivier Losson

Dr. Benjamin Mathon

Image couleur

- Formation de la couleur
- Perception de la couleur
- Elements de colorimétrie
- Représentation de la couleur



Couleur numérique
*acquisition, perception,
codage et rendu*

sous la direction de
Christine Fernandez-Maloigne
Frédérique Robert-Inacio
Ludovic Macaire



Imagerie numérique
avancées et perspectives pour la couleur

sous la direction de
Christine Fernandez-Maloigne
Ludovic Macaire
Frédérique Robert-Inacio

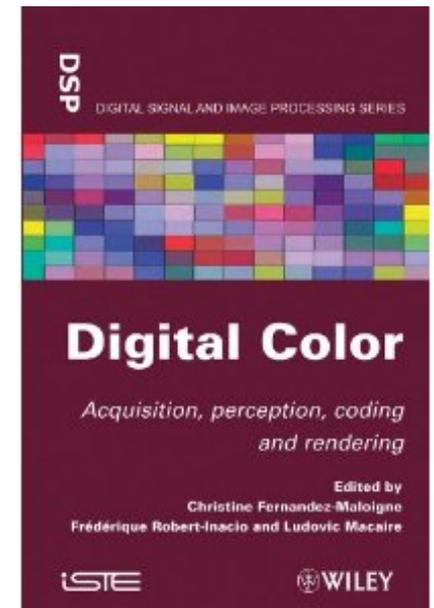
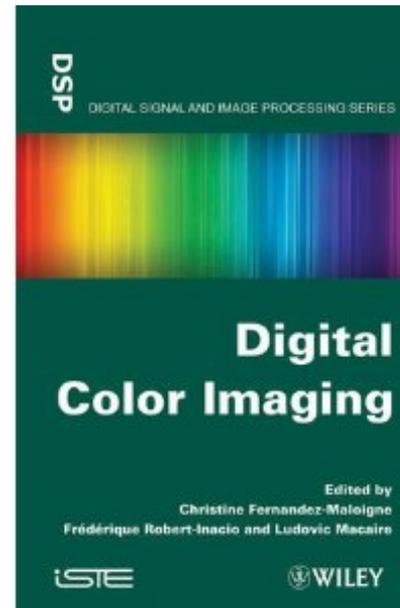


Image formation (Lambertian model)

Light



Illumination SPD

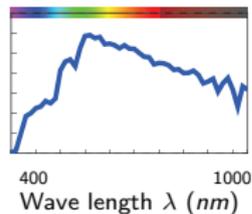


Image formation (Lambertian model)

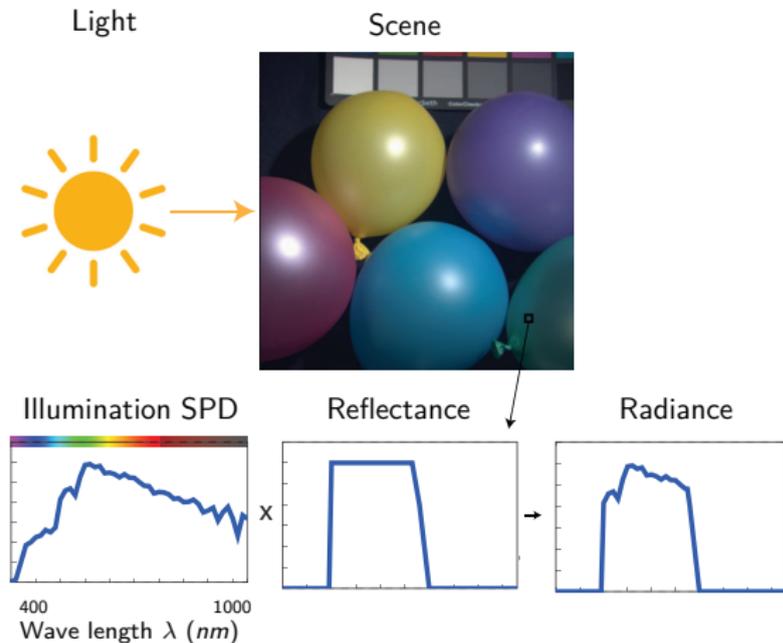
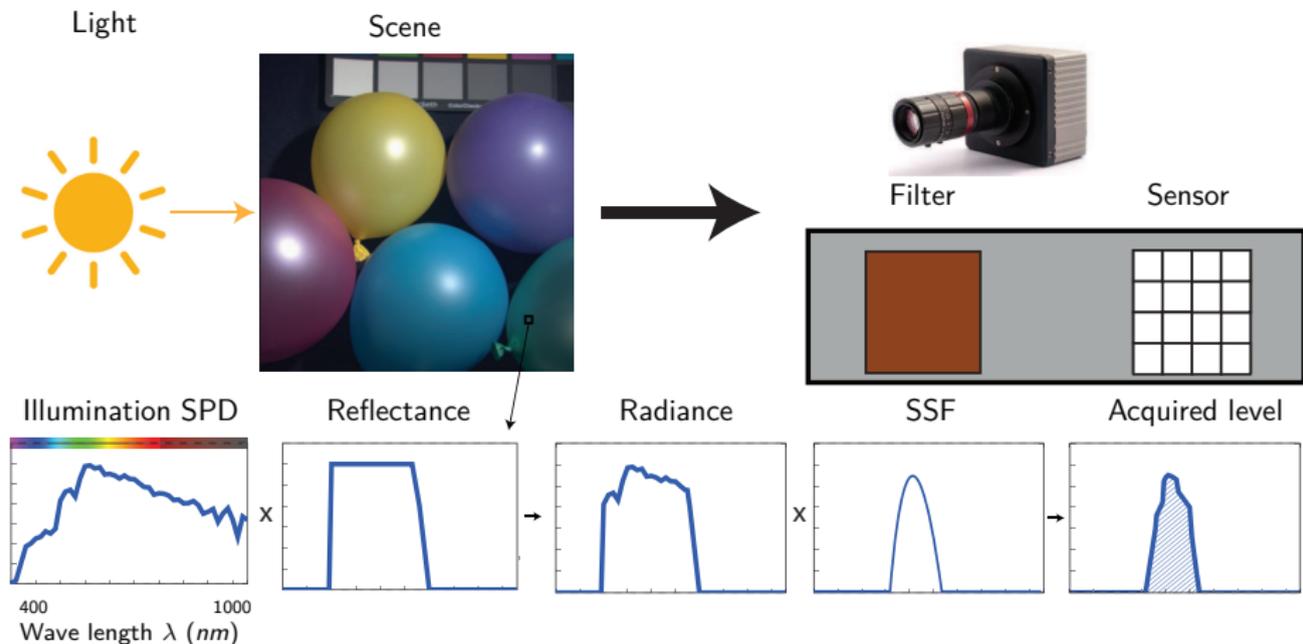


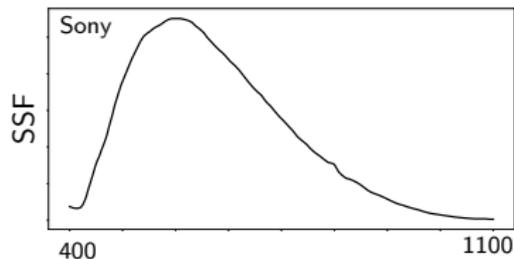
Image formation (Lambertian model)



Image

Panchromatic image :

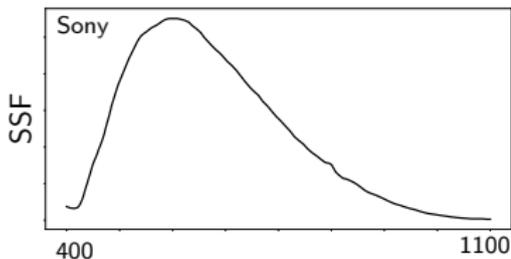
1 channel associated to the Spectral Sensitivity Function (SSF) of one sensor.



Image

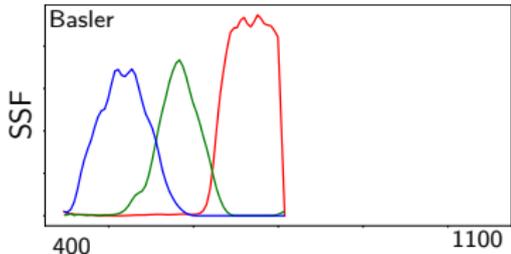
Panchromatic image :

1 channel associated to the Spectral Sensitivity Function (SSF) of one sensor.



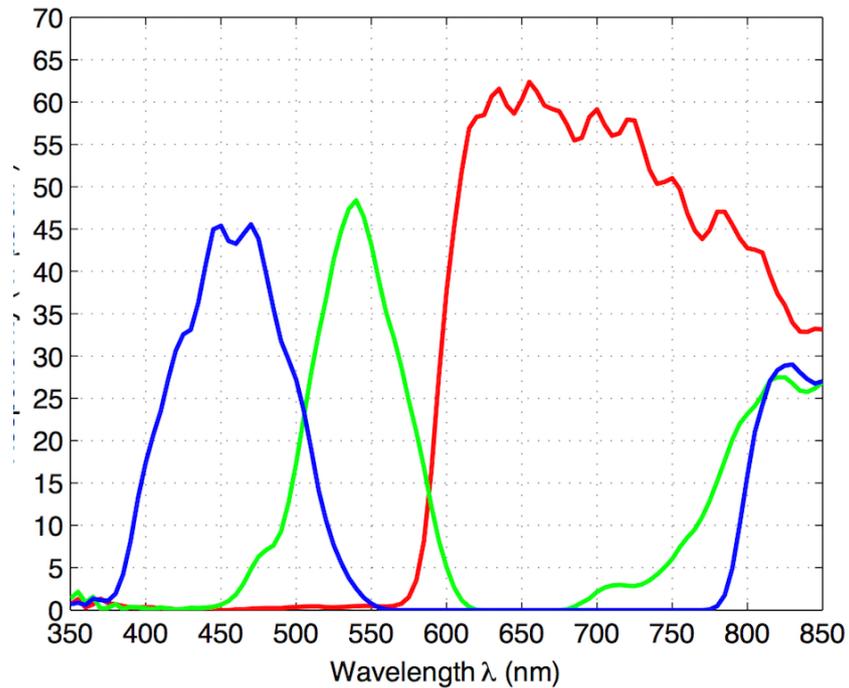
Color image :

3 color channels (R, G, B) in the visible spectral domain.



Sensibilités spectrales

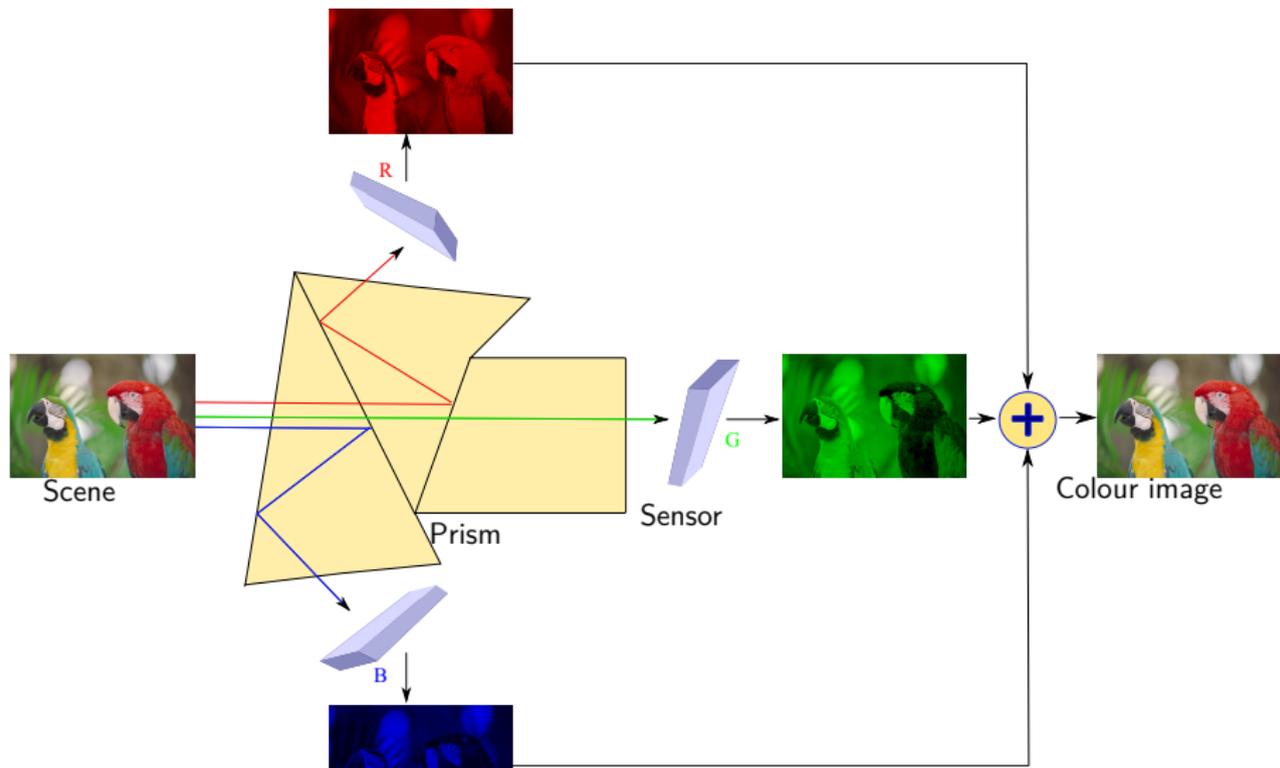
Caméra couleur



Sensibilité spectrale des photo-récepteurs

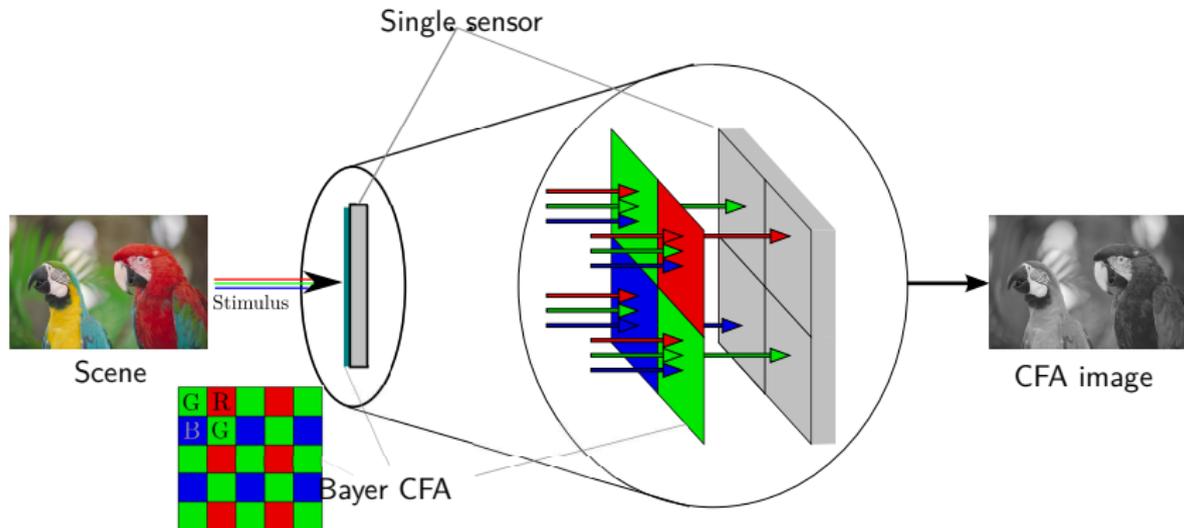
3-sensor colour camera

- Trichroic prism assembly splits stimulus into beams of differing wavelength on 3 sensors.



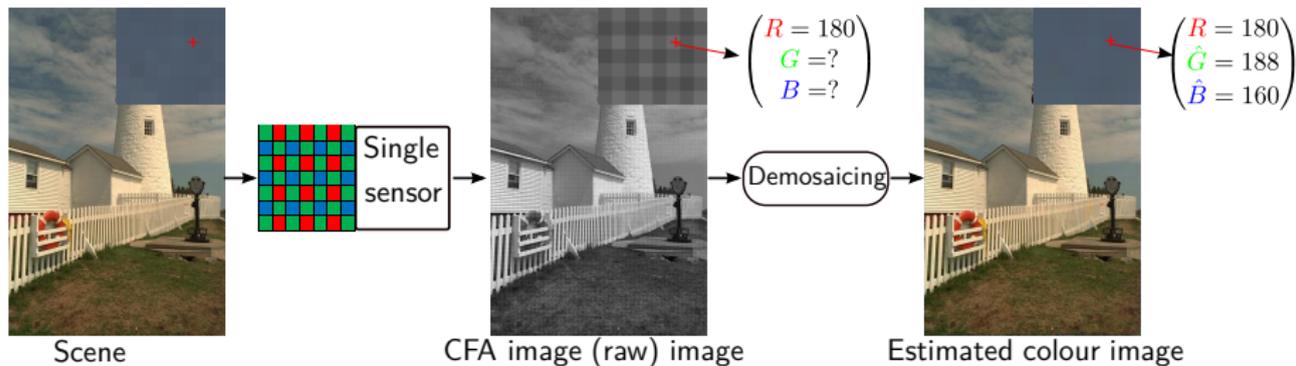
Single sensor colour camera

- Bayer Colour Filter Array (2x2 pattern) located in front of the sensor.



- Acquisition of CFA (raw) image in one shot.
- Requires demosaicing to retrieve the colour image.

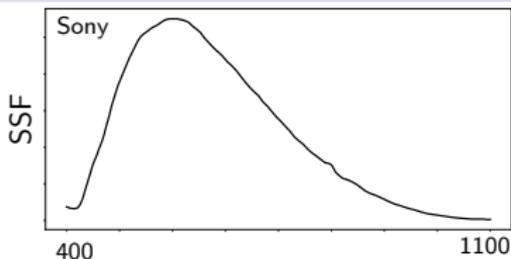
From CFA (raw) image to colour image : demosaicing



Image

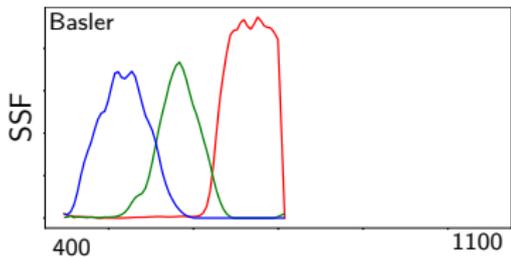
Panchromatic image :

1 channel associated to the Spectral Sensitivity Function (SSF) of one sensor.



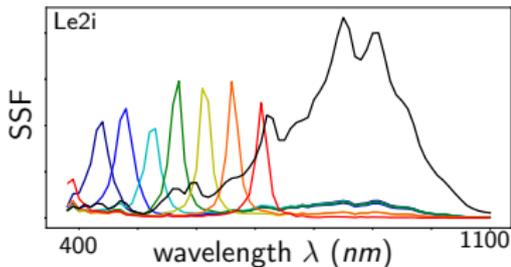
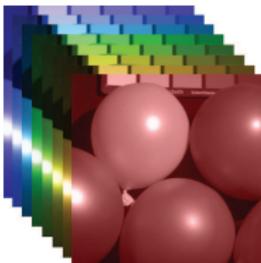
Color image :

3 color channels (R, G, B) in the visible spectral domain.

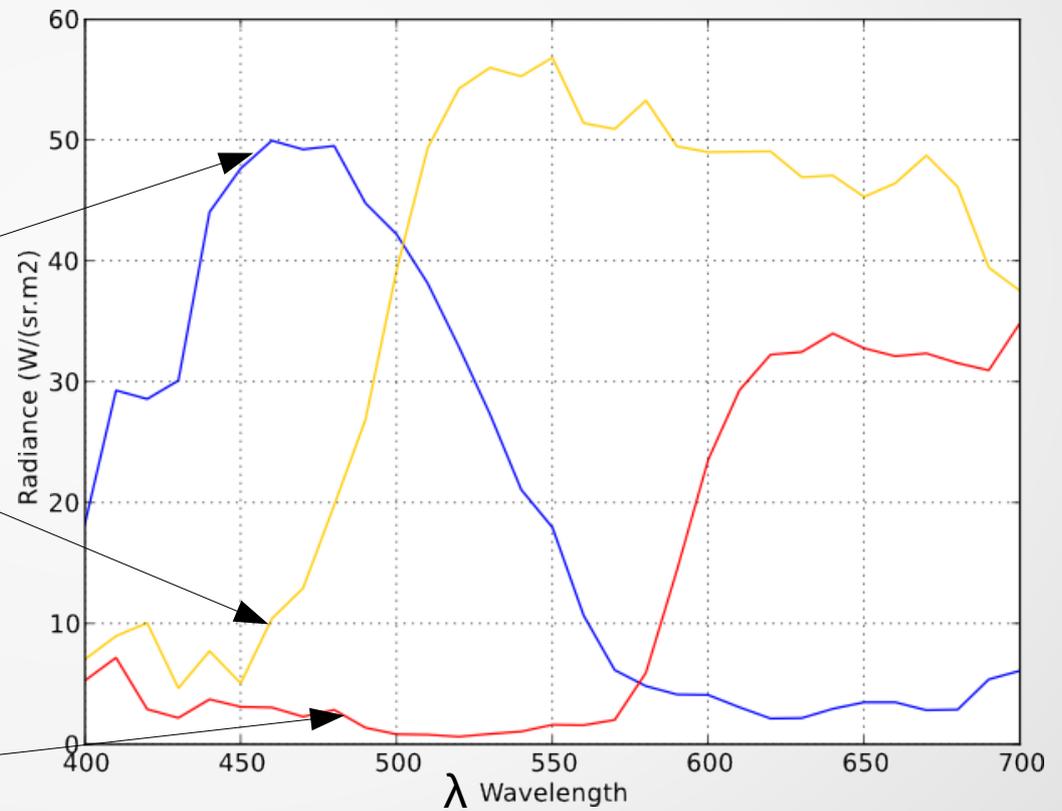


Multispectral (MS) image :

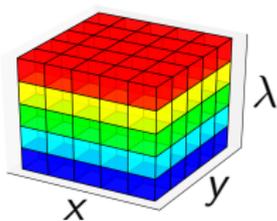
from 4 to 30 channels in the visible and near infra-red spectral domains.



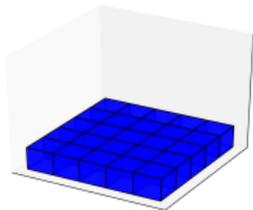
Reflectance spectrale et couleur



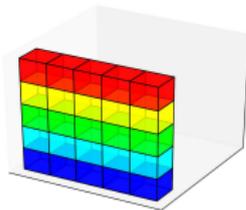
Multishot acquisition systems



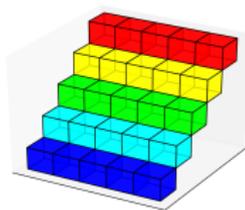
(a) Multispectral (MS) cube



(b) Tunable filter or illumination



(c) Push-broom

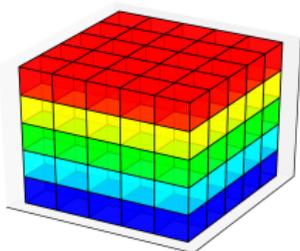


(d) Spatio-spectral linescan

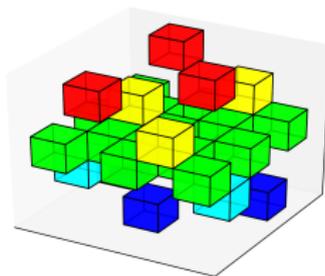
- Tunable filter or illumination (b)
 - One spectral channel is acquired at each time.
 - Spectral domain is scanned by changing the filter or illumination.
- *Push-broom* (c) or *spatio-spectral scan* (d)
 - One pixel row is acquired at each time.
 - Spatial plane is scanned by moving object or camera.

Multishot systems build cube from several acquisitions \Rightarrow still scene.

Snapshot acquisition devices



(a) Multi-sensors



(b) MSFA, raw image

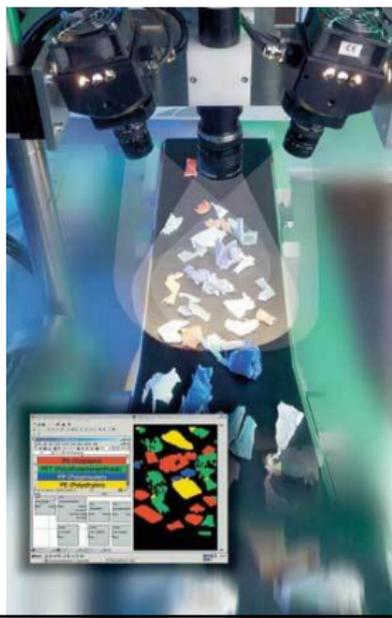
G	R	G	O	G
B	G	C	G	B
G	O	G	R	G
C	G	B	G	C
G	R	G	O	G

- Multi-sensor based technology (a)
 - All channels are simultaneously acquired.
 - Expensive and cumbersome - limited number of channels.
- MultiSpectral Filter Array (MSFA) based technology (b)
 - Spatio-spectral sampling of the incoming radiance according to the pixel location : **raw image**.
 - Fully-defined multispectral image is estimated by a demosaicing process.

Application domains

Recycling (PELLENC) :

Material identification ¹ .



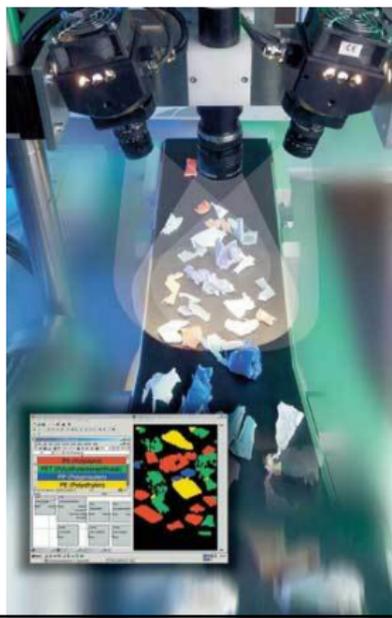
1. L. Jacquin *et al.*, Control of waste fragment sorting process. based on MIR imaging coupled with cautious classification, Resources, Conservation and Recycling, 2021.

2. M. Elias *et al.*, Multispectral camera ... to depict Leonardo's sfumato. Applied Optics, 2008.

Application domains

Recycling (PELLENC) :

Material identification ¹ .



Art and heritage :

Depict painting by pigment identification ² .



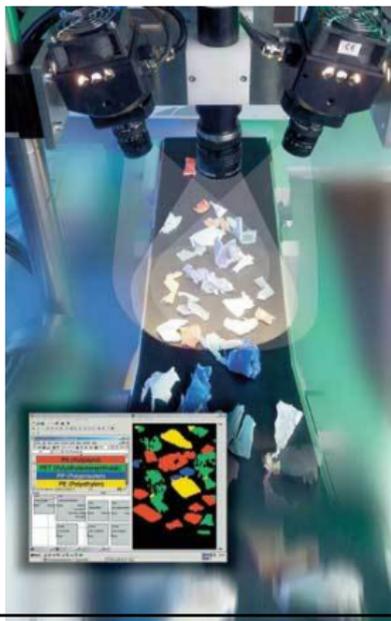
1. L. Jacquin *et al.*, Control of waste fragment sorting process. based on MIR imaging coupled with cautious classification, Resources, Conservation and Recycling, 2021.

2. M. Elias *et al.*, Multispectral camera ... to depict Leonardo's sfumato. Applied Optics, 2008.

Application domains

Recycling (PELLENC) :

Material identification ¹ .



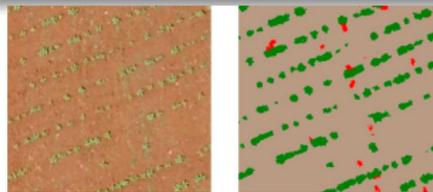
Art and heritage :

Depict painting by pigment identification ² .



Precision Agriculture :

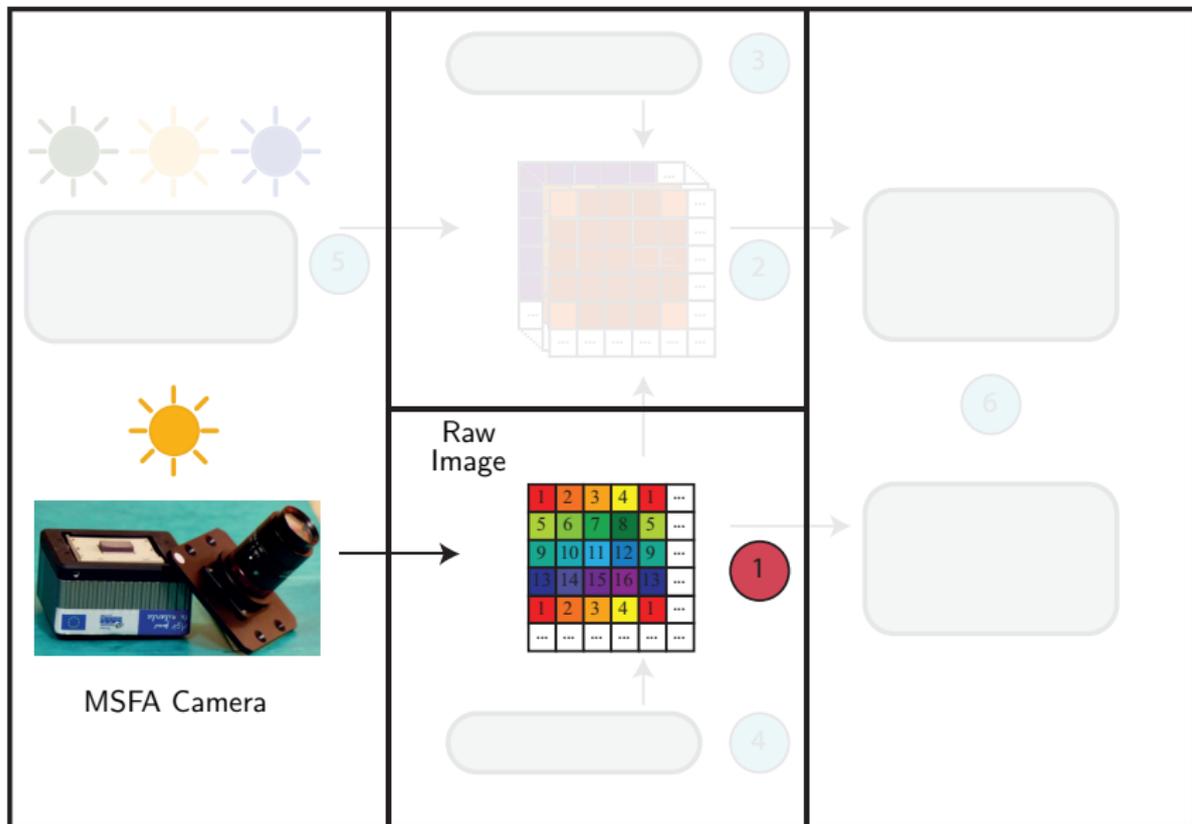
Weed detection.



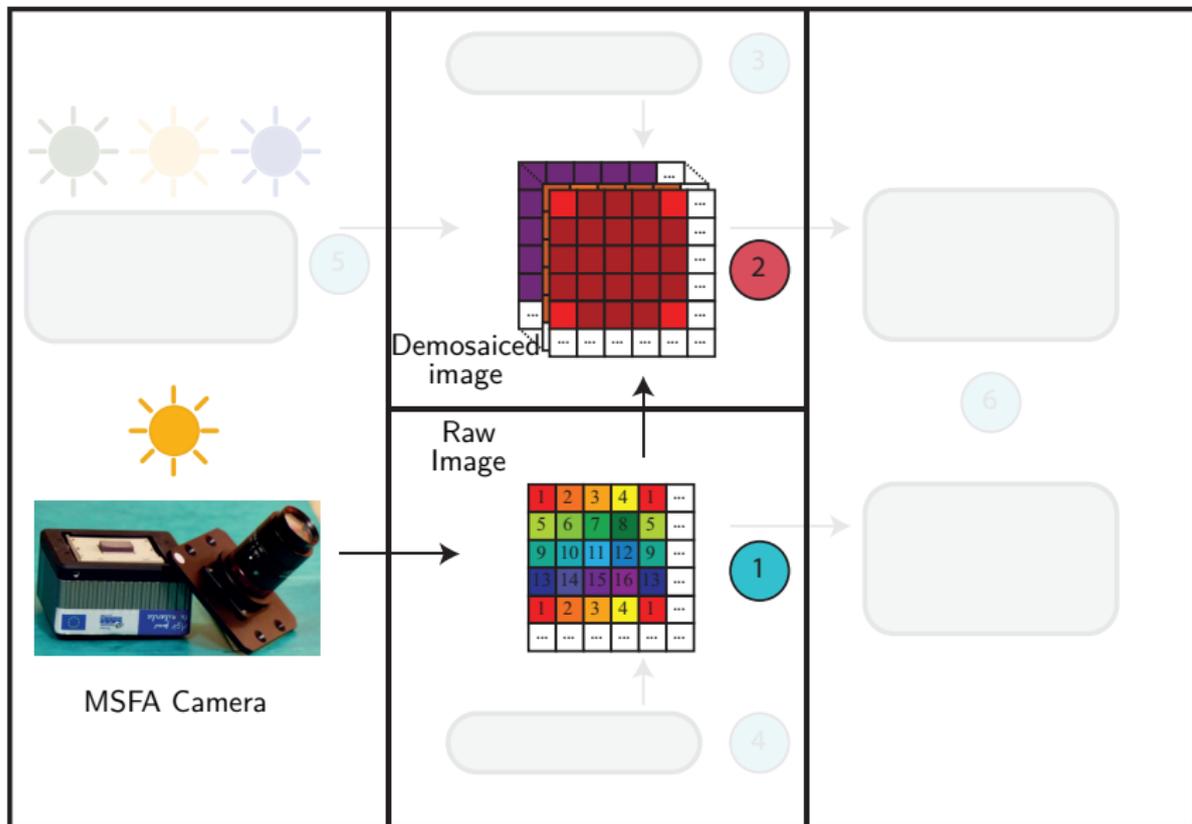
1. L. Jacquin *et al.*, Control of waste fragment sorting process. based on MIR imaging coupled with cautious classification, Resources, Conservation and Recycling, 2021.

2. M. Elias *et al.*, Multispectral camera ... to depict Leonardo's sfumato. Applied Optics, 2008.

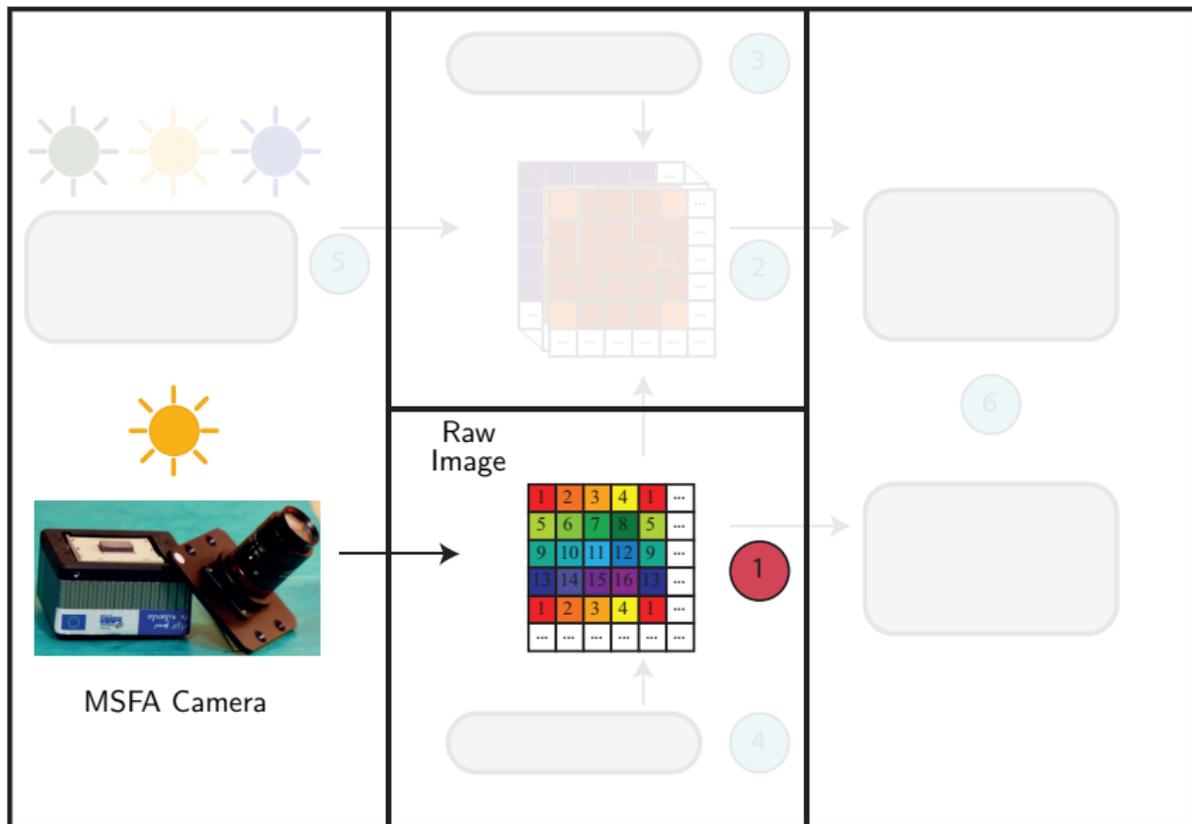
MS image acquired by a single-sensor camera



MS image demosaicing

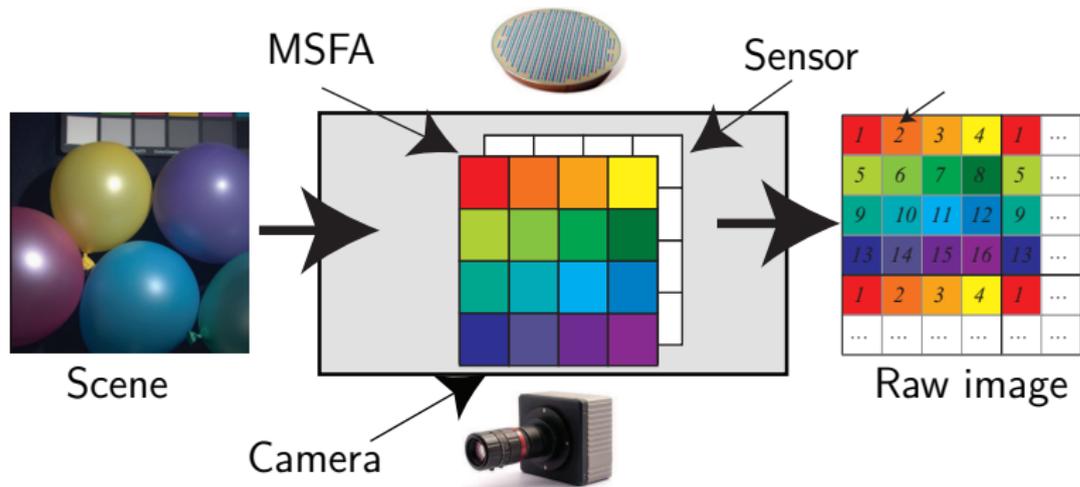


MS image acquired by a single-sensor camera



Snapshot mosaic camera

- Snapshot mosaic camera¹ (IMEC).
- MSFA represents a periodic 4×4 pattern composed of 16 spectral bands from 470nm to 630nm .



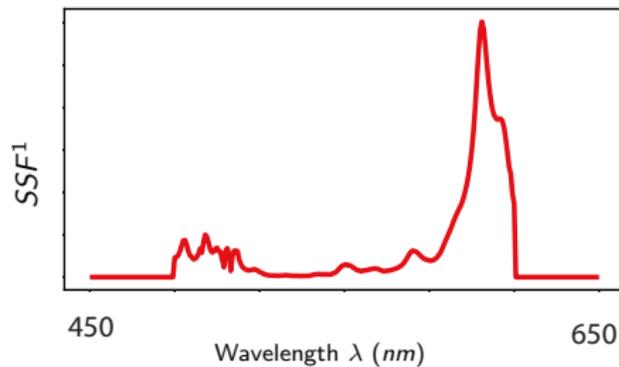
1. B.Geelen *et al.*, A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic, AFTMOP VII, 2014.

Raw image

...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...

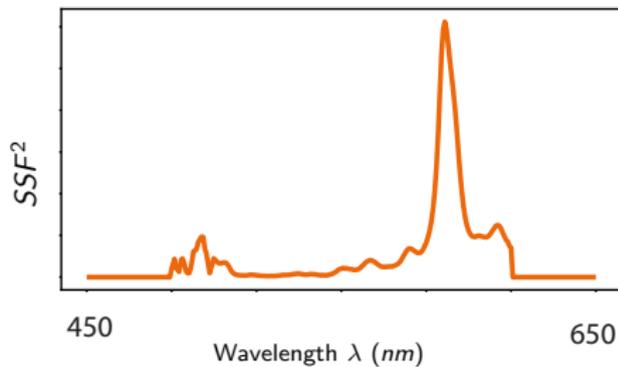
Raw image

...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...



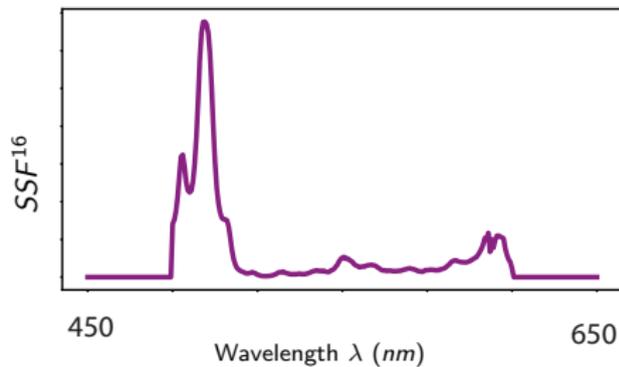
Raw image

...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...

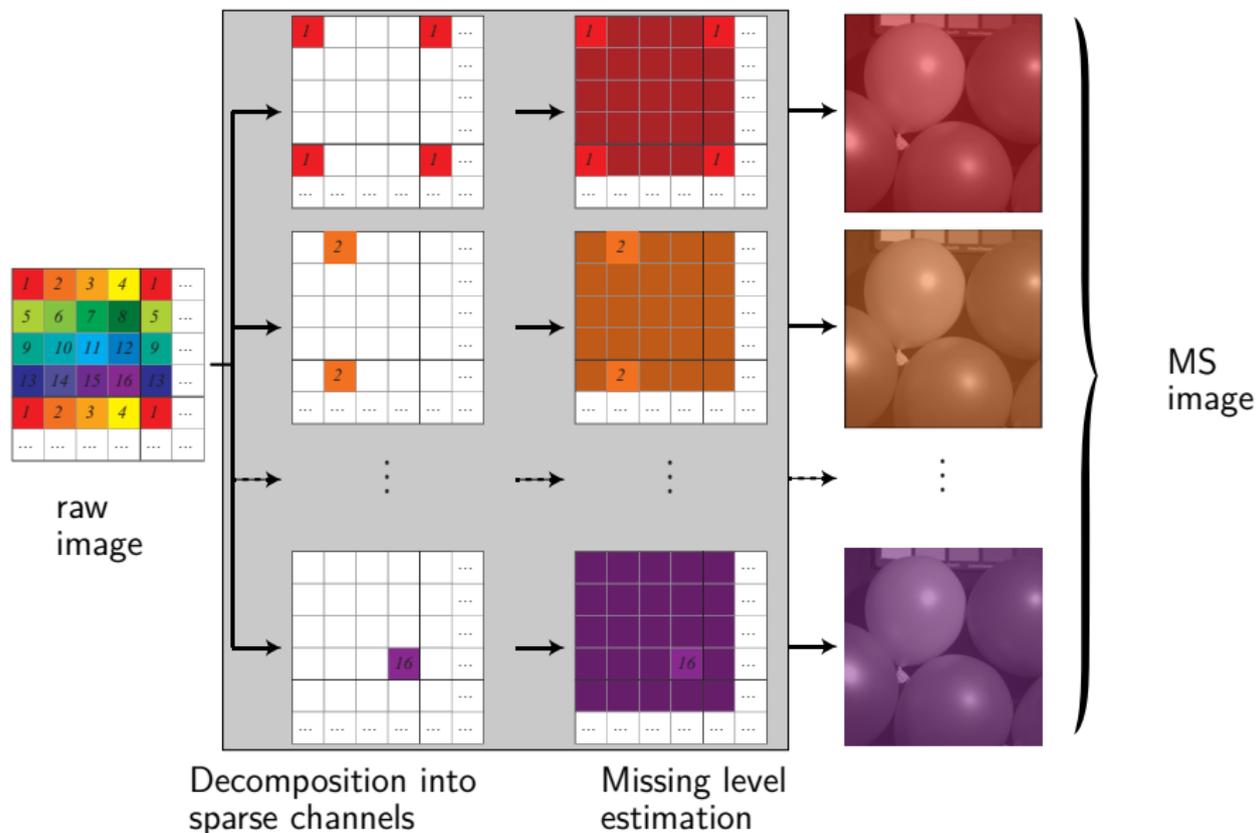


Raw image

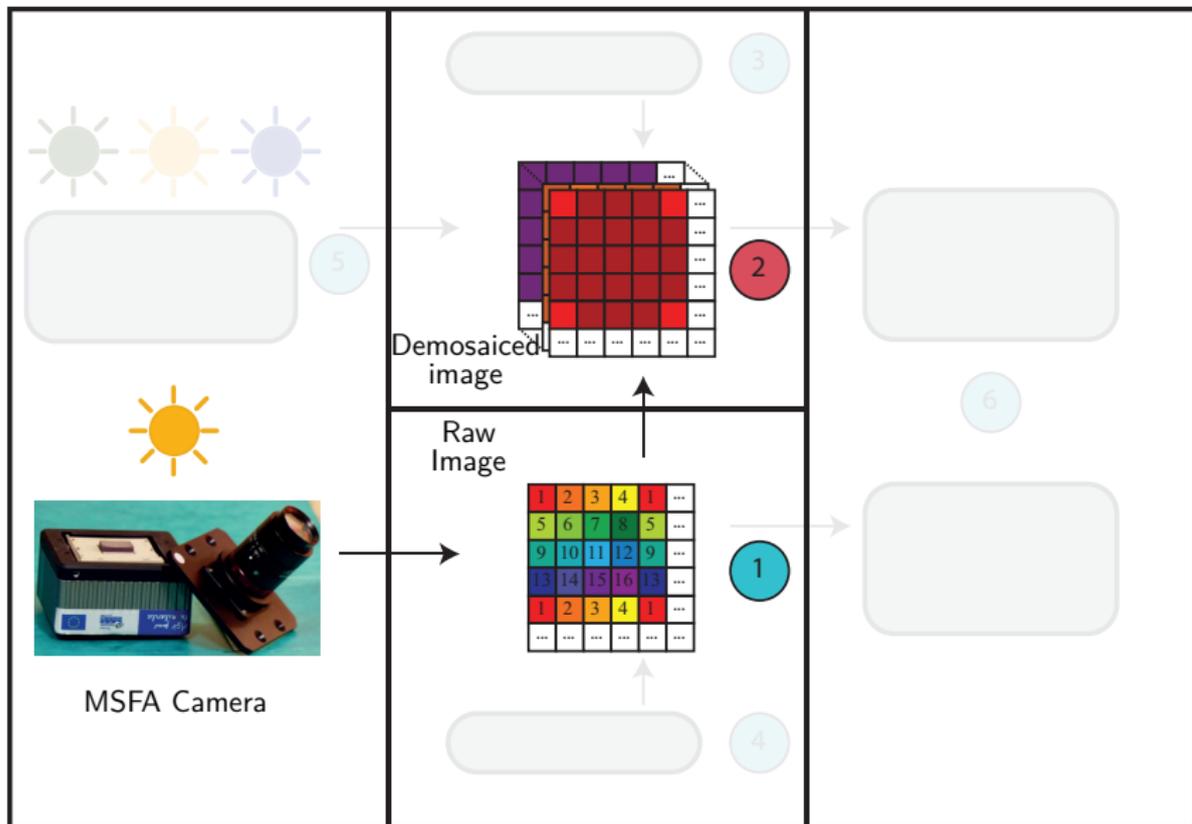
...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...	5	6	7	8	5	6	7	8	5	...
...	9	10	11	12	9	10	11	12	9	...
...	13	14	15	16	13	14	15	16	13	...
...	1	2	3	4	1	2	3	4	1	...
...



Demoisacing



MS image properties and demosaicing



Homogeneous regions and edges

Assumption

A natural image can be considered as set of **homogeneous regions** separated by **edges**.



Spatial correlation

Assumption :

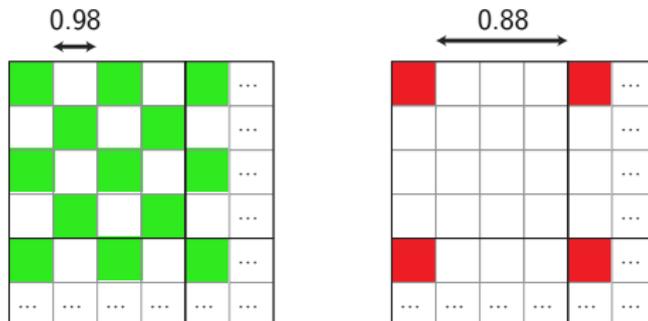
Within a channel, levels of neighbouring pixels in an homogeneous region are strongly correlated.

Demosaicing :

Pixel level can be estimated by analysing those of neighbouring pixels in the same homogeneous region.

Limits :

Probability to belong to the same region decreases with respect to pixel spatial distance.



Spectral correlation

Assumption :

Different channels are strongly correlated.

Demosaicing :

A channel can be estimated from the other ones.

Property :

Channels with nearby band centres are more correlated than those associated with distant centres. For example, blue channel is more correlated with cyan channel (0.95) than with red channel (0.82).



Spectral correlation

Assumption :

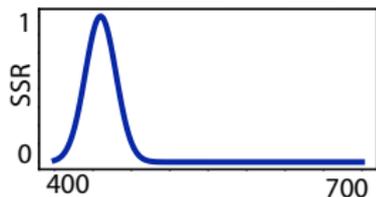
Different channels are strongly correlated.

Demosaicing :

A channel can be estimated from the other ones.

Property :

Channels with nearby band centres are more correlated than those associated with distant centres. For example, blue channel is more correlated with cyan channel (0.95) than with red channel (0.82).



Spectral correlation

Assumption :

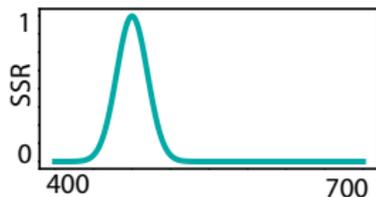
Different channels are strongly correlated.

Demosaicing :

A channel can be estimated from the other ones.

Property :

Channels with nearby band centres are more correlated than those associated with distant centres. For example, blue channel is more correlated with cyan channel (0.95) than with red channel (0.82).



Spectral correlation

Assumption :

Different channels are strongly correlated.

Demosaicing :

A channel can be estimated from the other ones.

Property :

Channels with nearby band centres are more correlated than those associated with distant centres. For example, blue channel is more correlated with cyan channel (0.95) than with red channel (0.82).

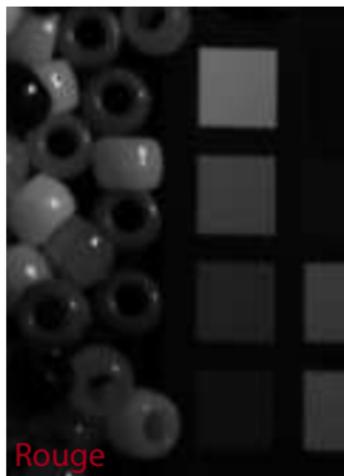


Image numérique et contrôle de l'aspect visuel

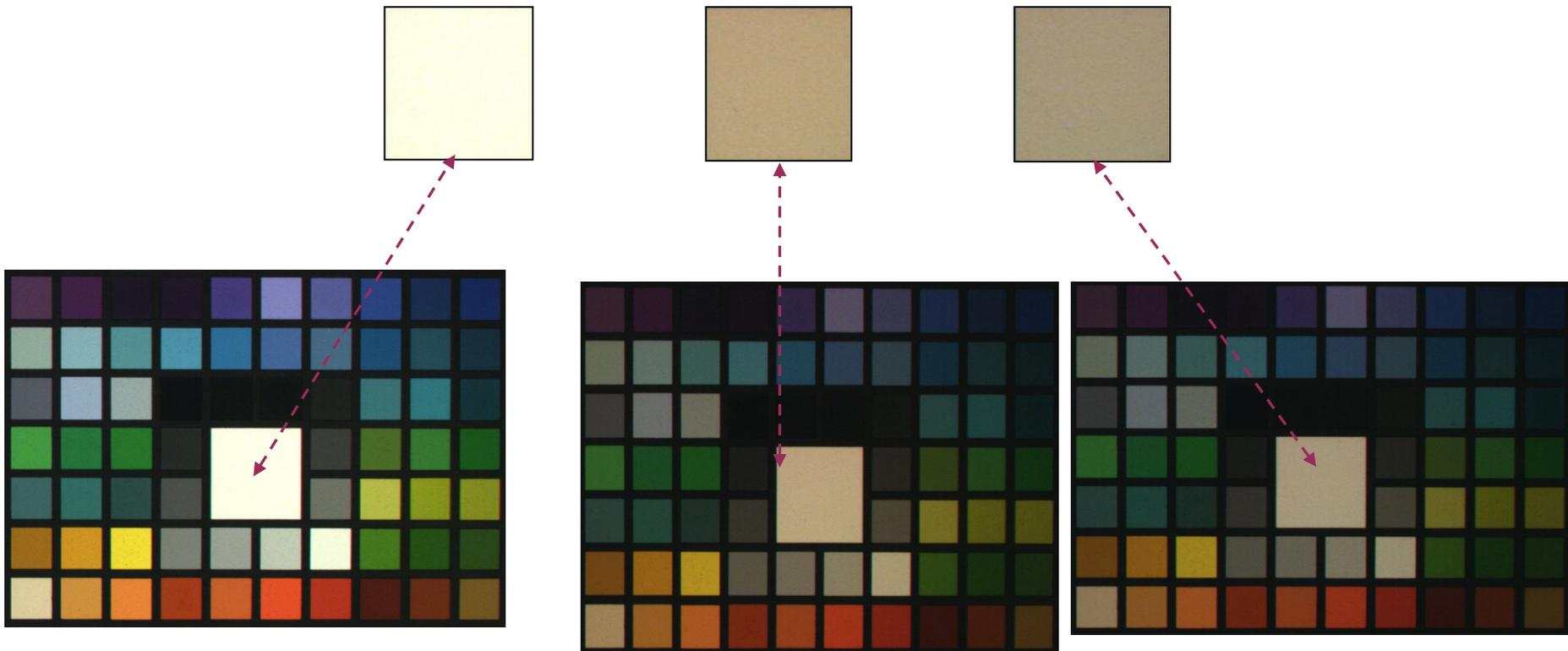


Image partielle de la mire de référence pour photo numérique de GretagMacbeth

Un même objet a été photographié dans trois conditions d'éclairage différentes

Valeurs RGB Constance des couleurs?

107 58 61



1

Ombre
du
vélo

70 46 62



2

122 92 190



3

Cadre
du
vélo

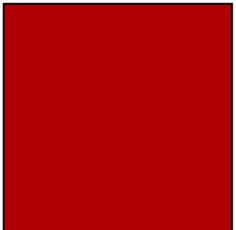
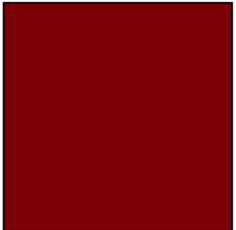
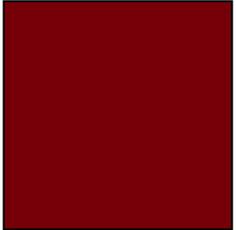
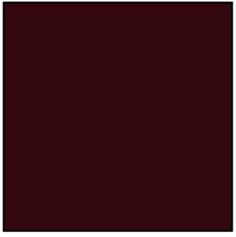
63 50 122



4

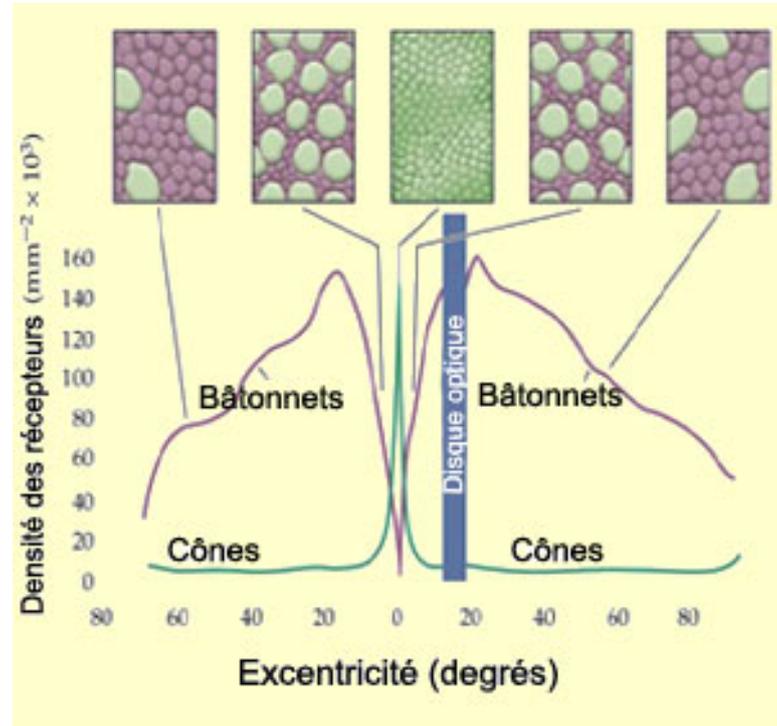
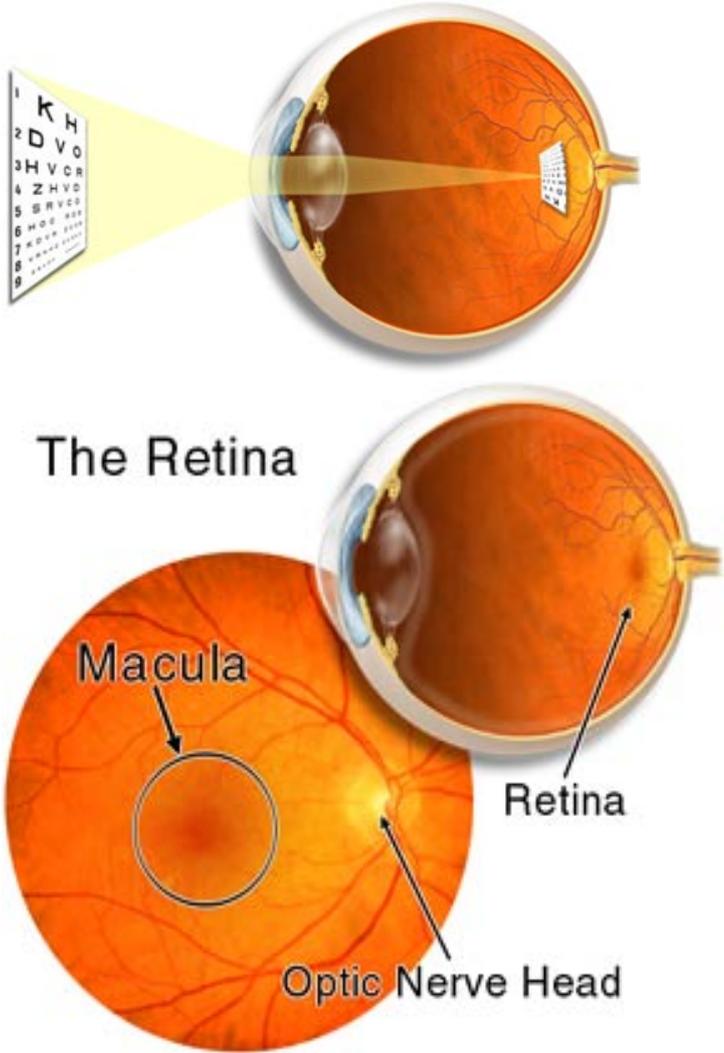


Constance des couleurs?



Perception de la couleur

Le fonctionnement de l'œil



Perception des couleurs – les cônes

Sensibilité maximale

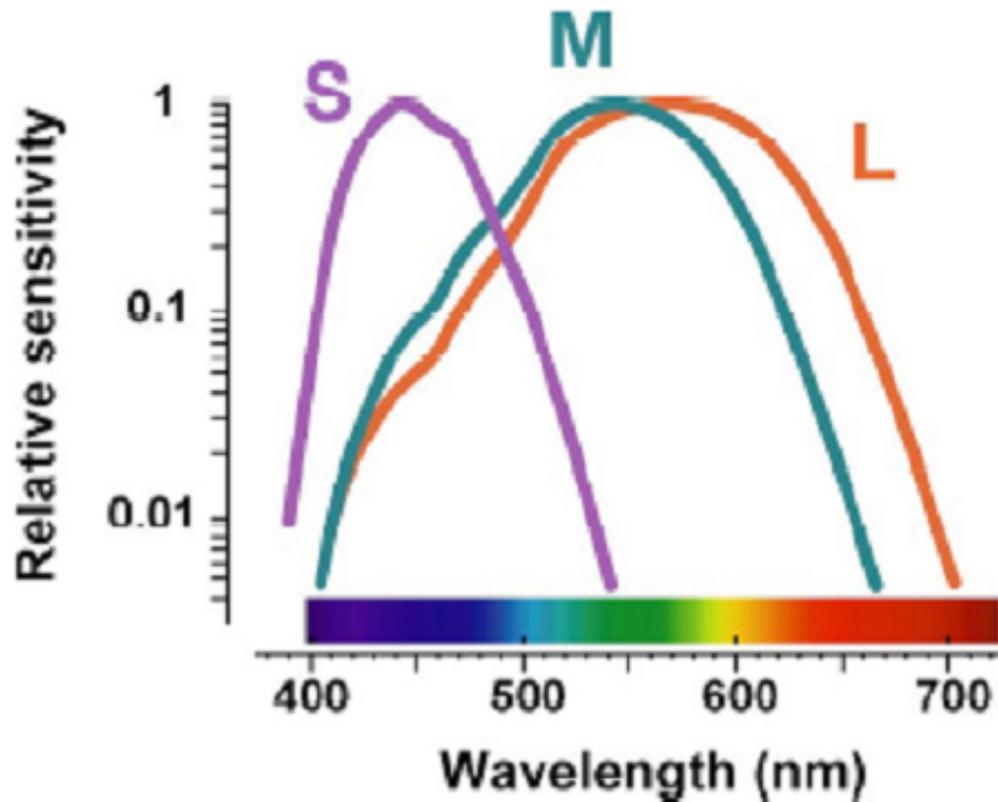
S : 440 nm

M : 535 nm

L : 565 nm

- Trois familles de cônes L, M, S (et non R, G, B)

- Large recouvrement des sensibilités des cônes L et M (30 nm)



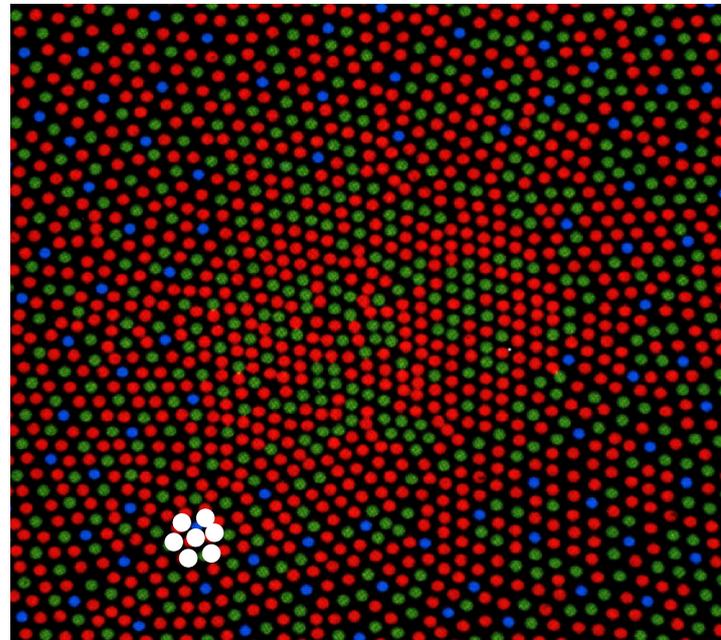
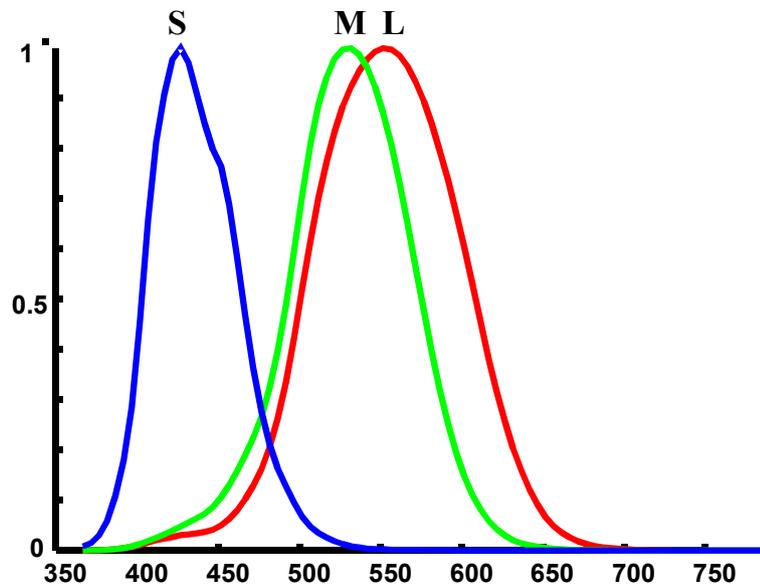
- Proportions relatives L:M très variables $\sim 0.25:01$ à $9:1$ (Roorda & Williams, 1999)

- Variation importante de la taille, de la densité optique des cônes entre fovéa et périphérie

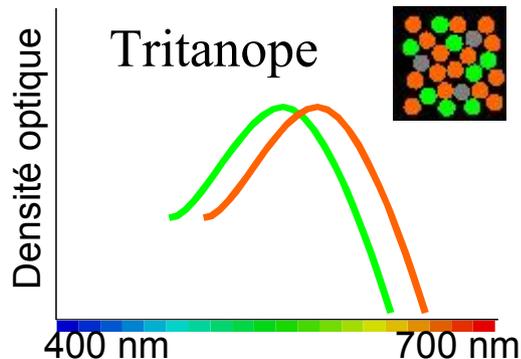
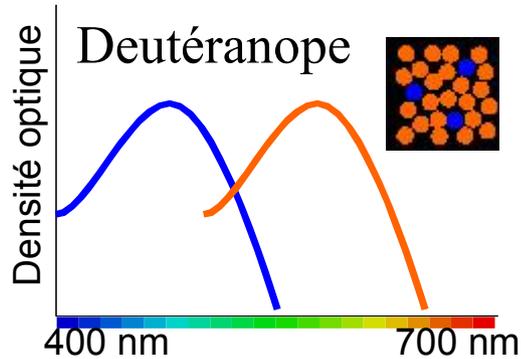
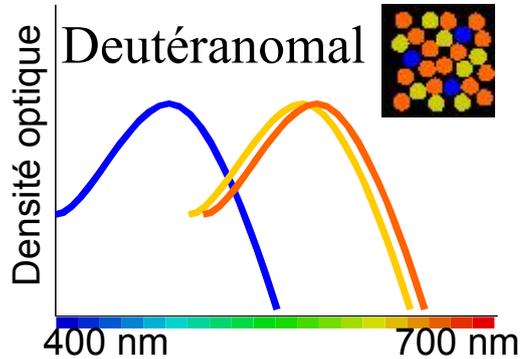
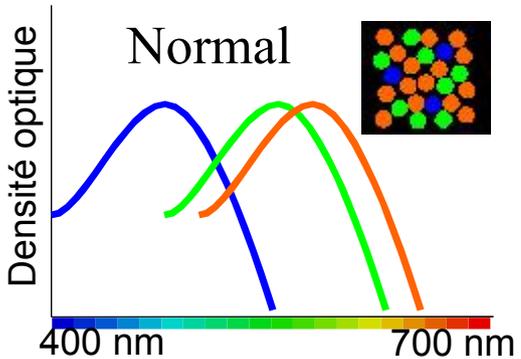
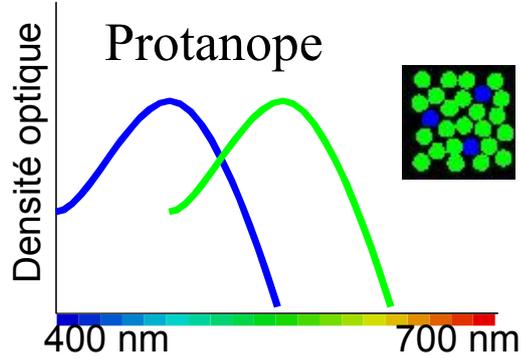
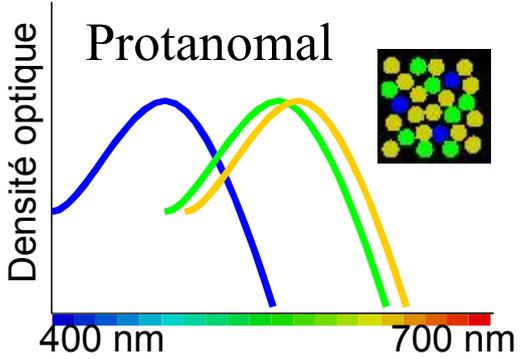
- Cônes S absents des 20' centrales, disposition selon un maillage parcimonieux ($\sim 10\%$) et régulier

The relative sensitivities of the Short- (S), Middle- (M), and Long- (L) wavelength sensitive cones as a function of wavelength. Each curve was normalized to its maximum. The curves show the cones sensitivity profiles derived by Stockman & Sharpe (2000).

Mosaïque de cônes



Vision normale et dyschromatopsies



Que voient les daltoniens ?

Protanope



Normal



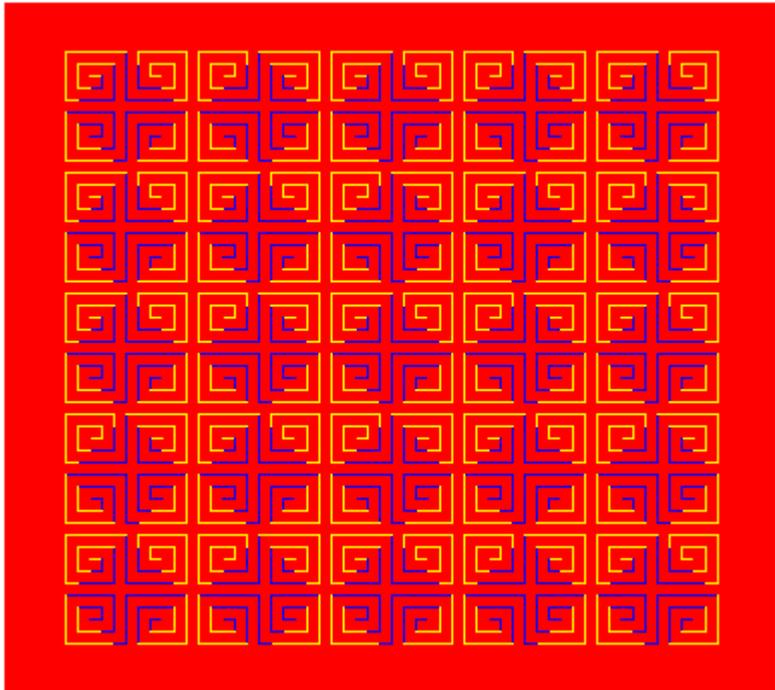
Deutéranope



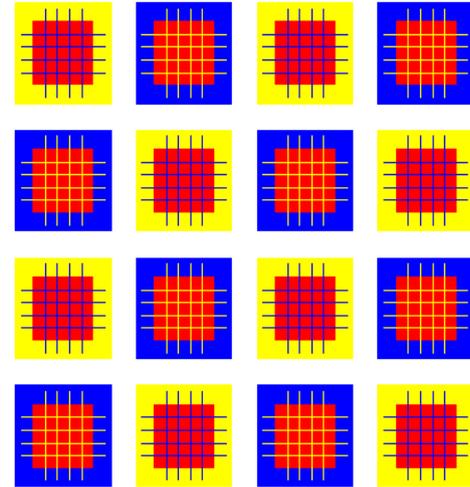
*Viénot, Brettel, Ott, Ben M'Barek, Mollon
(1995)*

www.tst.enst.fr/~brettel/

Perception – Illusion

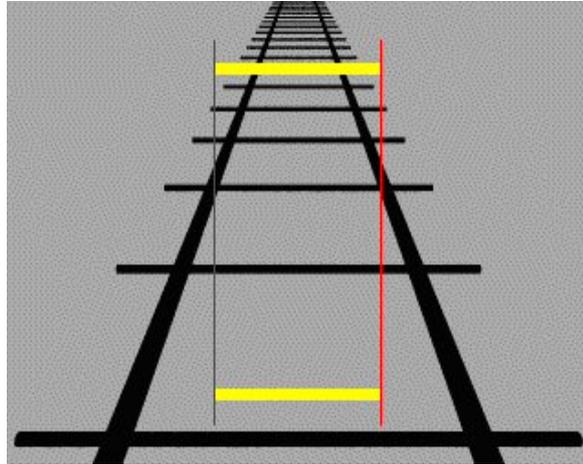


Assimilation
chromatique



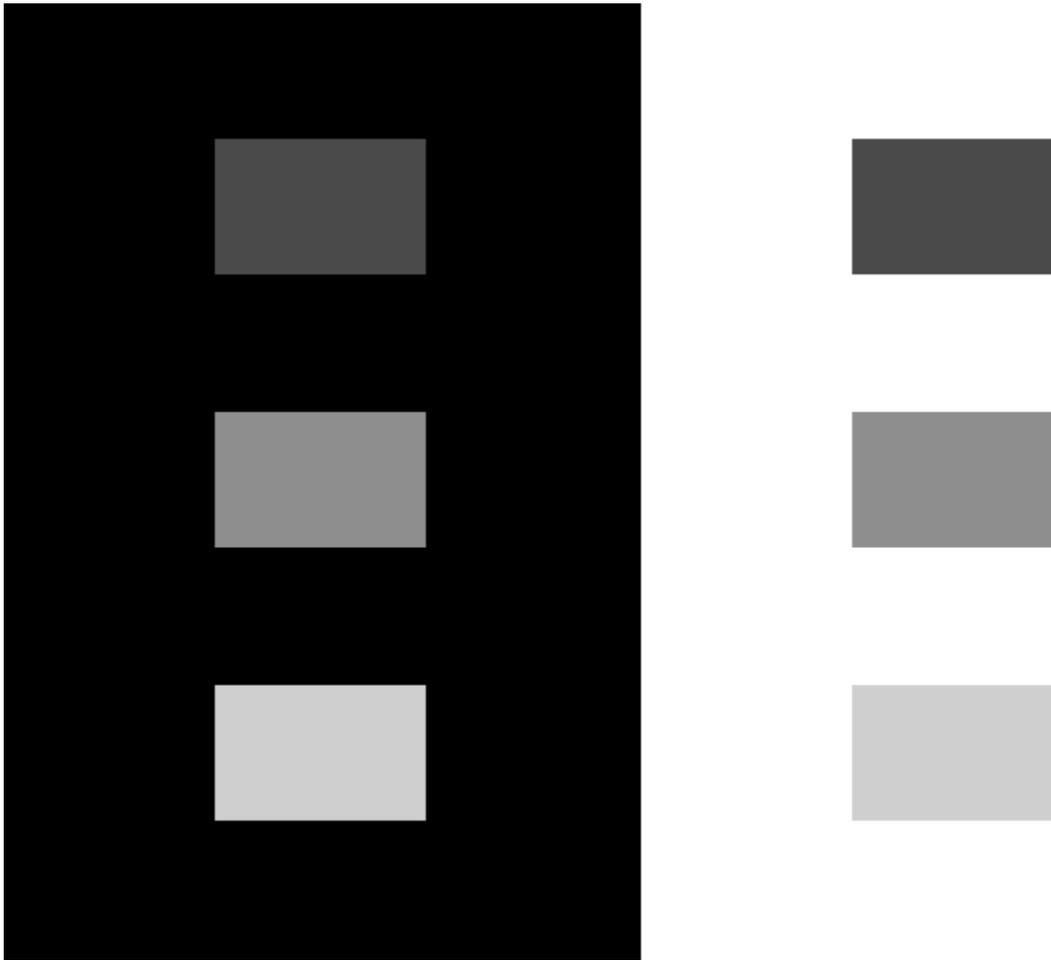
Le rouge est le même partout

Perception – Illusion (1)

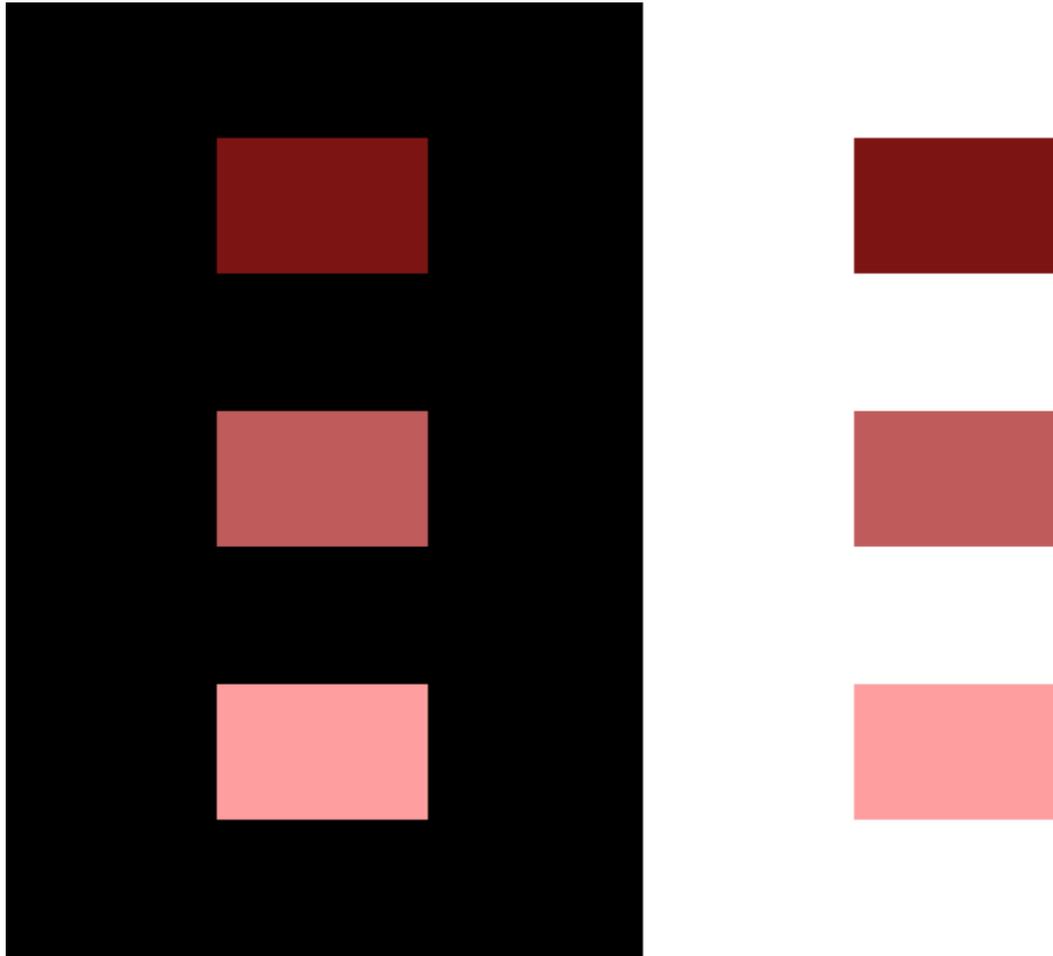


Perception – Illusion

Effet d'induction



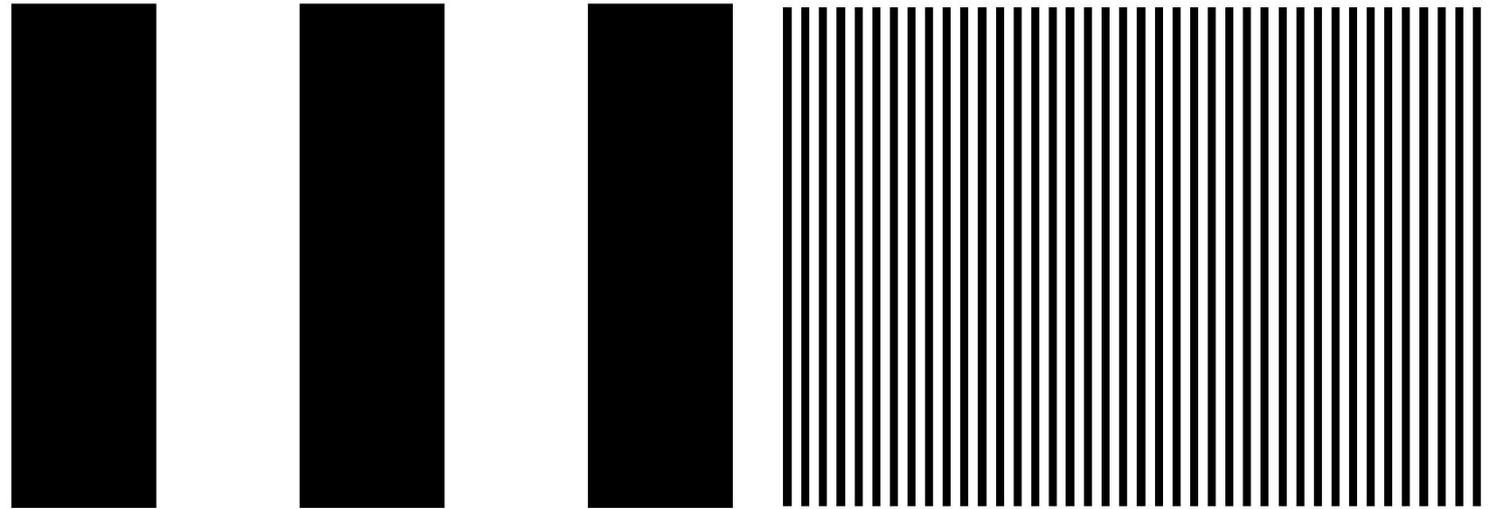
Ces carrés ont les mêmes valeurs de pixels à droite et à gauche . Mais ils apparaissent différents.



Ces carrés ont les mêmes valeurs de pixels à droite et à gauche. Mais ils apparaissent différents.

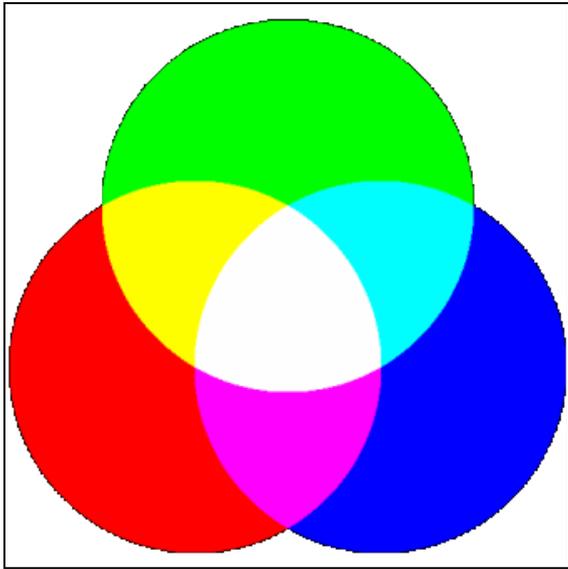
Sensibilité au contraste spatio-temporel

- La sensibilité au contraste diminue avec les hautes fréquences.



Elements de colorimétrie

Mélanges additifs et soustractifs

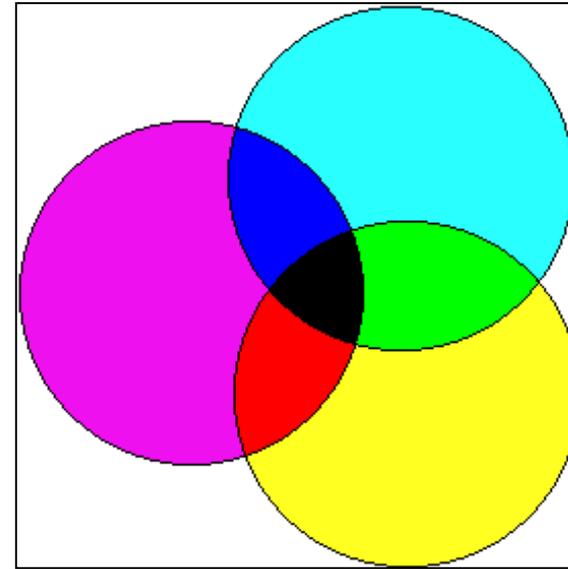


Combinaison de rayons lumineux
(émission)

mélange additif
 $\Sigma \lambda = \text{blanc}$



Télévision
rouge - vert - bleu



Combinaison de pigments
(absorption)

mélange soustractif
 $\Sigma \lambda = \text{noir}$

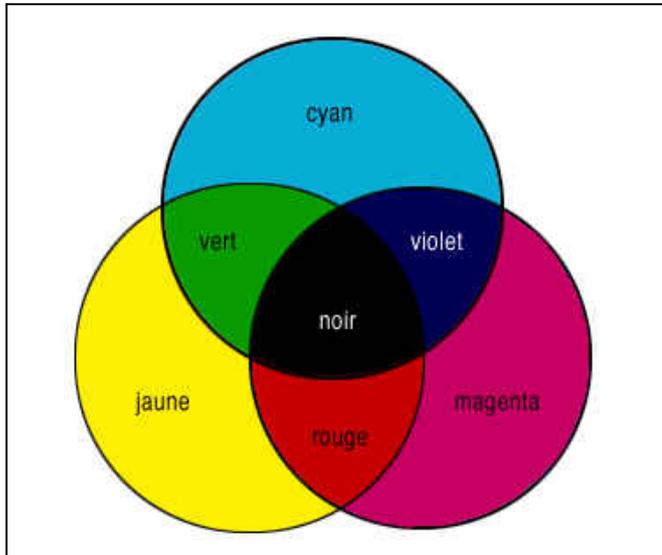


Imprimerie
cyan - magenta - jaune

J.C. Le Blon

3 couleurs

(rouge, jaune, bleu)



Imprimerie moderne

cyan - magenta - jaune

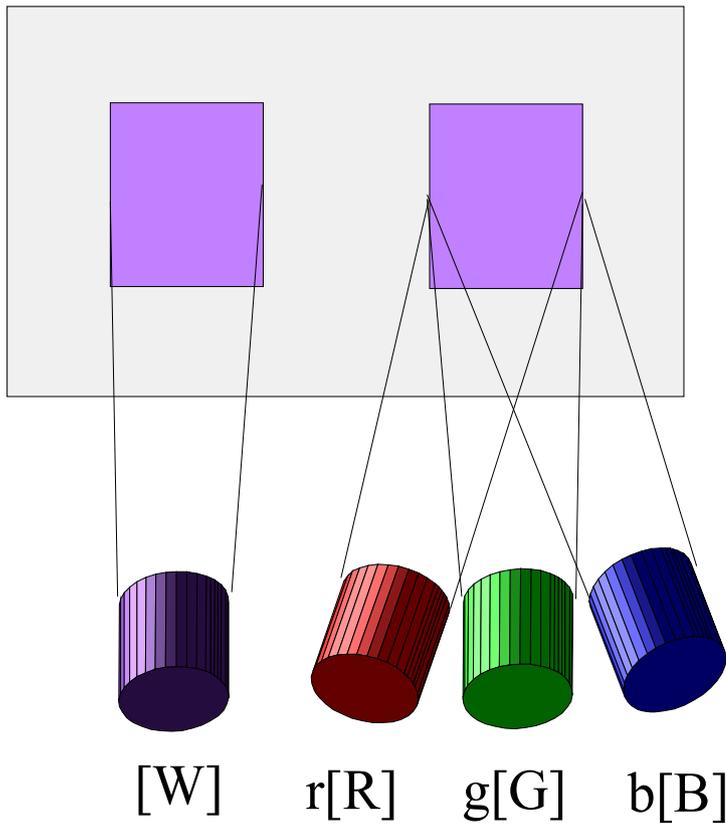
+

noir

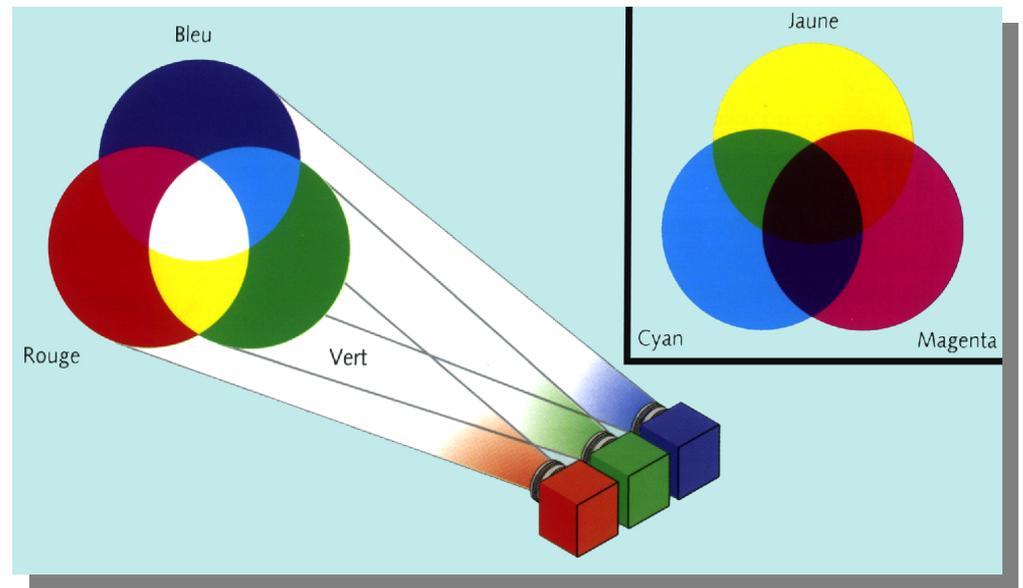
A sequence of four panels illustrating the printing process of a parrot. Panel 1 shows a cyan image with a green registration bar and label 'C'. Panel 2 shows a magenta image with a green registration bar and label 'C M'. Panel 3 shows a yellow image with a green registration bar and label 'C M Y'. Panel 4 shows the final color image with a green registration bar and label 'C M Y + K'.

- Trivalance de la perception des couleurs chez l'homme.

Expérience d'égalisation visuelle

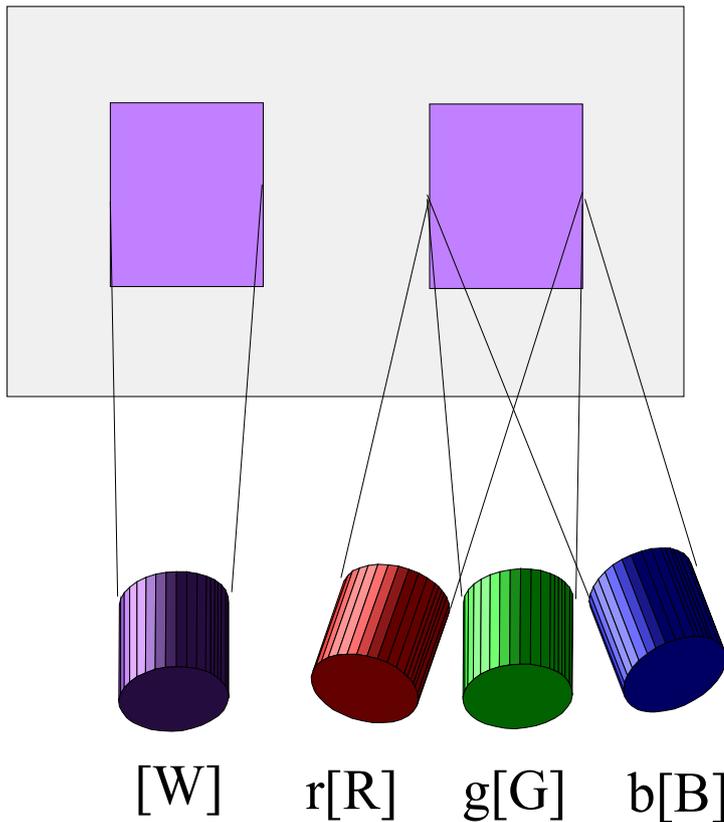


$$\text{Equivalence } [W] = r[R] + g[G] + b[B]$$



- Trivalance de la perception des couleurs chez l'homme.

Expérience d'égalisation visuelle



Equivalence $[W] = r[R] + g[G] + b[B]$

Loi de Grassman

Symétrie

$$[A]=[B] \Rightarrow [B]=[A]$$

Transistivité

$$[A]=[B], [B]=[C] \Rightarrow [A]=[C]$$

Proportionalité

$$[A]=[B] \Rightarrow k[A]=k[B]$$

Additivité

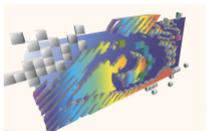
$$[A]=[B], [C]=[D] \Rightarrow [A]+[C]=[B]+[D]$$

→ Espace Vectoriel

La trivariance visuelle

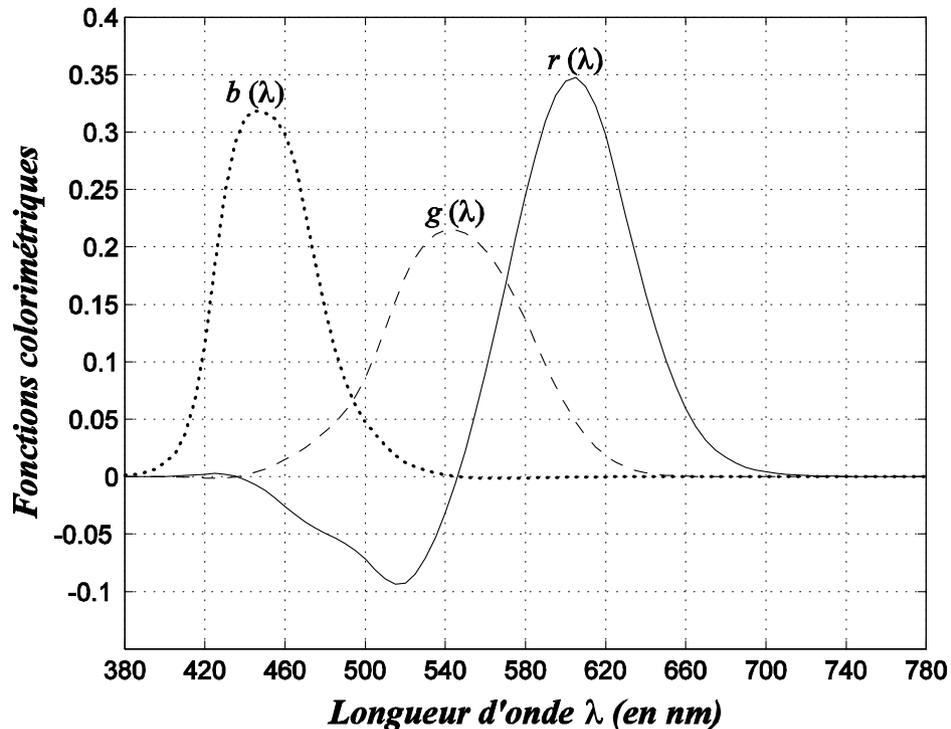


Simulation de la réponse des cônes : L (440 nm), M (545 nm) et S (580 nm)



Les différents systèmes de primaire

Le système RGB CIE 1931 est défini à partir de trois primaires monochromatiques : 700 nm, 546.1 nm et 435.8 nm (**Guild et Wright**). Il traduit le comportement de l'observateur de référence CIE 1931 à 2°.



$$\int_{\lambda_{min}}^{\lambda_{max}} \bar{r}(\lambda) d\lambda = \int_{\lambda_{min}}^{\lambda_{max}} \bar{g}(\lambda) d\lambda = \int_{\lambda_{min}}^{\lambda_{max}} \bar{b}(\lambda) d\lambda$$

Les quantités unitaires ont été normalisées de
que leur mélange unitaire soit
équi-énergétique. Ces
1,0000 pour R,

$$E(\lambda) = 1 = \bar{r}(\lambda).R + \bar{g}(\lambda).G + \bar{b}(\lambda).B$$

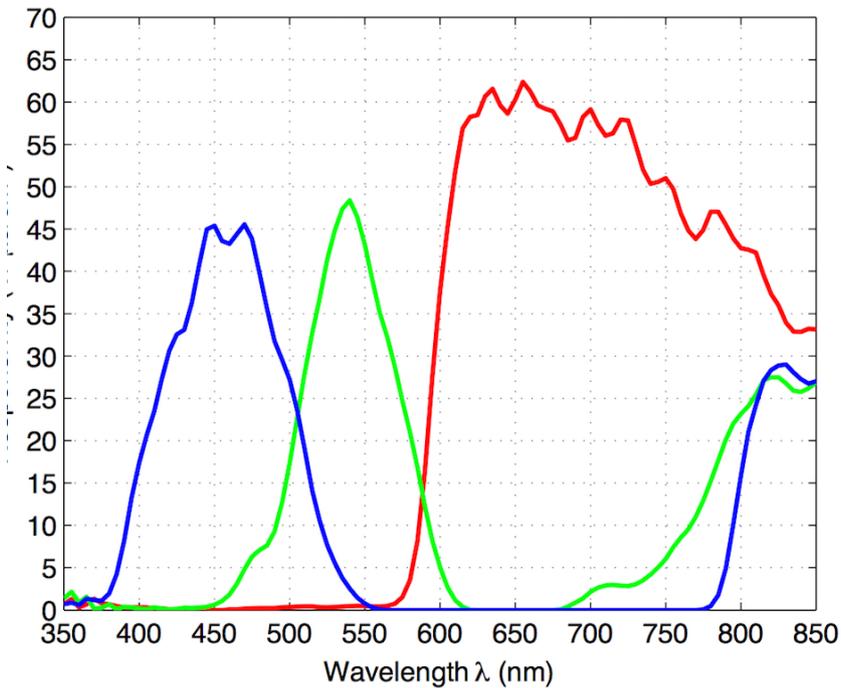
Fonctions colorimétriques du système de primaires RGB de l'observateur de référence CIE 1931 à 2°.

Nb. La courbe $r(\lambda)$ présente une partie négative, ce qui contredit le principe de la synthèse additive, d'où :

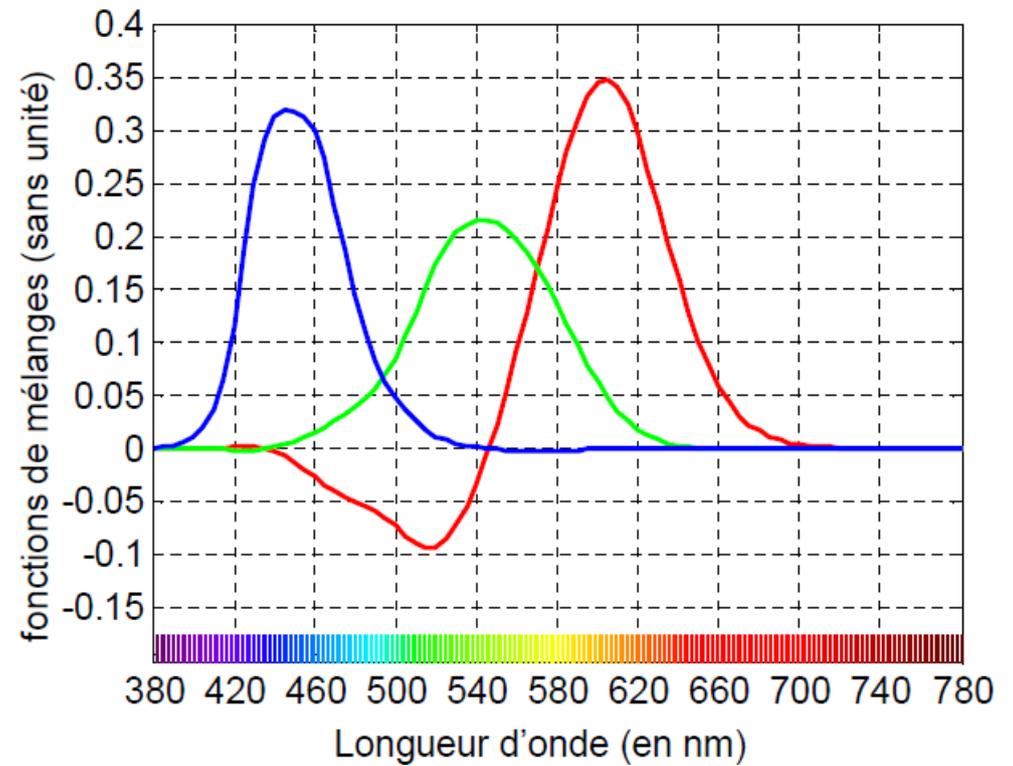
$$S(\lambda_i) + r(\lambda_i) R \equiv g(\lambda_i) G + b(\lambda_i) B \quad \Rightarrow \quad S(\lambda_i) \equiv -r(\lambda_i) R + g(\lambda_i) G + b(\lambda_i) B$$

Sensibilités spectrales

Caméra couleur



Observateur moyen

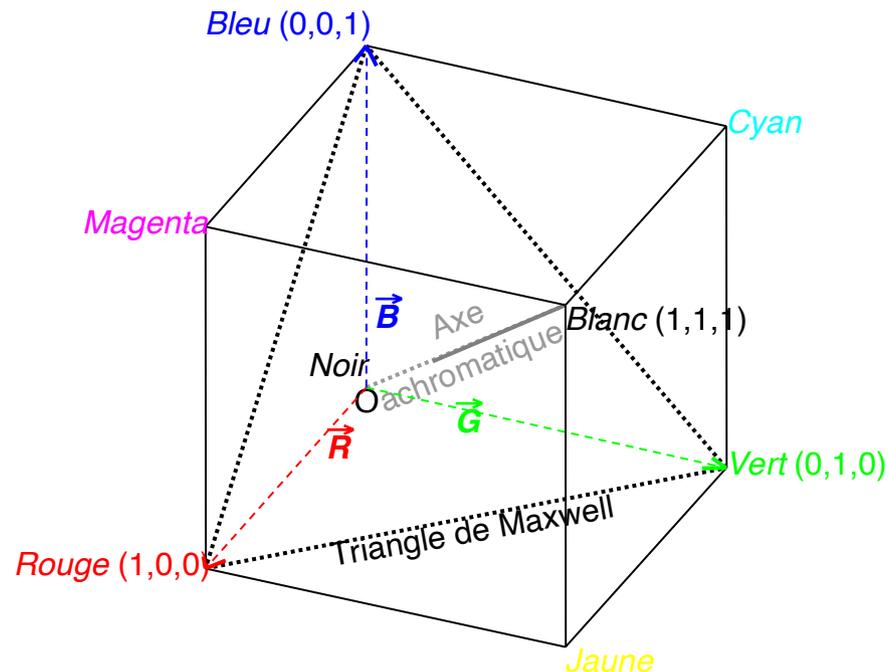


Sensibilité spectrale des photo-récepteurs

Fonction colorimétrique RVB

Systemes (R,G,B)

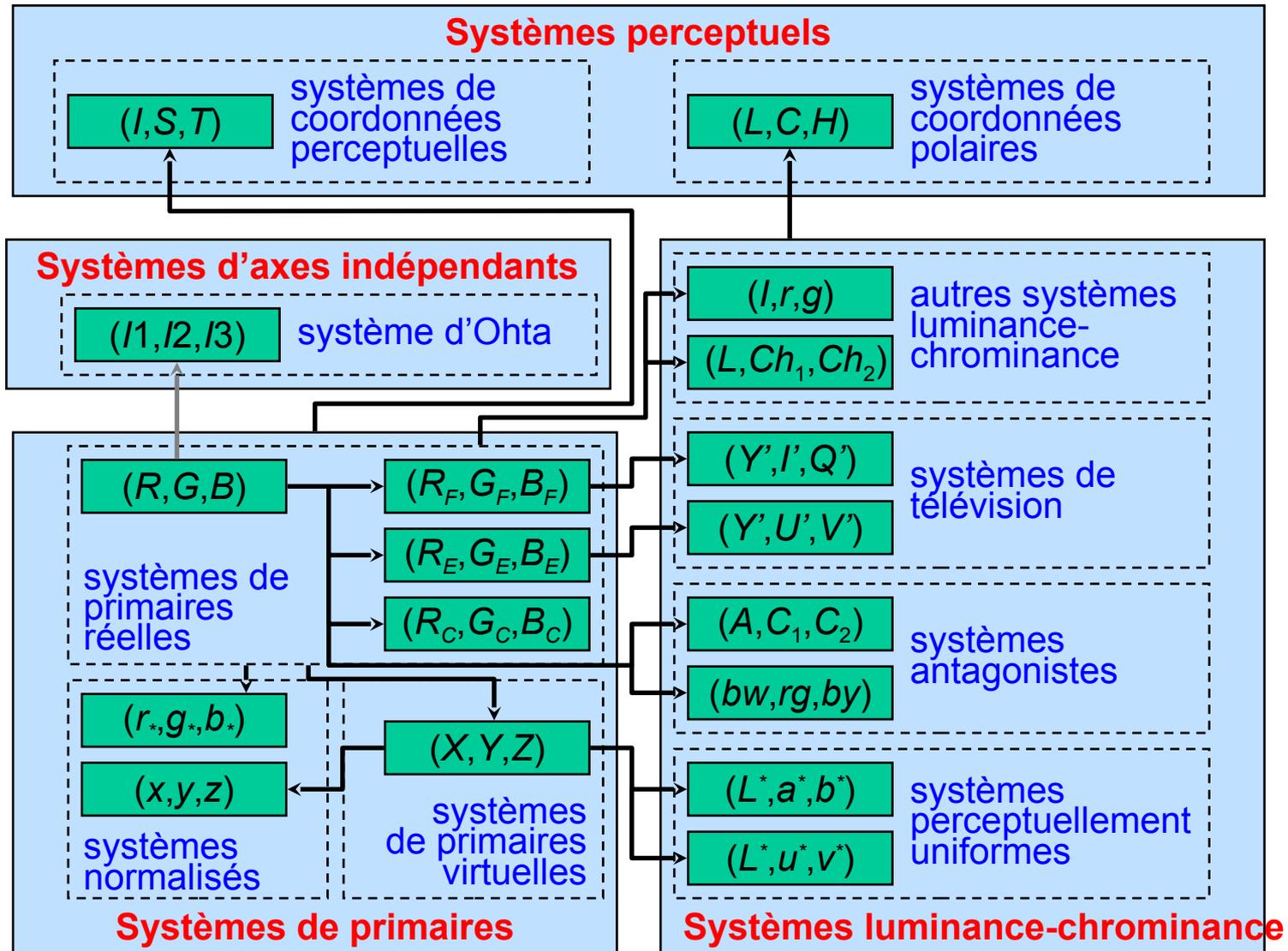
- Cube des couleurs
 - Les primaires correspondent à des vecteurs
 - Les composantes trichromatiques sont les coordonnées d'un point



Systemes (R, G, B)

- Inconvénients des systèmes (R, G, B)
 - Valeurs négatives des composantes trichromatiques
 - Valeurs des composantes trichromatiques liées à la luminance et corrélées
 - Multitude de systèmes (R, G, B)
 - Pour les écrans à tubes cathodiques ou LCD
 - Pour la colorimétrie (observateur de référence)
 - Pour les caméras
 - Pour les scanners
 - Pour les imprimantes (modèle CMYK)

Espaces couleur



Les coordonnées trichromatiques RGB

Aux trois primaires R, G, B on peut faire correspondre trois vecteurs directeurs normés, lesquels vecteurs forment alors un espace vectoriel à trois dimensions, appelé l'espace RGB, d'origine O.

Toute couleur peut alors être décrite par :

$$\vec{C} = R.\vec{R} + G.\vec{G} + B.\vec{B}$$

Les coordonnées trichromatiques RGB s'expriment alors de la façon suivante :

$$R = \int_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{r}(\lambda)d\lambda, \quad G = \int_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{g}(\lambda)d\lambda$$

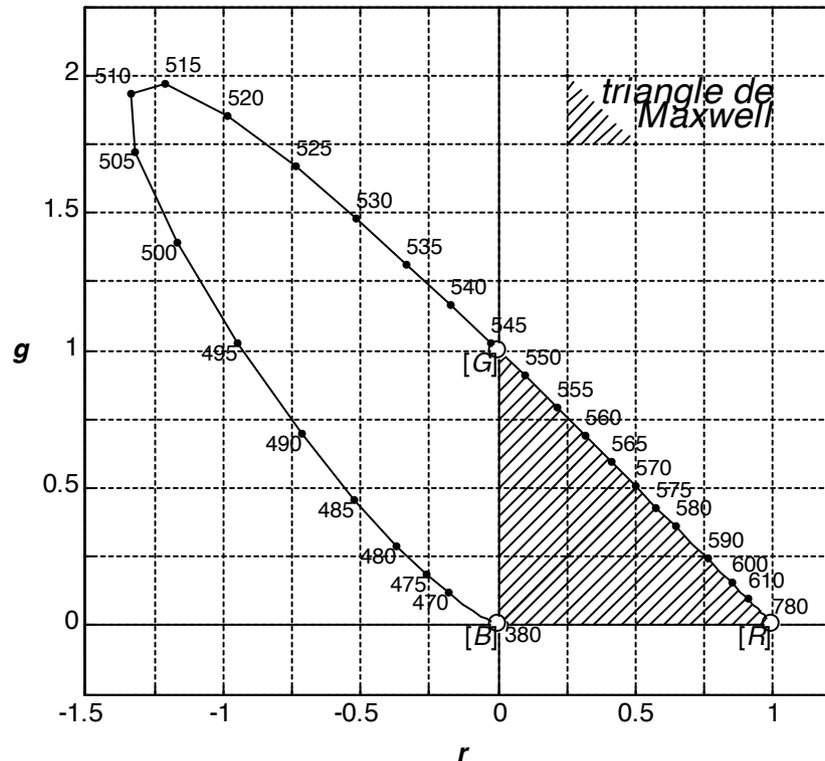
$$B = \int_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{b}(\lambda)d\lambda$$

Systemes (R,G,B)

- Diagramme de chromaticité
 - Composantes normalisées ou coordonnées trichromatiques (ou réduites)

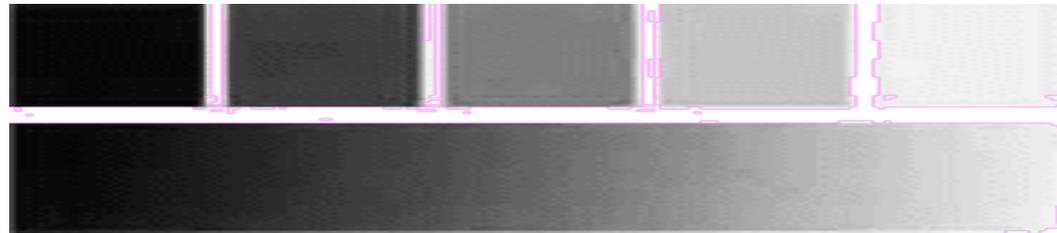
$$\left\{ \begin{array}{l} r = \frac{R}{R+G+B}, \\ g = \frac{G}{R+G+B}, \\ b = \frac{B}{R+G+B}. \end{array} \right.$$

Deux stimuli de même chrominance, mais de luminances différentes, peuvent avoir des coordonnées trichromatiques différentes.

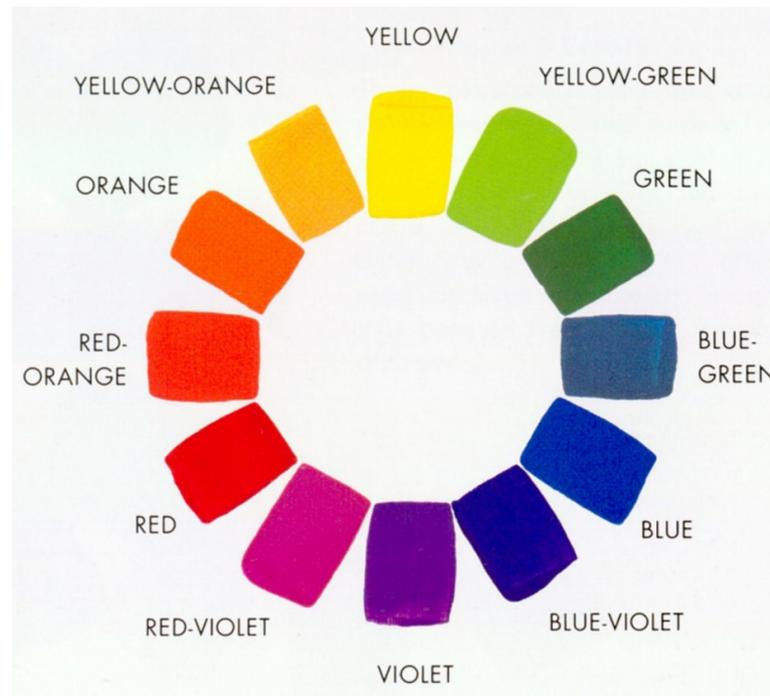


La couleur, c'est quoi ?

- Clarté



- Teinte

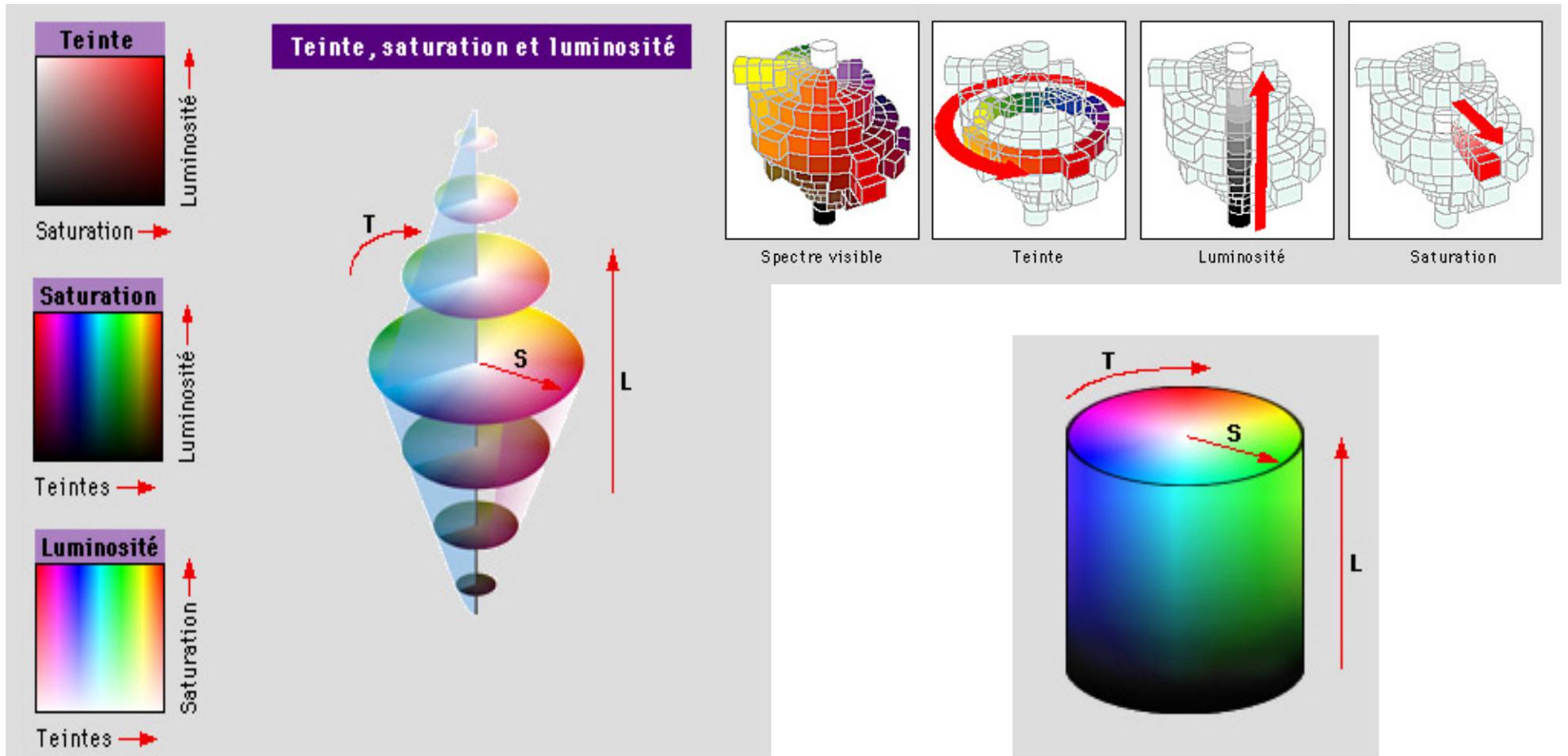


- Saturation



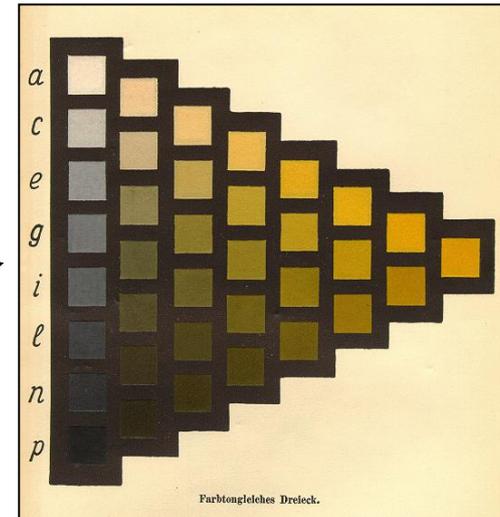
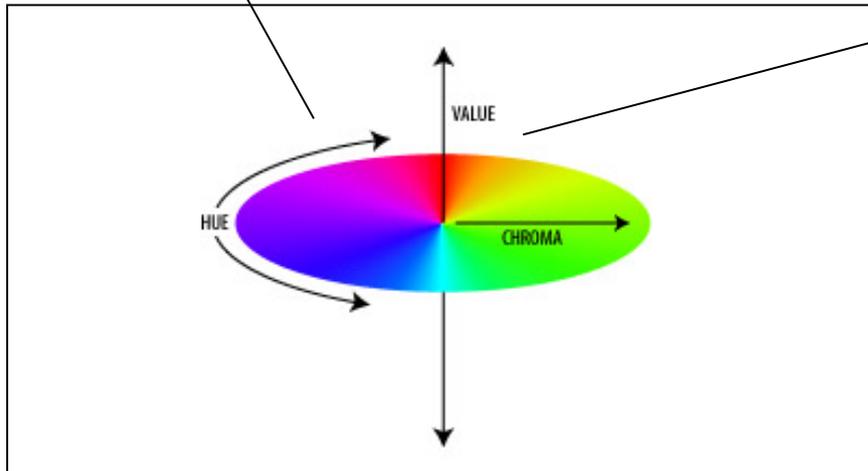
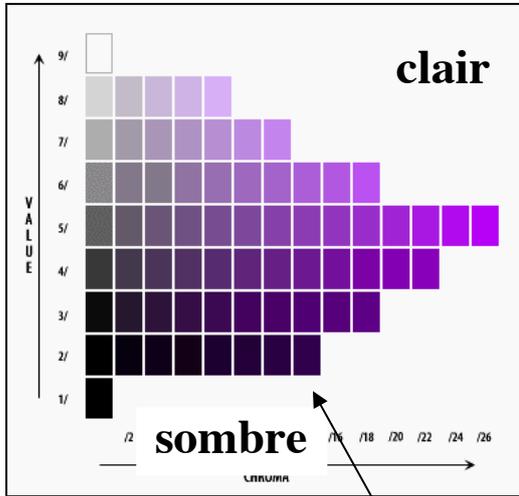
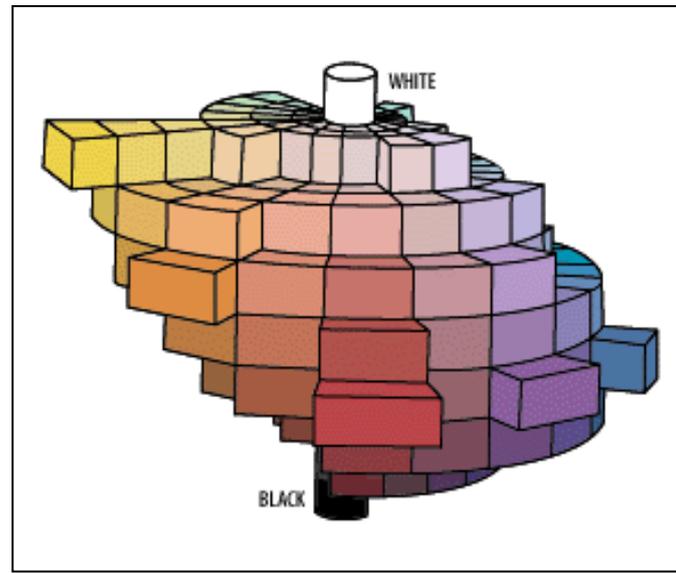
- Il faut trois quantités pour caractériser complètement une couleur !
- Dans la pratique, ces paramètres sont liés, et il n'est pas toujours facile de les distinguer.

La couleur des objets



Atlas de Munsell (1915)

ensemble de planches 2D obtenues par coupe 3D



Atlas de Munsell

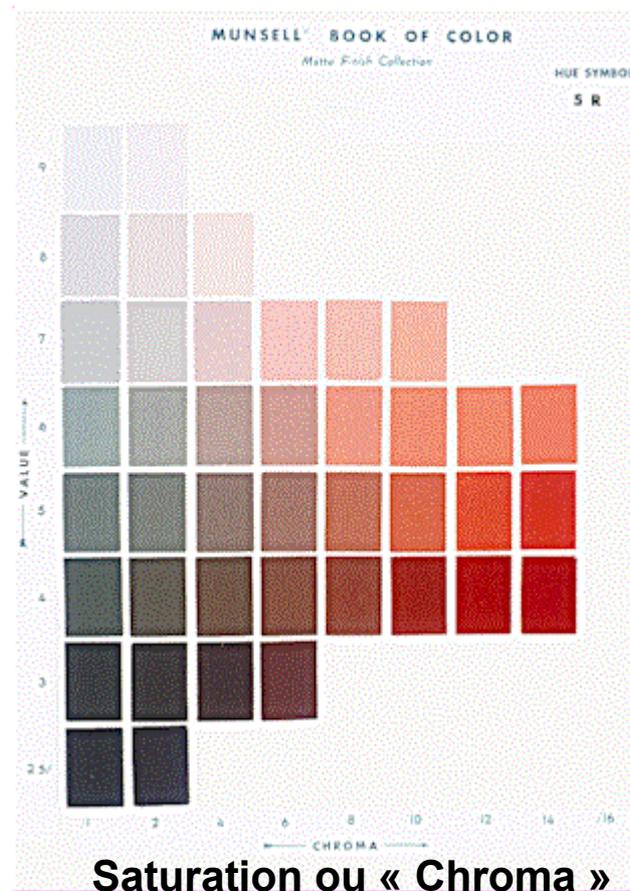
- A partir d'estimations « d'équidistance perceptive », subdivision fine de la roue des couleurs donnant les pages de l'atlas :

Une page = 1 teinte ou « Hue »

Actuellement : environ 1500 échantillons dans l'atlas

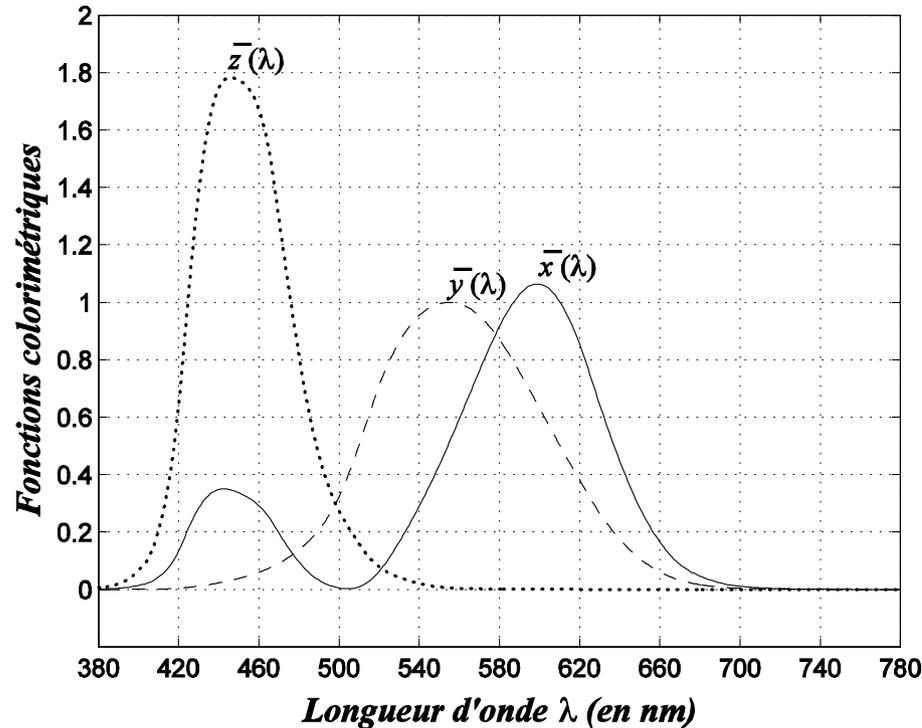
Un coloriste expérimenté distingue environ 10^6 couleurs !!

Clarté ou « Value »



Les différents systèmes de primaire

Les systèmes de primaire XYZ CIE 1931 (à 2°) et XYZ CIE 1964 (à 10°)



$$\int_{\lambda_{min}}^{\lambda_{max}} \bar{x}(\lambda) d\lambda = \int_{\lambda_{min}}^{\lambda_{max}} \bar{y}(\lambda) d\lambda = \int_{\lambda_{min}}^{\lambda_{max}} \bar{z}(\lambda) d\lambda$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.7690 & 1.7518 & 1.1300 \\ 1.0000 & 4.5907 & 0.0601 \\ 0.0000 & 0.0565 & 5.5943 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Fonctions colorimétriques du système de primaire XYZ de l'observateur de référence CIE 1931 à 2°.

Nb. Y représente explicitement la luminance de la couleur.

Les coordonnées trichromatiques XYZ

Aux trois primaires X, Y, Z on peut faire correspondre trois vecteurs directeurs normés, lesquels vecteurs forment alors un espace vectoriel à trois dimensions, appelé l'espace XYZ, d'origine O.

Toute couleur peut alors être décrite par :

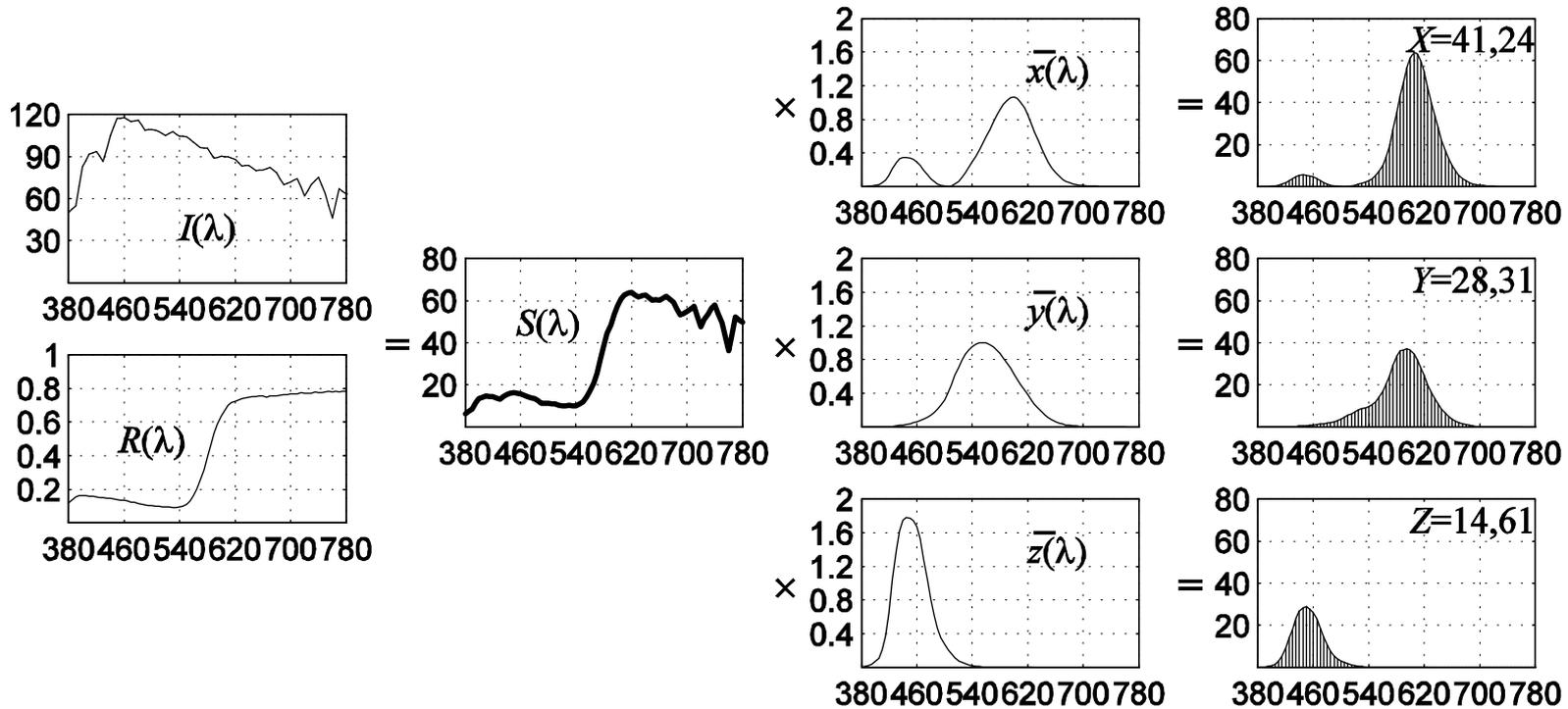
$$\vec{C} = X.\vec{X} + Y.\vec{Y} + Z.\vec{Z}$$

Les coordonnées trichromatiques XYZ s'expriment alors de la façon suivante :

$$X = \sum_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{x}(\lambda)d\lambda, \quad Y = \sum_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{y}(\lambda)d\lambda$$

$$Z = \sum_{\lambda_{min}}^{\lambda_{max}} C(\lambda).\bar{z}(\lambda)d\lambda$$

Les coordonnées trichromatiques XYZ



Exemple : calcul des coordonnées trichromatiques d'une couleur $C(\lambda)$ produit de la réflectance d'une source $I(\lambda)$ et une surface $R(\lambda)$. Dans cet exemple, $k = 0,009463$.

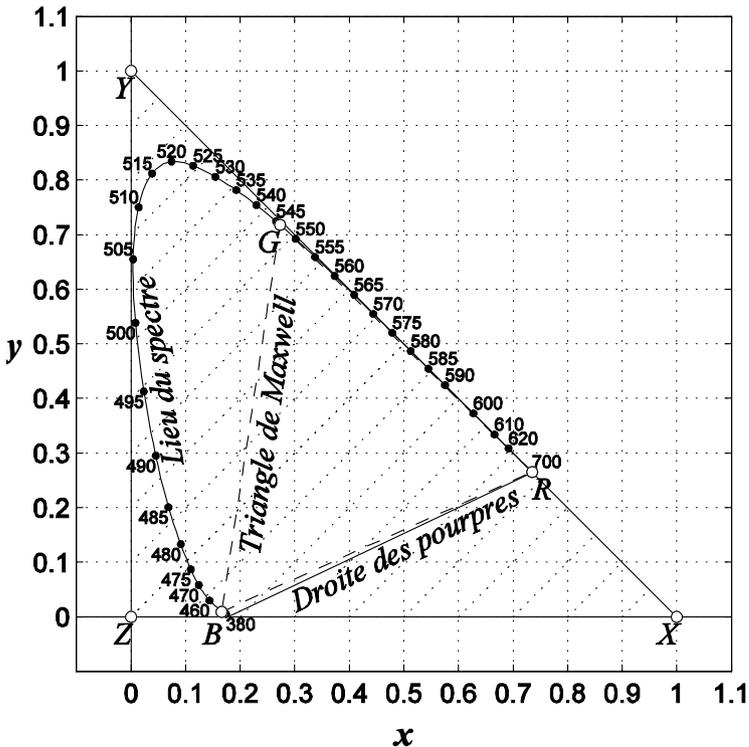
Le diagramme de chromaticité xyY

La normalisation des coordonnées XYZ donne :

$$x = \frac{X}{X+Y+Z} \quad y = \frac{Y}{X+Y+Z} \quad z = \frac{Z}{X+Y+Z}$$

Etant donné que $x + y + z = 1$, la transformation inverse est donnée par :

$$X = x \cdot \frac{Y}{y} \quad Z = (1 - x - y) \cdot \frac{Y}{y}$$



Représentation du diagramme de chromaticité à Y constant.

- La chromaticité correspond à la projection des coordonnées XYZ sur le plan de chromaticité xy.
- L'axe perpendiculaire à ce plan est l'axe des luminances Y.
- Domaine des couleurs reproductibles par synthèse additive des trois primaires XYZ (en pointillé RGB).
- Domaine des couleurs du visible (courbe des couleurs pures et droite des pourpres).

Diagramme de chromaticité *CIE* (1931)

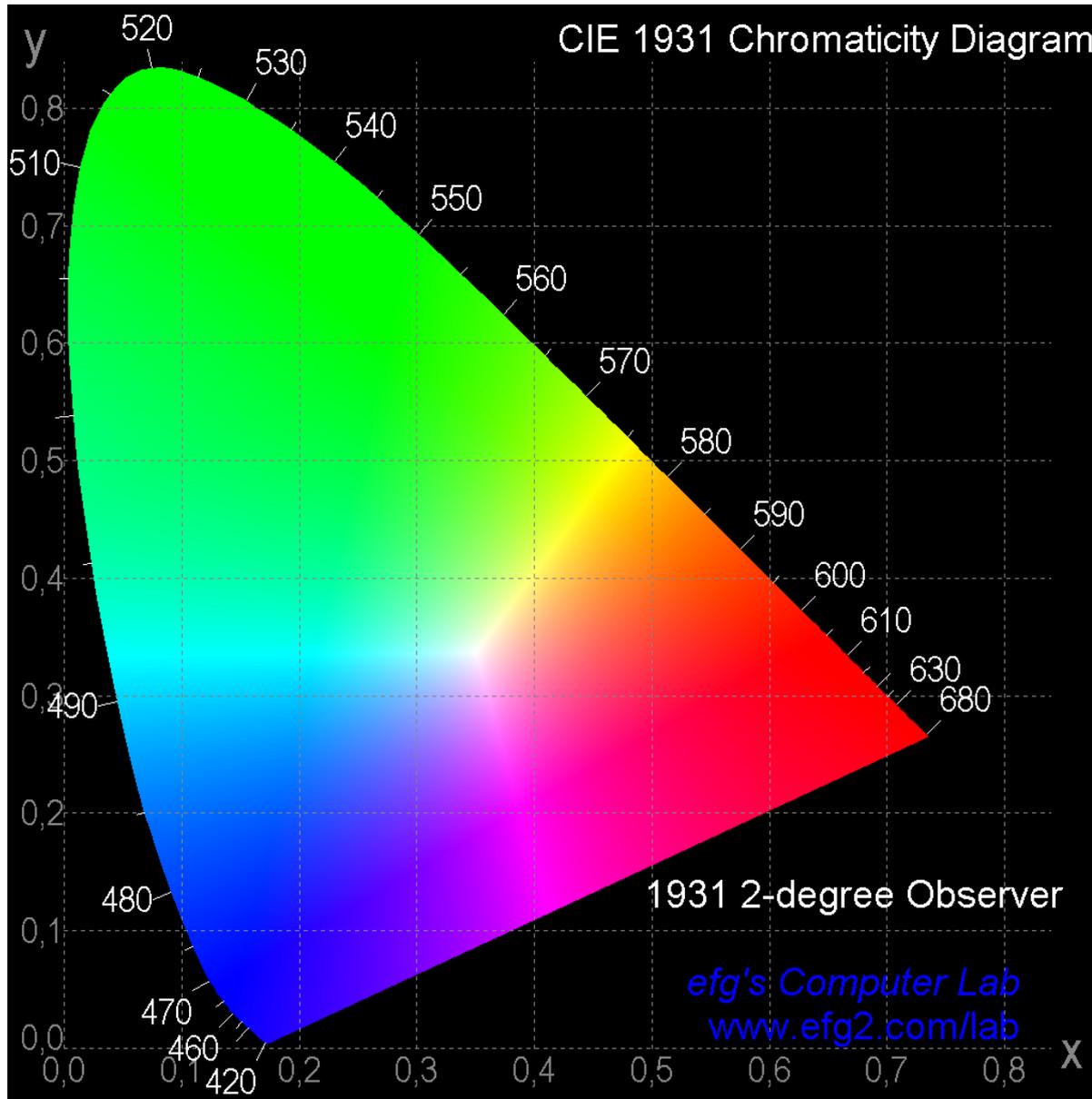
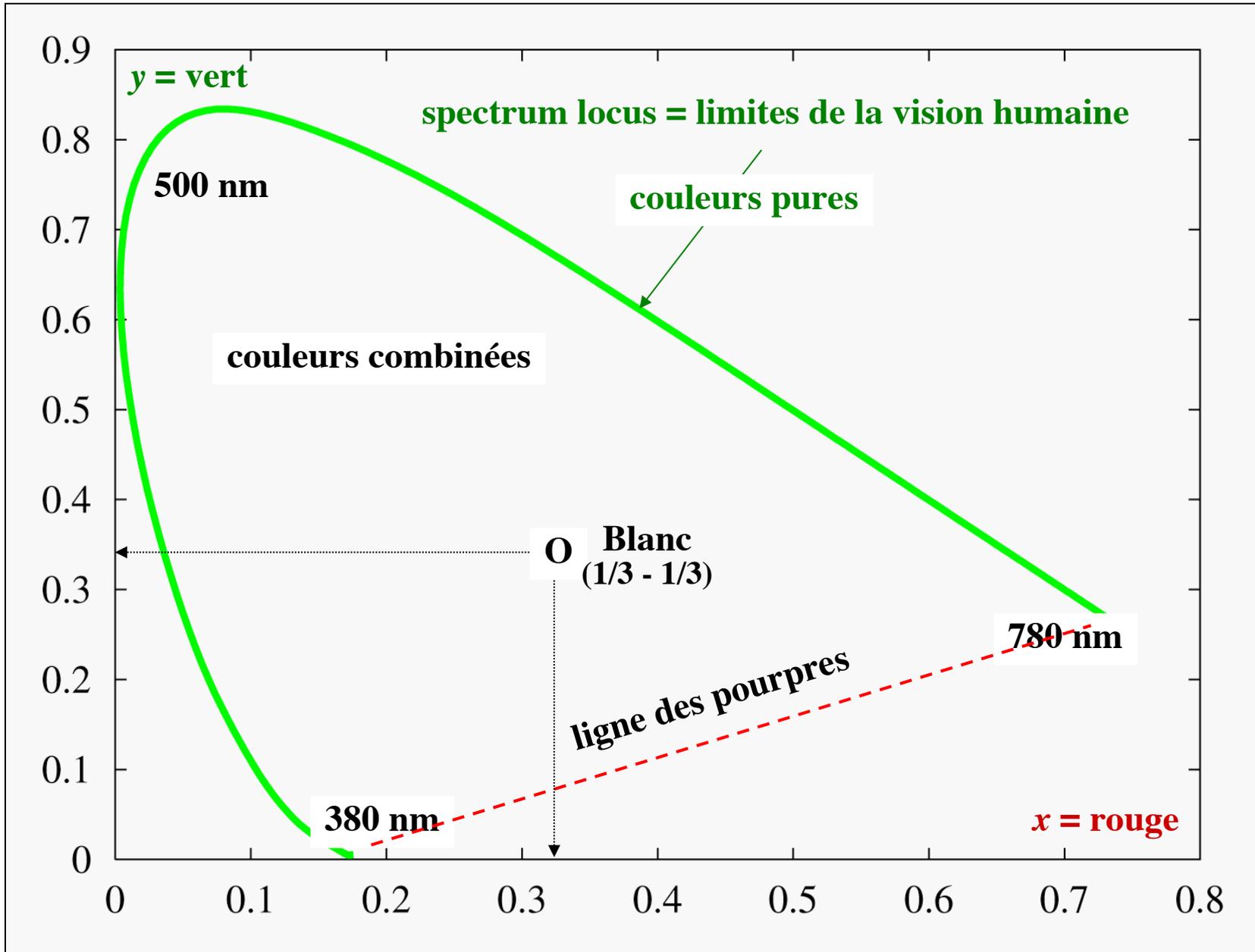
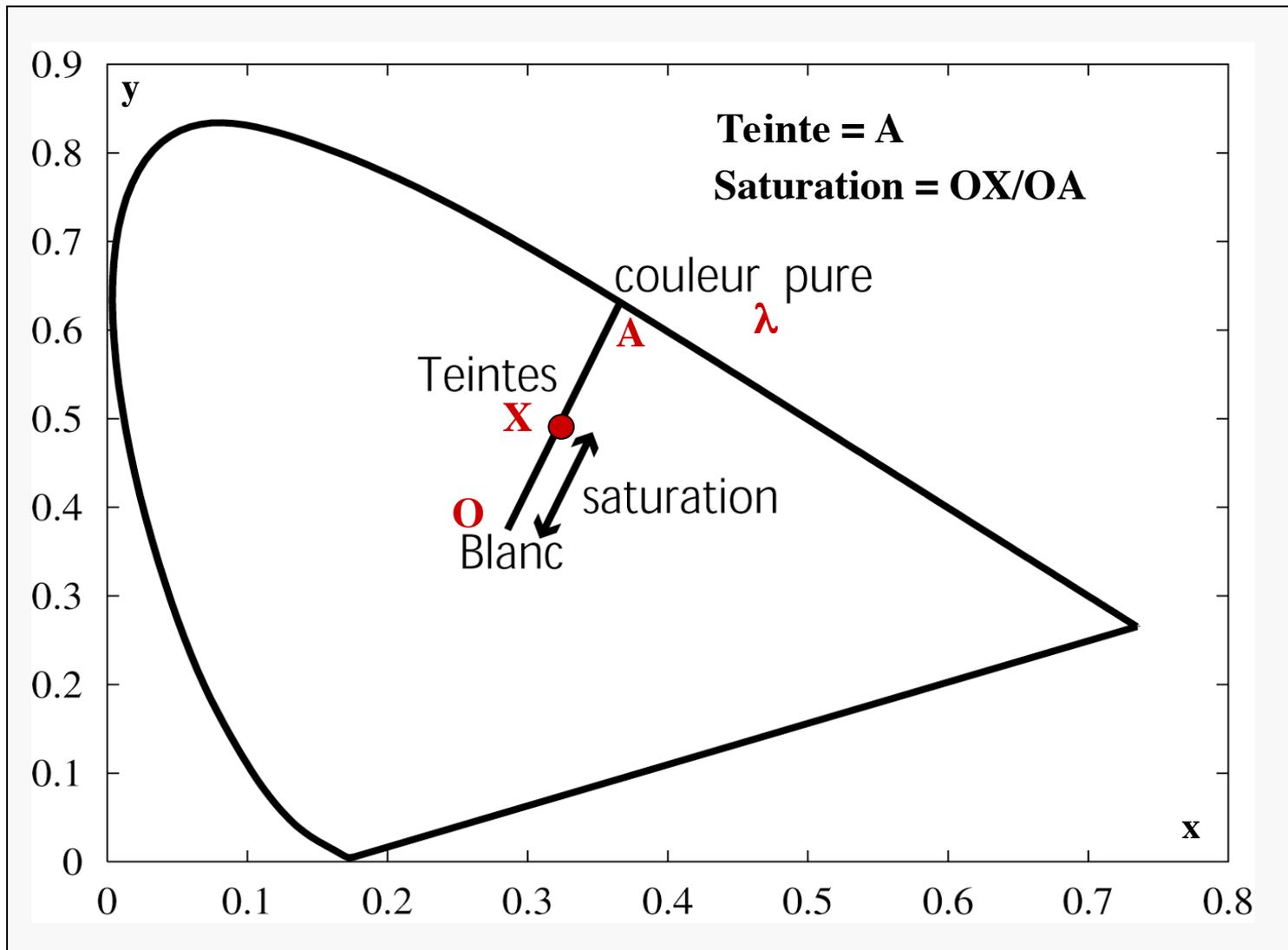


Diagramme de chromaticité



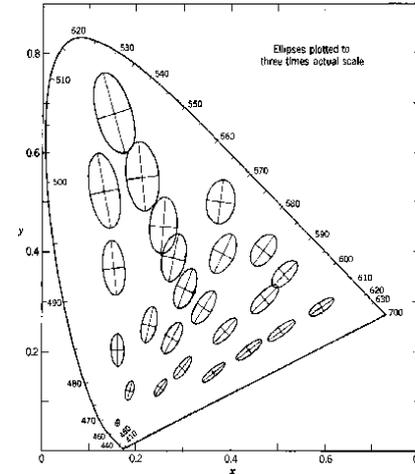


**Une couleur est définie par ses 2 coordonnées chromatiques (x,y)
 et sa luminance Y**

Systeme CIELAB₇₆

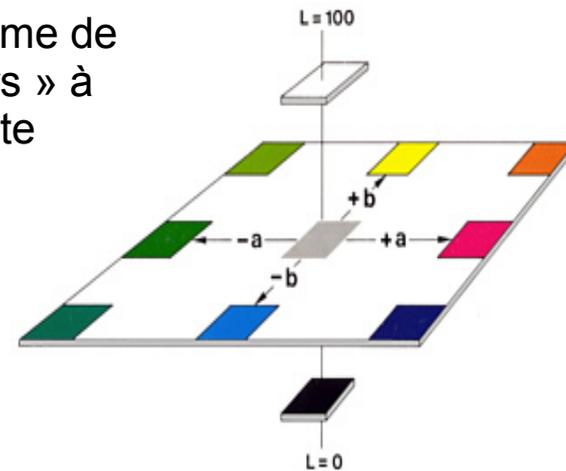
L'espace ainsi obtenu n'est toujours pas pourvu d'une métrique homogène, comme le montre le diagramme des seuils de perception :

Après plusieurs tentatives une transformation non linéaire est couramment adoptée :



$$L^* = 116 \left(\frac{\Psi}{\Psi_I} \right)^{1/3} - 16$$
$$a^* = 500 \left[\left(\frac{X}{X_I} \right)^{1/3} - \left(\frac{Y}{Y_I} \right)^{1/3} \right]$$
$$b^* = 200 \left[\left(\frac{Y}{Y_I} \right)^{1/3} - \left(\frac{Z}{Z_I} \right)^{1/3} \right]$$

Qui redonne une forme de « solide des couleurs » à peu près satisfaisante



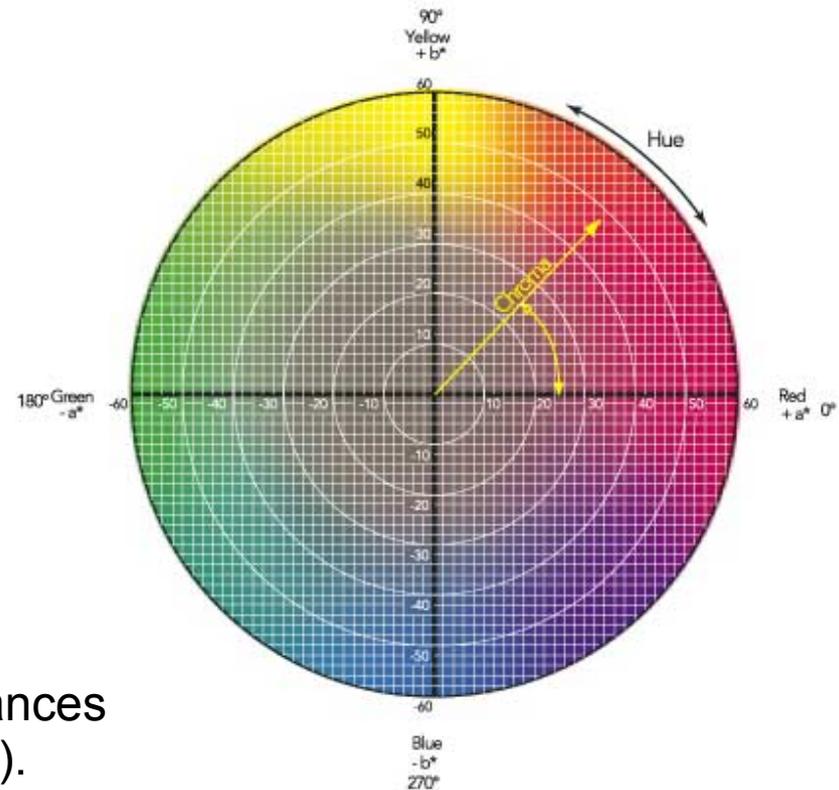
Ce système est conçu pour caractériser la couleur des objets observés en réflexion (mode objet) sous un *illuminant standard*. Par construction on a $L_I^* = 100$ $a_I^* = 0$ $b_I^* = 0$ pour tenir compte des effets d'adaptation.

On définit la chroma :

$$C^* = (a^{*2} + b^{*2})^{1/2}$$

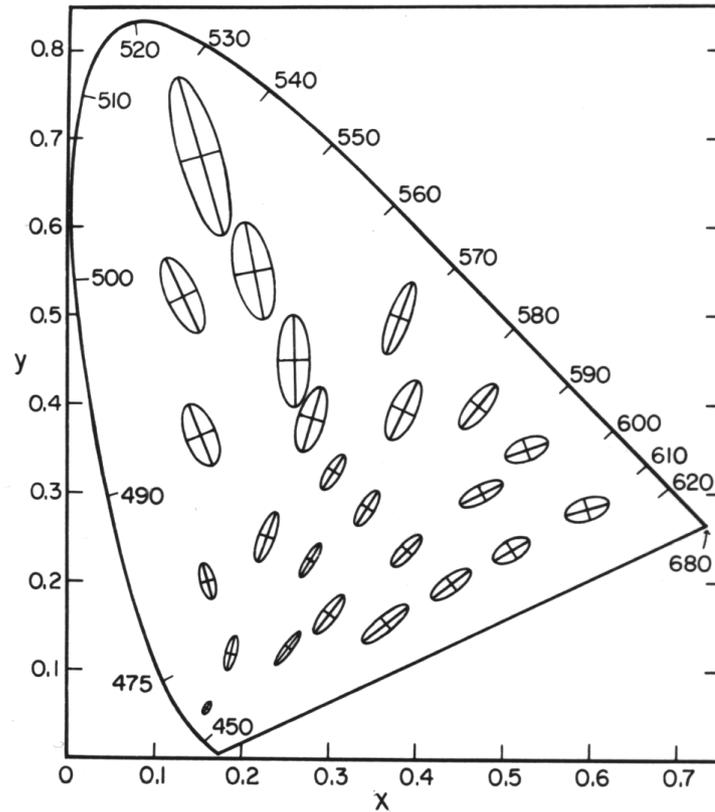
et l'angle de teinte :

$$h = \arctan(b^* / a^*)$$



- la métrique correspond mieux aux distances colorimétriques perçues par l'œil (Munsell).
- est couramment adopté par les professionnels de la couleur.

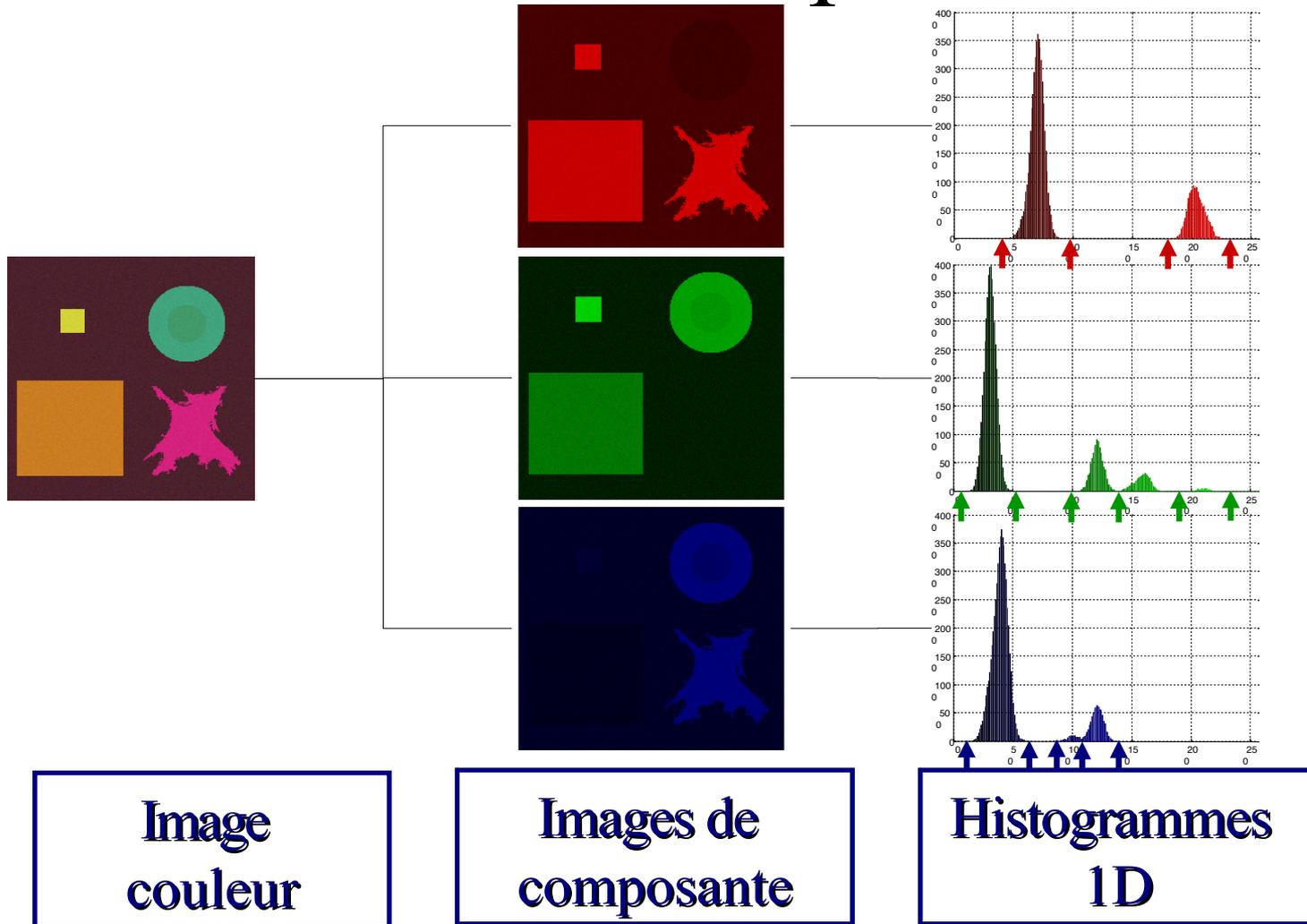
La sensibilité visuelle aux écarts de chromaticité



Seuils différentiels de chromaticité selon Macadam (agrandies 10 fois).

Tandis que le rapport de longueur du grand axe sur le petit axe tend jusqu'à 15:1, ce rapport vaut en moyenne 3,12 avec un écart type de 1,03.

Principe



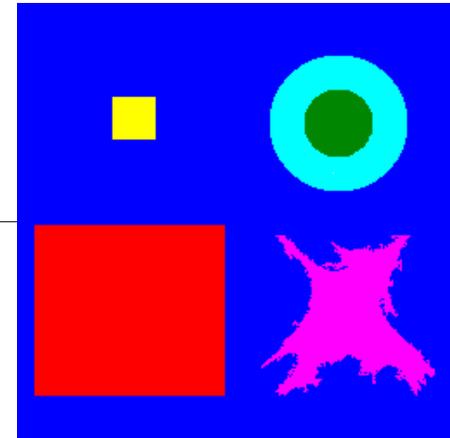
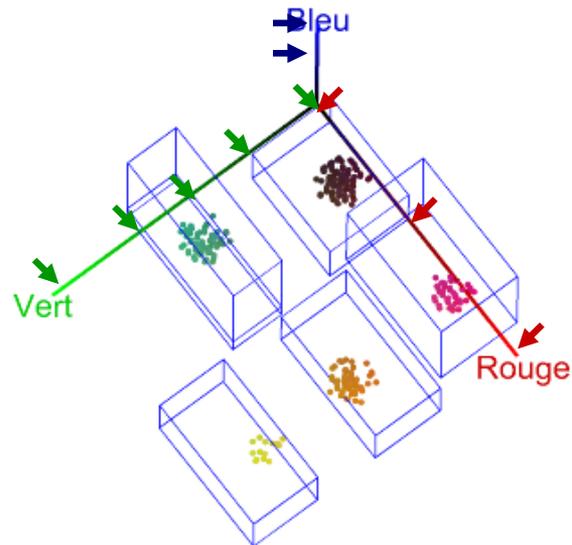
Contexte

Sélection d'espaces

Application

Conclusion

Principe



**Partitionnement de
l'espace**

**Image des classes de
pixels**

Contexte

Sélection d'espaces

Application

Conclusion

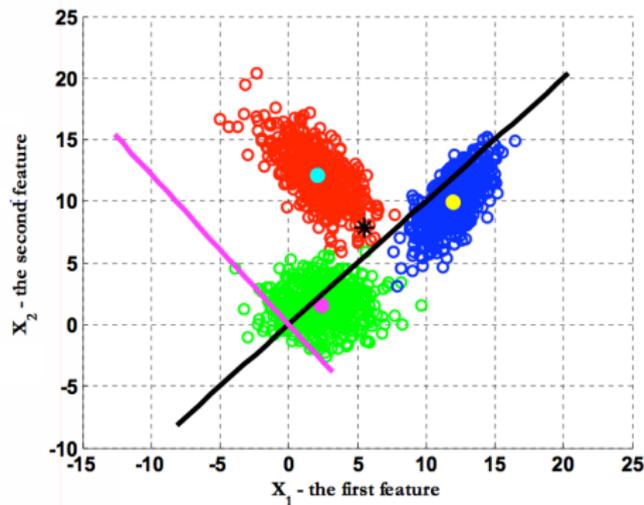
Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

6 avril 2021

Motivations et objectifs du cours

- Réduction de la dimension de l'espace de représentation
- Analyse en composantes principales (non supervisé)
- Mise en oeuvre du critère de Fisher (supervisé)
- Basé sur le cours de Farag, CVIP Lab, Louisville University



Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

1 Analyse en Composantes Principales (non supervisé)

- Matrice de co-variance
- Base de vecteurs orthogonaux
- Vecteurs propres
- Vecteurs propres de la matrice de co-variance

2 Analyse Factorielle Discriminante à 2 classes

- Mesure de séparation entre classes
- Critère de Fisher
- Recherche du max du critère de Fisher
- Exemple

3 Analyse Factorielle Discriminante à 3 classes

Matrice des données discrètes X

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance
Base de vecteurs orthogonaux
Vecteurs propres
Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soient N observations (ou évènements) pour chacun des D attributs $X_i, i = 1, \dots, D$.
- On peut donc représenter les N observations de l'attribut x_i sous la forme d'un vecteur

$$x_i = (x_{i,1}, \dots, x_{i,j}, \dots, x_{i,N})^T$$

- On peut alors rassembler les D vecteurs d'attributs x_i dans une matrice X de dimension $D \times N$.
- $x_{i,j}$ est donc le i ème attribut de la j ème observation.
- Réduction de la dimension consiste à réduire D vers $d < D$ et à trouver la fonction $Y = f(X)$ où Y de dimension $d \times N$

Exemple : Image couleur

- Soit une image de N pixels p caractérisés par un vecteur couleur $(R(p), G(p), B(p))^T$.
- Les composantes RGB sont considérées comme des variables aléatoires ($D = 3$).
- L'image est donc représentée comme une matrice X de dimension $3 \times N$.



Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Co-variance

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit les moyennes des D attributs $M = (\mu_1, \dots, \mu_D)^T$ avec $\mu_i = \frac{1}{N} \sum_{j=1}^N X_{i,j}$
- La co-variance Σ mesure la relation entre les attributs i et l

$$\Sigma_{i,l} = \frac{1}{N} \sum_{j=1}^N (X_{i,j} - \mu_i) \cdot (X_{l,j} - \mu_l)$$

- Une valeur positive signifie que les 2 attributs augmentent ou diminuent en même temps. Par exemple, le nombre d'heures à étudier RDF ($X_{i,j}$) et la note à l'examen $X_{l,j}$ des étudiants j .
- Une valeur négative signifie que quand un attribut augmente ($X_{i,j}$ le nombre d'heures aux zinzins), l'autre diminue (la note à l'examen $X_{l,j}$).
- Une valeur nulle indique l'indépendance ($X_{i,j}$ la taille de l'étudiant et $X_{l,j}$ sa note à l'examen).

Matrice de co-variance Σ

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

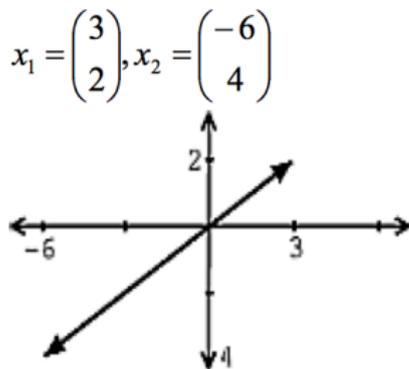
Mesure de séparation entre classes

Critère de Fisher

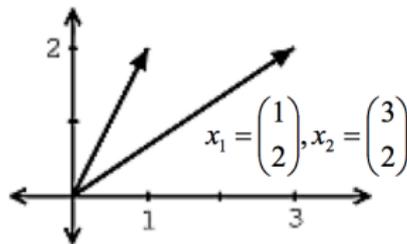
- Des données représentées par D attributs sont représentées par une matrice de co-variance de dimension $(D \times D)$.
- Les termes diagonaux de la matrice $\Sigma_{i,i}$ représentent la variance de l'attribut i .
- La matrice Σ est symétrique à savoir $\Sigma_{i,l} = \Sigma_{l,i}$

Indépendance linéaire

- Un ensemble de k vecteurs x_i de dimension D sont dits linéairement indépendants si aucun d'entre eux ne s'exprime comme une combinaison linéaire des autres :
- $c_1 \cdot x_1 + \dots + c_k \cdot x_k = 0$ ssi $c_1 = \dots = c_k = 0$



Vecteurs non indépendants



Vecteurs indépendants

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Base de vecteurs

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Une base de \mathbb{R}^D est un ensemble de vecteurs x_i de dimension D

- qui permet d'exprimer un vecteur quelconque comme une combinaison linéaire des x_i ,
- qui sont linéairement indépendants entre eux.
- Ainsi tout ensemble de D vecteurs linéairement indépendants entre eux est une base de \mathbb{R}^D .
- Une base est dite normée si $x_i^T \cdot x_i = 1$
- Une base est dite orthogonale si $x_i^T \cdot x_j = 0$

Matrices de transformation

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

- Soit la matrice de transformation (3, 2)
- Prenons un multiple de cette matrice de transformation

$$2 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \times \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 24 \\ 16 \end{bmatrix} = 4 \times \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Valeur propre

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$A \cdot v = \lambda \cdot v$

- (3, 2) est le vecteur propre de la matrice A.
- 4 est la valeur propre de la matrice A.
- Pour une matrice A, comment calculer les vecteurs propres et leurs valeurs propres ?

Valeur propre par recherche de racines

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

$$\Leftrightarrow \mathbf{A} \cdot \mathbf{v} - \lambda \cdot \mathbf{I} \cdot \mathbf{v} = \mathbf{0}$$

$$\Leftrightarrow (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} = \mathbf{0}$$

Valeur propre par recherche de racines

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 0 & 1 \\ -2 & -3 \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ -2 & -3 \end{vmatrix} - \begin{vmatrix} \lambda & 0 \\ 0 & \lambda \end{vmatrix} \\ &= \begin{vmatrix} -\lambda & 1 \\ -2 & -3 - \lambda \end{vmatrix} = (-\lambda \times (-3 - \lambda)) - (-2 \times 1) = \lambda^2 + 3\lambda + 2 \end{aligned}$$

- Les racines sont $\lambda_1 = -1$ et $\lambda_2 = -2$.

Vecteur propre de $\lambda_1 = -1$

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$(A - \lambda_1 \cdot I) \cdot v_1 = 0$$

$$\begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} \cdot \begin{bmatrix} v_{1:1} \\ v_{1:2} \end{bmatrix} = 0$$

$$v_{1:1} + v_{1:2} = 0$$

$$-2v_{1:1} - 2v_{1:2} = 0$$

$$v_{1:1} = -v_{1:2}$$

- Le vecteur propre associé est $(1, -1)$ et pour $\lambda_2 = -2$: $(1, -2)$.

Propriétés des vecteurs propres de Σ

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- On peut extraire d vecteurs propres d'une matrice de dimension $D \times D$.
- Les vecteurs propres extraits d'une matrice symétrique sont orthogonaux.

Réduction de la dimension D de l'espace de représentation en d

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- On veut retenir les attributs les plus caractéristiques des données parmi les D disponibles.
- On peut ignorer les attributs constants, ceux qui ont une faible dispersion, ceux qui sont linéairement dépendants les uns des autres ($x_1 = a.x_2 + b.x_3$)
- On va **conserver ceux avec une variance importante**, et ceux qui ne sont pas corrélés (co-variance faible).

Changement de base

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit W de dimension $(D \times d)$ la matrice de d vecteurs représentant une base.
- Soit le vecteur y qui est la projection de la donnée dans cette base $x : y = W^T \cdot x$
- avec x de dimension $(D \times 1)$
- avec y de dimension $(d \times 1)$
- la matrice Y des projections des données X dans cette base est alors : $Y = W^T \cdot X$
- avec X de dimension $(D \times N)$
- avec Y de dimension $(d \times N)$

Analyse en Composantes Principales

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Calcul de la matrice Σ de co-variance des données X
- Recherche des d vecteurs propres w_i de Σ : $\Sigma \cdot w_i = \lambda_i \cdot w_i$ où λ_i sont les d valeurs propres les plus élevées et $i = 1, \dots, d$.
- La matrice solution représentant la base est $W = (w_1, \dots, w_d)$
- Les données projetées sont alors $Y = W^T \cdot X$

Recherche du vecteurs w_1 avec la plus grande valeur propre

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

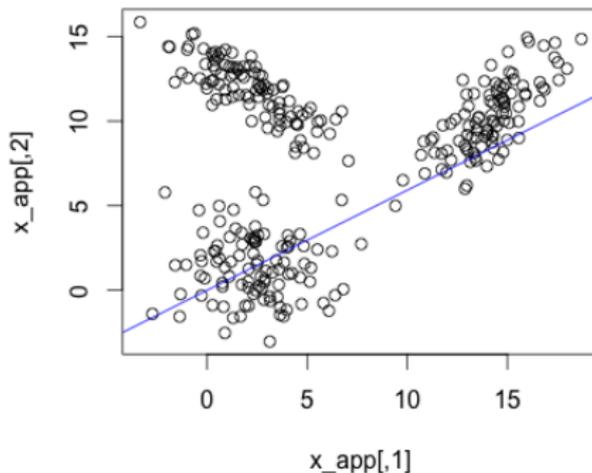
Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher



Apprentissage supervisé avec 2 classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

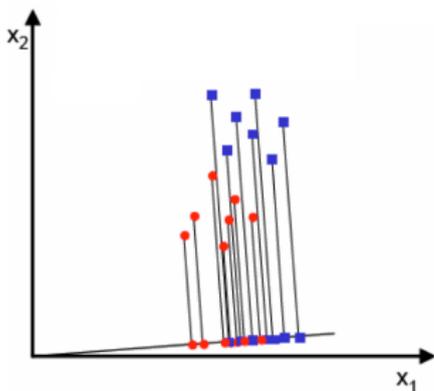
Mesure de séparation entre classes

Critère de Fisher

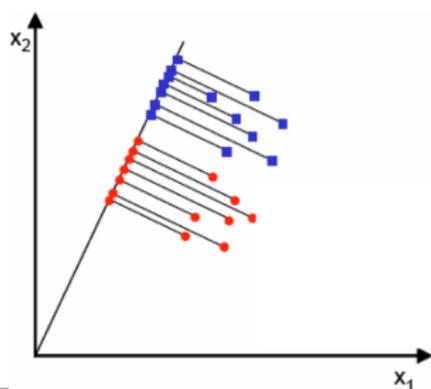
- Nous disposons d'un ensemble d'apprentissage
- Chaque observation est associée à une des $K = 2$ classes ω_k , $k = 1, 2$, à savoir $\omega(x_{*,j})$.
- Chaque classe ω_k est représentée par N_k observations $\{X^1, \dots, X^{N_k}\}$.
- La matrice X se définit comme $\{x^1, \dots, x^N\}$, la succession des données de ω_1 et ω_2 .
- Il s'agit de réduire $d < D$ tel que les 2 classes soient encore identifiées.

AFD pour 2 classes

- Nous recherchons une transformation de X vers Y en projetant les données de X sur un hyper-plan de dimension $D - 1$.
- Nous recherchons le vecteur $w = (w_1, \dots, w_D)$ tel que le scalaire $y = w^T \cdot x_{*,j}$ donne le numéro de classe $\omega(x_{*,j})$.



Données mal séparées



Données bien séparées

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Mesure de séparation entre classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit $\boldsymbol{\mu}_k$ et $\tilde{\mu}_k$ la moyenne de la classe ω_k dans les deux espaces d'attributs : $\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{x \in \omega_k} x$ et $\tilde{\mu}_k = \frac{1}{N_k} \sum_{y \in \omega_k} y = \mathbf{w}^T \cdot \boldsymbol{\mu}_k$
- Soit la distance entre les moyennes projetées comme fonction objective : $J(\mathbf{w}) = |\tilde{\mu}_1 - \tilde{\mu}_2|^2 = |\mathbf{w}^T (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)|^2$

Mesure de séparation entre classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

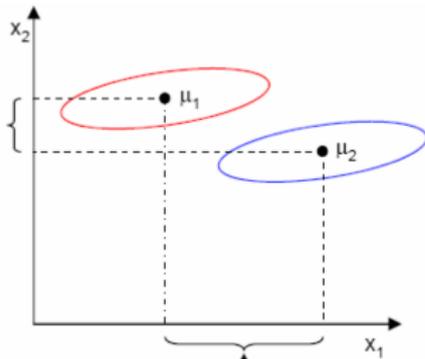
Analyse en Composantes Principales (non supervisé)

Matrice de co-variance
Base de vecteurs orthogonaux
Vecteurs propres
Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

- La distance entre les moyennes ne tient pas compte de la dispersion des données au sein des classes.



- Exemple :
- $|\tilde{\mu}_1 - \tilde{\mu}_2|$ plus élevée selon l'axe X_1 que l'axe X_2 alors que les projections sur l'axe X_1 se mélangent.

Dispersion intra-classe (Fisher)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

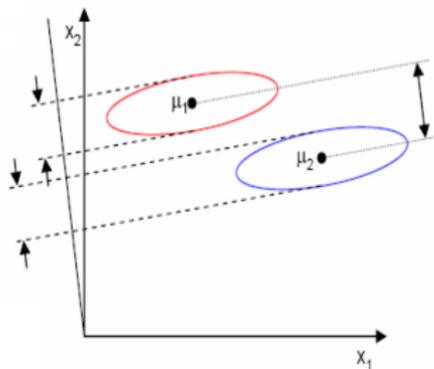
- Soit la dispersion intra-classe des prototypes de la classe ω_k sur le y -espace : $\tilde{s}_k^2 = \frac{1}{N_k} \sum_{y \in \omega_k} (y - \tilde{\mu}_k)^2$
- La dispersion intra-classe est alors $\tilde{s}^2 = \tilde{s}_1^2 + \tilde{s}_2^2$

Critère de Fisher

- Soit la fonction objective de Fisher à maximiser est

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}$$

- Recherche de l'attribut qui **maximise la distance entre moyennes** et **minimise la dispersion intra-classe**.



Critère de Fisher

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Afin de trouver la projection optimale w^* , il faut exprimer J en fonction de w .
- Soit S_k la matrice de dispersion de la classe ω_k dans l'espace d'origine (matrice de co-variance de ω_k)

$$S_k = \frac{1}{N_k} \sum_{x \in \omega_k} (x - \mu_k) \cdot (x - \mu_k)^T$$

- La matrice dispersion intra-classe (within) est alors $S^W = S_1 + S_2$
- La matrice dispersion inter-classe (between) est $S^B = (\mu_1 - \mu_2) \cdot (\mu_1 - \mu_2)^T$

Critère de Fisher (AFD)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

- La dispersion \tilde{s}_k de la projection y peut être exprimée en fonction de S_k :

- $\tilde{s}_k = \frac{1}{N_k} \sum_{y \in \omega_k} (y - \tilde{\mu}_k)^2 = \sum_{x \in \omega_k} (w^T \cdot x - w^T \mu_k)^2$

- $\tilde{s}_k = \sum_{x \in \omega_k} w^T \cdot (x - \mu_k) \cdot (x - \mu_k)^T \cdot w$

- $\tilde{s}_k = w^T \cdot (S_k) \cdot w$

$$\tilde{s} = \tilde{s}_1 + \tilde{s}_2 = w^T \cdot S_1 \cdot w + w^T \cdot S_2 \cdot w$$

- $\tilde{s} = w^T \cdot (S_1 + S_2) \cdot w$

- $\tilde{s} = w^T \cdot (S^W) \cdot w$

Critère de Fisher (AFD)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- de même : $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (\mathbf{w}^T \boldsymbol{\mu}_1 - \mathbf{w}^T \boldsymbol{\mu}_2)^2$
- $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = \mathbf{w}^T \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \cdot \mathbf{w}$
- $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = \mathbf{w}^T \cdot \mathbf{S}^B \cdot \mathbf{w}$
- On peut donc exprimer le critère de Fisher :

$$J(\mathbf{w}) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{\mathbf{w}^T \cdot \mathbf{S}^B \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S}^W \cdot \mathbf{w}}$$

Recherche du max du critère (AFD)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Le maximum de $J(w)$ correspond à $\frac{dJ(w)}{dw} = 0$
- $(S^W)^{-1} \cdot S^B \cdot w - J(w) \cdot w = 0$
- Solution du problème généralisé de valeur propre
- La solution w^* est le vecteur propre.
- $(S^W)^{-1} \cdot S^B \cdot w = \lambda \cdot w$ où $\lambda = J(w)$ est scalaire.

$$w^* = (S^W)^{-1} \cdot (S^B) = (S^W)^{-1} \cdot (\mu_1 - \mu_2)$$

Exemple d'AFD - Echantillons

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

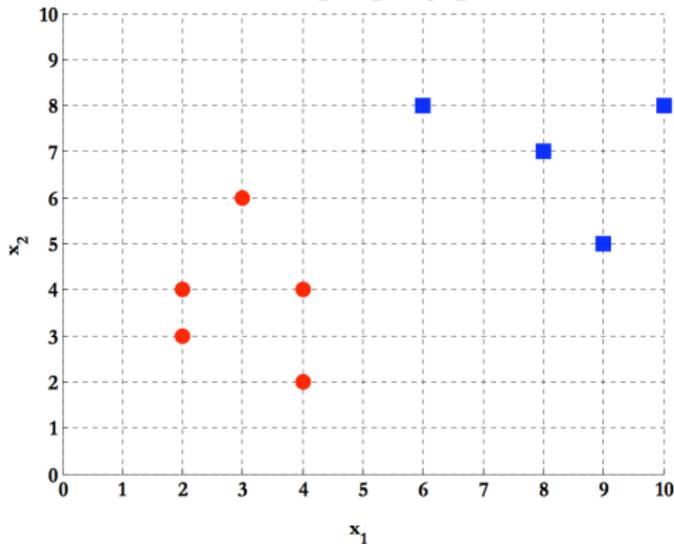
Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Echantillons de la classe ω_1 : $\mathbf{X}_1 = (x_1, x_2) = \{(4,2), (2,4), (2,3), (3,6), (4,4)\}$

Echantillons de la classe ω_2 : $\mathbf{X}_2 = (x_1, x_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$



Exemple d'AFD - Moyenne des classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x = \frac{1}{5} \left[\begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 6 \end{pmatrix} + \begin{pmatrix} 4 \\ 4 \end{pmatrix} \right] = \begin{pmatrix} 3 \\ 3.8 \end{pmatrix}$$
$$\mu_2 = \frac{1}{N_2} \sum_{x \in \omega_2} x = \frac{1}{5} \left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} + \begin{pmatrix} 6 \\ 8 \end{pmatrix} + \begin{pmatrix} 9 \\ 5 \end{pmatrix} + \begin{pmatrix} 8 \\ 7 \end{pmatrix} + \begin{pmatrix} 10 \\ 8 \end{pmatrix} \right] = \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix}$$

Exemple d'AFD - Covariance de la classe ω_1

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{aligned} S_1 &= \sum_{x \in \omega_1} (x - \mu_1)(x - \mu_1)^T = \left[\begin{pmatrix} 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ 3.8 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 3.8 \end{pmatrix} \right]^2 \\ &\quad + \left[\begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 3.8 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 3 \\ 6 \end{pmatrix} - \begin{pmatrix} 3 \\ 3.8 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 4 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 3.8 \end{pmatrix} \right]^2 \\ &= \begin{pmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{pmatrix} \end{aligned}$$

Exemple d'AFD - Covariance de la classe ω_2

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{aligned} S_2 &= \sum_{x \in \omega_2} (x - \mu_2)(x - \mu_2)^T = \left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 6 \\ 8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^2 \\ &+ \left[\begin{pmatrix} 9 \\ 5 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 8 \\ 7 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^2 + \left[\begin{pmatrix} 10 \\ 8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^2 \\ &= \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix} \end{aligned}$$

Exemple d'AFD - Covariance intra-classe

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{aligned} S_w = S_1 + S_2 &= \begin{pmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{pmatrix} + \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix} \\ &= \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix} \end{aligned}$$

Exemple d'AFD - Covariance inter-classe

$$\begin{aligned} S_B &= (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \\ &= \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right] \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]^T \\ &= \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix} \begin{pmatrix} -5.4 & -3.8 \end{pmatrix} \\ &= \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} \end{aligned}$$

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Exemple d'AFD - Résolution équation valeur propre

$$S_W^{-1} S_B w = \lambda w$$

$$\Rightarrow |S_W^{-1} S_B - \lambda I| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 9.2213 - \lambda & 6.489 \\ 4.2339 & 2.9794 - \lambda \end{pmatrix} \right|$$

$$= (9.2213 - \lambda)(2.9794 - \lambda) - 6.489 \times 4.2339 = 0$$

$$\Rightarrow \lambda^2 - 12.2007\lambda = 0 \Rightarrow \lambda(\lambda - 12.2007) = 0$$

$$\Rightarrow \lambda_1 = 0, \lambda_2 = 12.2007$$

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Exemple d'AFD - Résolution équation valeur propre

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{pmatrix} 9.2213 & 6.489 \\ 4.2339 & 2.9794 \end{pmatrix} w_1 = \underbrace{0}_{\lambda_1} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$$\begin{pmatrix} 9.2213 & 6.489 \\ 4.2339 & 2.9794 \end{pmatrix} w_2 = \underbrace{12.2007}_{\lambda_2} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$$w_1 = \begin{pmatrix} -0.5755 \\ 0.8178 \end{pmatrix}$$

$$w_2 = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix} = w^*$$

Exemple d'AFD - Vecteur de projection optimal

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\begin{aligned}w^* &= S_W^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right] \\ &= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix} \\ &= \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}\end{aligned}$$

Exemple d'AFD - Projection sur w_1

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

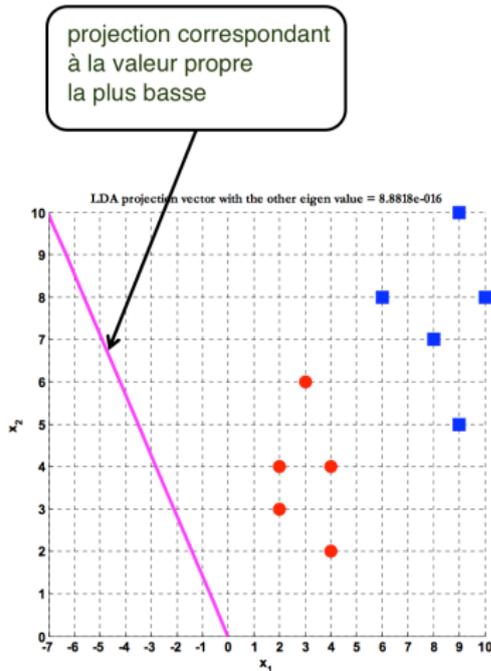
Vecteurs propres

Vecteurs propres de la matrice de co-variance

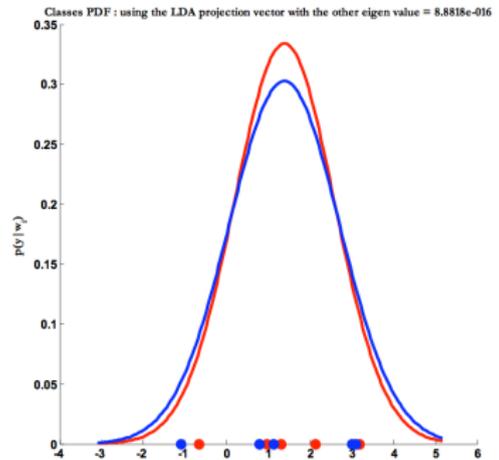
Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher



Données mal séparées



Modes mal séparés

Exemple d'AFD - Projection sur w_2

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

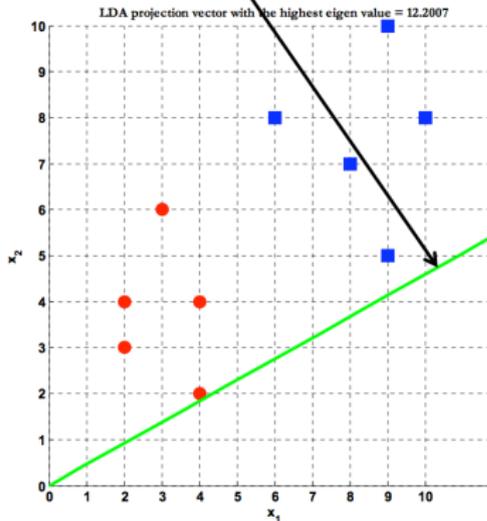
Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

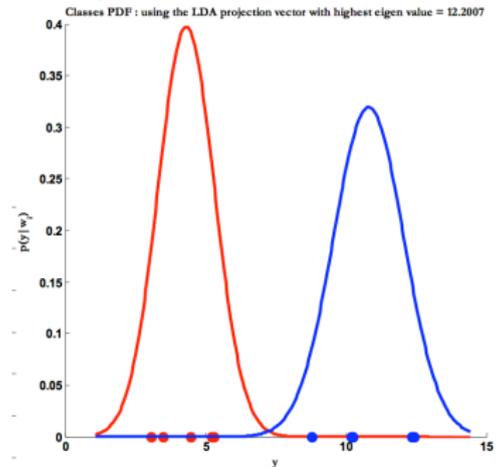
Mesure de séparation entre classes

Critère de Fisher

Vecteur de projection correspondant à la valeur propre la + élevée



Données bien séparées



Bonne séparation des modes des données projetées sur ce vecteur

Modes bien séparés

Problème à K classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- La matrice des données est X composée de N données x de dimension D .
- Chaque classe ω_k est représentée par N_k observations $\{x^1, \dots, x^{N_k}\}$.
- Pour les séparer, il s'agit de rechercher les $(K - 1)$ vecteurs de projection $W = (w_1, \dots, w_{K-1})$.
- Soit la matrice de projection $W = [w_1, \dots, w_{K-1}]$ de dimension $(D \times (K - 1))$
- Soit y_i la valeur de projection de la donnée x sur l'axe w_i :
$$y_i = w_i^T \cdot x$$
- Soit y_i le vecteur des N projections sur l'axe w_i :
$$y_i = w_i^T \cdot X$$

Problème à K classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit le vecteur y des projections de la donnée x : $y = W^T \cdot x$
- avec x de dimension $(D \times 1)$
- avec W de dimension $(D \times (K - 1))$
- avec y de dimension $((K - 1) \times 1)$
- Soit la matrice Y des projections : $Y = W^T \cdot X$
- avec X de dimension $(D \times N)$
- avec W de dimension $(D \times (K - 1))$
- avec Y de dimension $((K - 1) \times N)$

Co-variance intra-classe (K classes)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit S_k la matrice de dispersion de la classe ω_k dans l'espace d'origine (matrice de co-variance de ω_k)

$$S_k = \frac{1}{N_k} \sum_{x \in \omega_k} (x - \mu_k) \cdot (x - \mu_k)^T$$

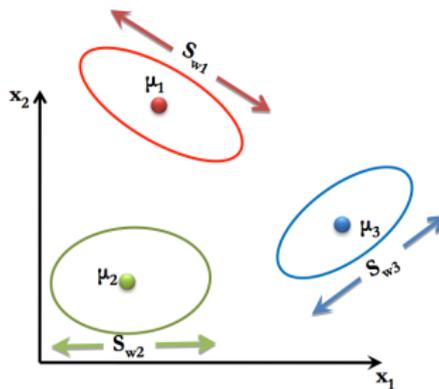
- La matrice dispersion intra-classe (within) est alors

$$S^W = \sum_{k=1}^K S_k$$

Co-variance intra-classe (K classes)

- La matrice dispersion intra-classe (within) est alors

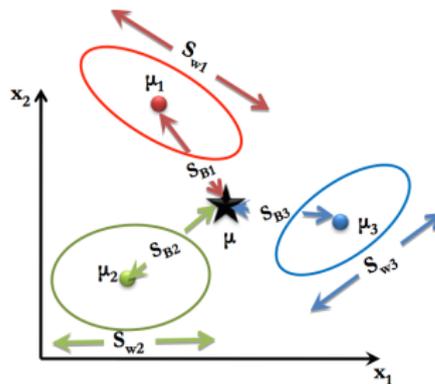
$$S^W = \sum_{k=1}^K S_k$$



Co-variance inter-classe S^B (K classes)

$$S_B = \frac{1}{N} \sum_{k=1}^K N_k (\boldsymbol{\mu}_k - \boldsymbol{\mu}) \cdot (\boldsymbol{\mu}_k - \boldsymbol{\mu})^T$$

- avec $\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{\mathbf{x} \in \omega_k} \mathbf{x}$ et $\boldsymbol{\mu} = \frac{1}{N} \sum_{\mathbf{x}} \mathbf{x}$



Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

Co-variance intra-classe des projections (K classes)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Soit les moyennes $\tilde{\boldsymbol{\mu}}_k = \frac{1}{N_k} \sum_{\mathbf{y} \in \omega_k} \mathbf{y}$ et $\tilde{\boldsymbol{\mu}} = \frac{1}{N} \sum \mathbf{y}$
- Rappel $\tilde{\boldsymbol{\mu}}_k$ de dimension $((K - 1) \times 1)$ Soit S_k la matrice de dispersion de la classe ω_k dans l'espace d'origine (matrice de co-variance de ω_k)

$$\tilde{S}_k = \frac{1}{N_k} \sum_{\mathbf{y} \in \omega_k} (\mathbf{y} - \tilde{\boldsymbol{\mu}}_k) \cdot (\mathbf{y} - \tilde{\boldsymbol{\mu}}_k)^T$$

- La matrice dispersion intra-classe (within) est alors

$$S^{\tilde{W}} = \sum_{k=1}^K \tilde{S}_k$$

Co-variance inter-classe des projections (K classes)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\tilde{S}_B = \frac{1}{N} \sum_{k=1}^K N_k (\tilde{\boldsymbol{\mu}}_k - \tilde{\boldsymbol{\mu}}) \cdot (\tilde{\boldsymbol{\mu}}_k - \tilde{\boldsymbol{\mu}})^T$$

- avec $\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{\mathbf{x} \in \omega_k} \mathbf{x}$ et $\boldsymbol{\mu} = \frac{1}{N} \sum_{\mathbf{x}} \mathbf{x}$

Discrimination de Fisher (AFD) (K classes)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- A partir du problème à 2 classes, on peut généraliser les matrices de co-variance des données projetées :
- $\tilde{S}^W = W^T.(S^W).W$
- $\tilde{S}^B = W^T.(S^B).W$
- On recherche la projection W^* qui maximise le ratio entre la dispersion inter-classe et intra-classe.
- Comme la projection est de dimension $K - 1$, on utilise les déterminants des matrices :

$$J(W) = \frac{|\tilde{S}^B|}{|\tilde{S}^W|} = \frac{|W^T.(S^B).W|}{|W^T.(S^W).W|}$$

Recherche du max du critère de Fisher (K classes)

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- On recherche W^* qui maximise $J(W)$
- Rappel pour 2 classes : la solution w^* est le vecteur propre.
- $(S^W)^{-1} \cdot S^B \cdot w = \lambda \cdot w$ où $\lambda = J(w)$ est scalaire.
- Pour K classes, et la recherche de $K - 1$ vecteurs propres, le problème peut être généralisé à :

$$(S^W)^{-1} \cdot S^B \cdot w_k = \lambda_k \cdot w_k$$

- où $\lambda_k = J(w_k)$ est scalaire. et $k = 1, \dots, K - 1$.
- la solution est donc le vecteur propre w_k^* associé à la valeur propre la plus élevée.
- La matrice solution est $W^* = (w_1^*, \dots, w_{K-1}^*)$

Caractéristiques des 3 classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

$$\mu = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$\mu_1 = \mu + \begin{bmatrix} -3 \\ 7 \end{bmatrix}, \quad \mu_2 = \mu + \begin{bmatrix} -2.5 \\ -3.5 \end{bmatrix}, \quad \mu_3 = \mu + \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

$$S_1 = \begin{pmatrix} 5 & -1 \\ -3 & 3 \end{pmatrix}$$

$$S_2 = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$$

$$S_3 = \begin{pmatrix} 3.5 & 1 \\ 3 & 2.5 \end{pmatrix}$$

Points des 3 classes

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

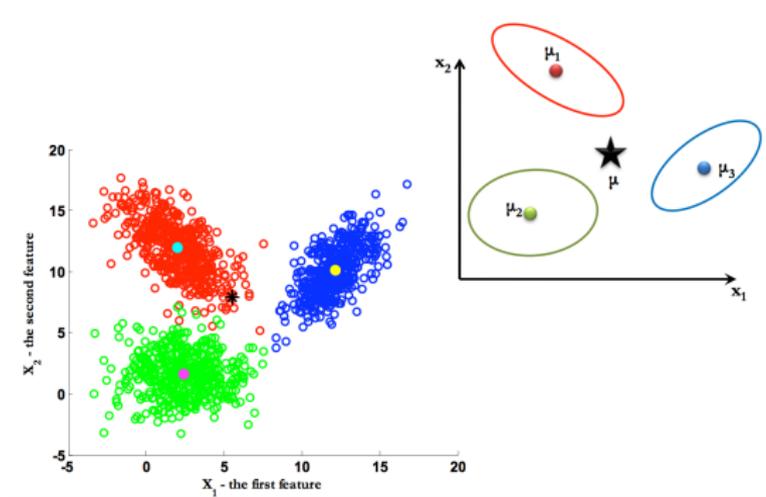
Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher



Vecteurs propres W^*

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

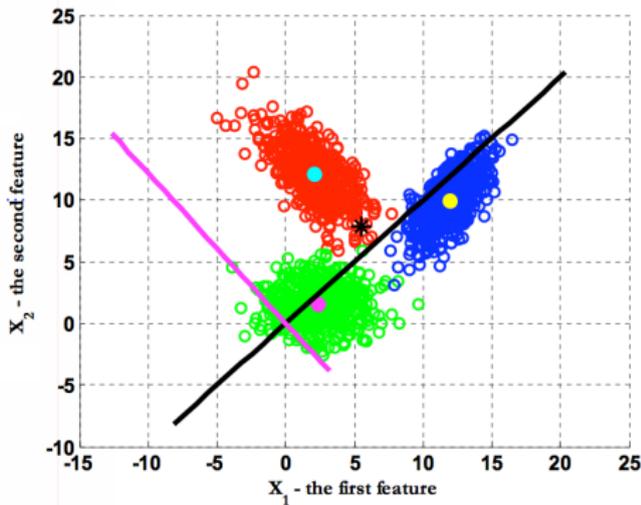
Vecteurs propres

Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher



pdf des projections $y = W^T \cdot x$

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

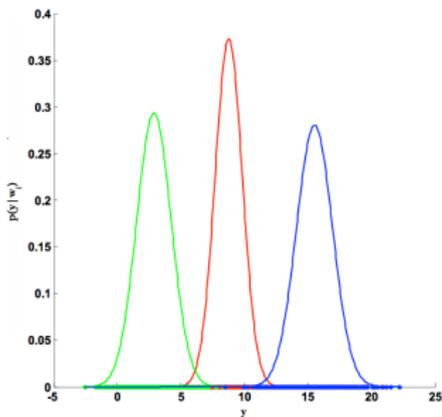
Vecteurs propres

Vecteurs propres de la matrice de co-variance

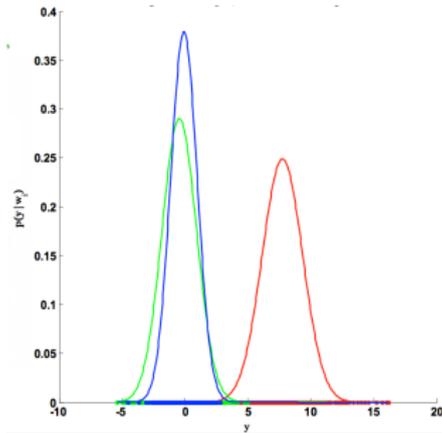
Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher



val propre : 4508



val propre : 1878

Limites de l'Analyse Factorielle Discriminante

Réduction de la dimension de l'espace de représentation

M1 AIV1 - C3

Analyse en Composantes Principales (non supervisé)

Matrice de co-variance

Base de vecteurs orthogonaux

Vecteurs propres

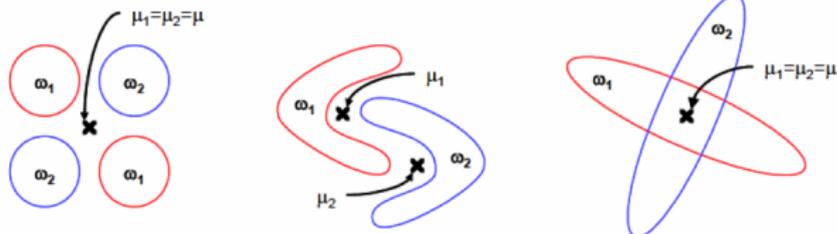
Vecteurs propres de la matrice de co-variance

Analyse Factorielle Discriminante à 2 classes

Mesure de séparation entre classes

Critère de Fisher

- Hypothèse de distribution Gaussienne
- Sélection de $C - 1$ attributs pas toujours suffisants



Segmentation par kmeans

L. Macaire - M1 IVI - RDF - Cours 8

13 avril 2021

Une application industrielle

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Contrôle qualité par vision
- Apprentissage supervisé
- Segmentation d'image couleur
- Acquisition d'images ligne couleur

Défauts à détecter



Image d'un verre
sans défaut



Image d'un verre avec un
défaut de tâché

Contexte

Sélection d'espaces

Application

Conclusion

Défauts à détecter



Image d'un verre
sans défaut



Image d'un verre avec un
défaut de manque

Contexte

Sélection d'espaces

Application

Conclusion

Défauts à détecter



Image d'un verre
sans défaut



Image d'un verre avec un
défaut de décalage

Contexte

Sélection d'espaces

Application

Conclusion

Défauts à détecter



Image d'un verre
sans défaut



Image d'un verre avec un
défaut de nuance

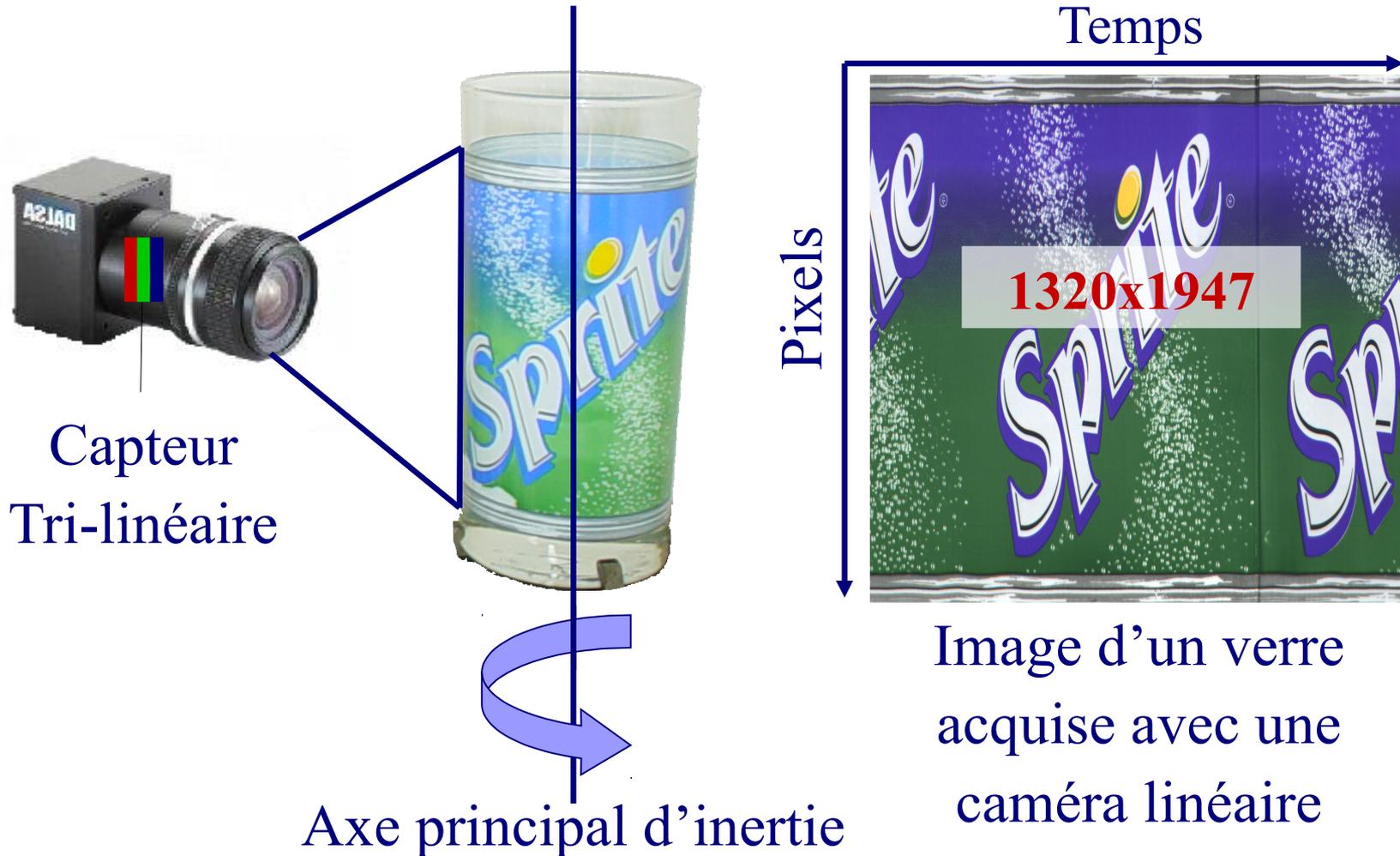
Contexte

Sélection d'espaces

Application

Conclusion

Poste d'acquisition



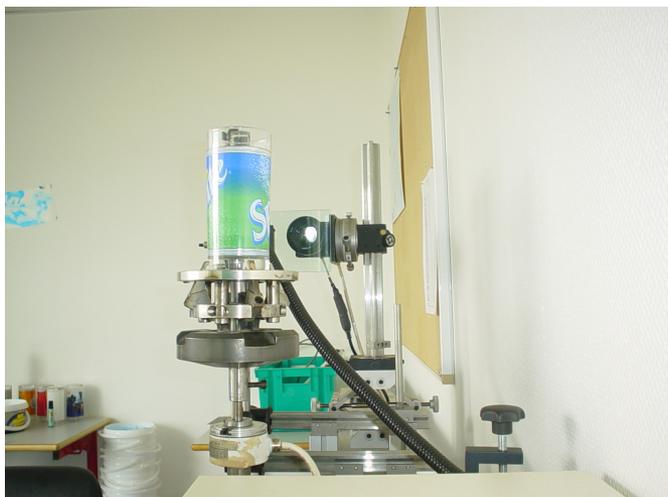
Contexte

Sélection d'espaces

Application

Conclusion

Poste d'acquisition



- Verre à inspecter
- Système de préhension
- Caméra linéaire
- Système d'éclairage

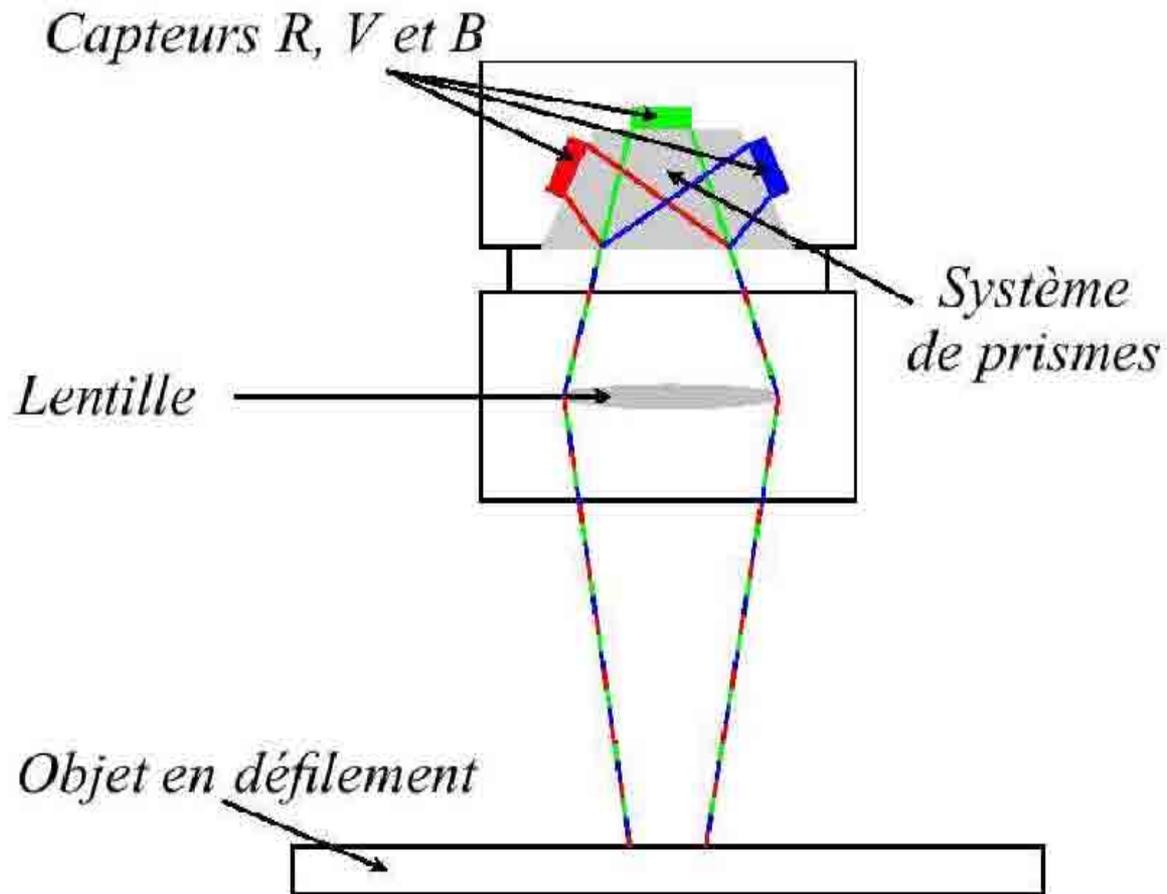
Contexte

Sélection d'espaces

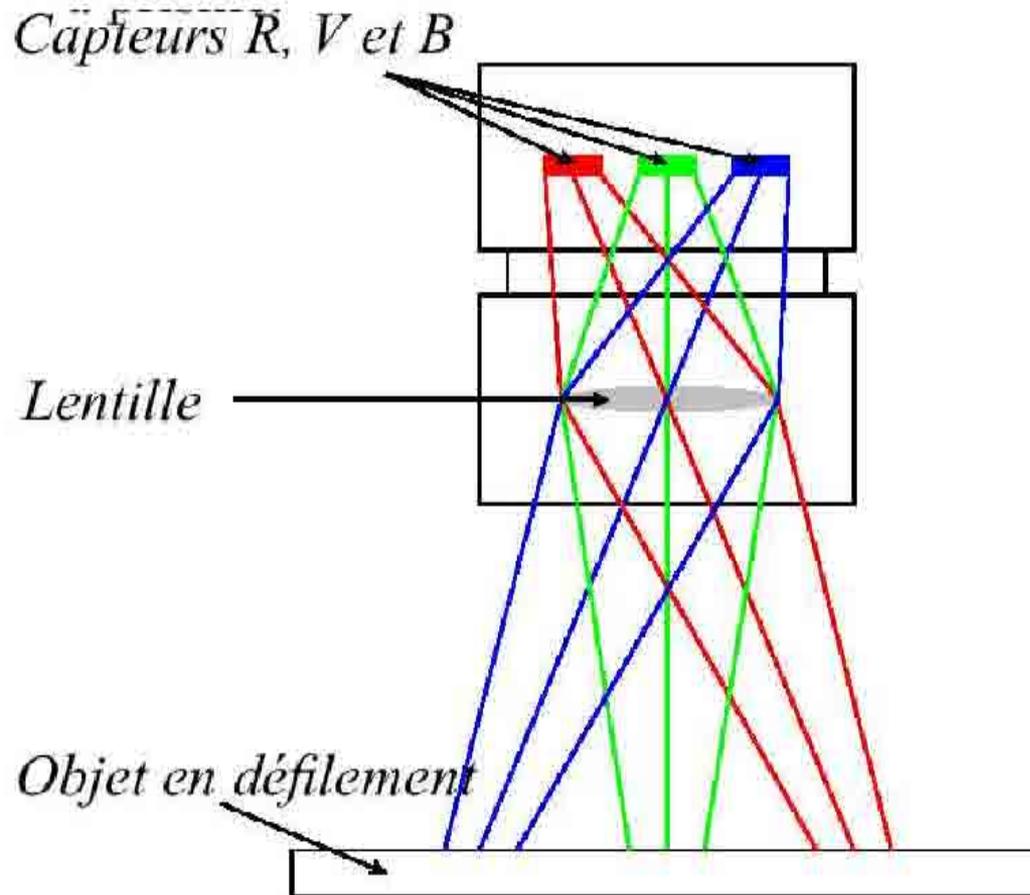
Application

Conclusion

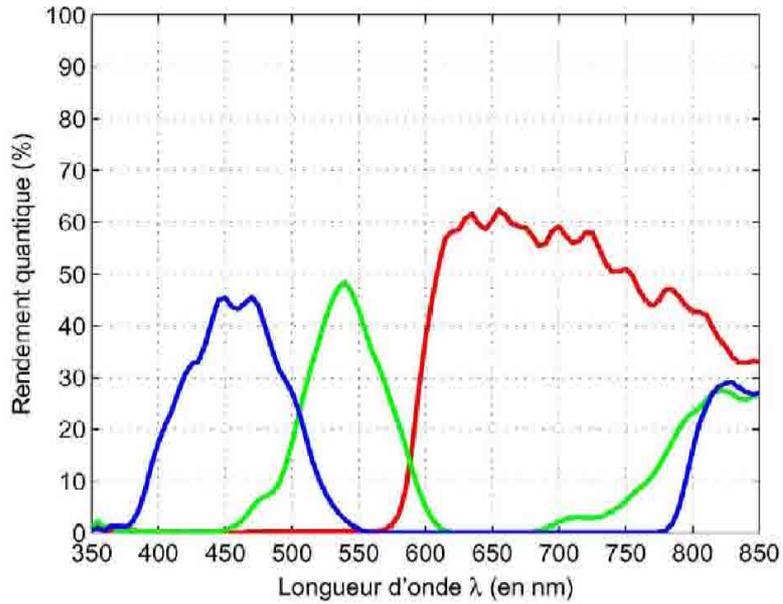
Caméra linéaire couleur 3CCD



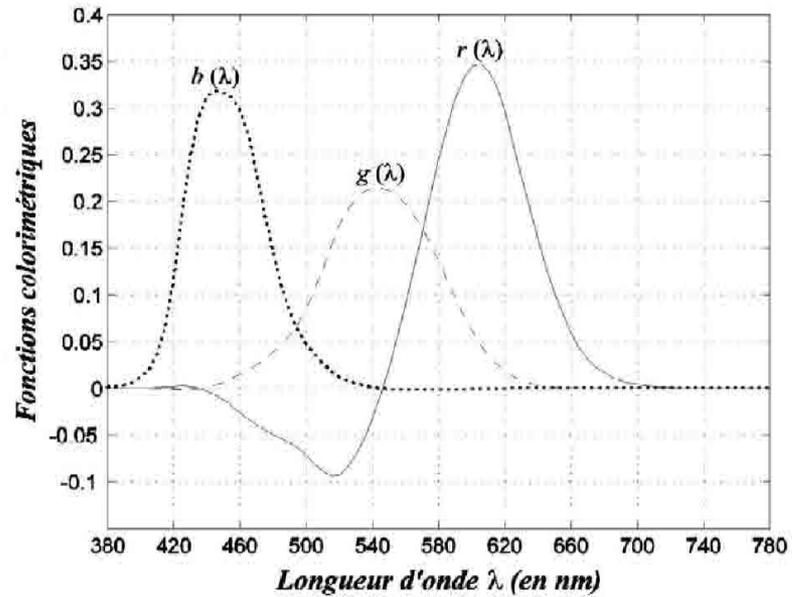
Caméra linéaire couleur tri-linéaire



Sensibilité des capteurs

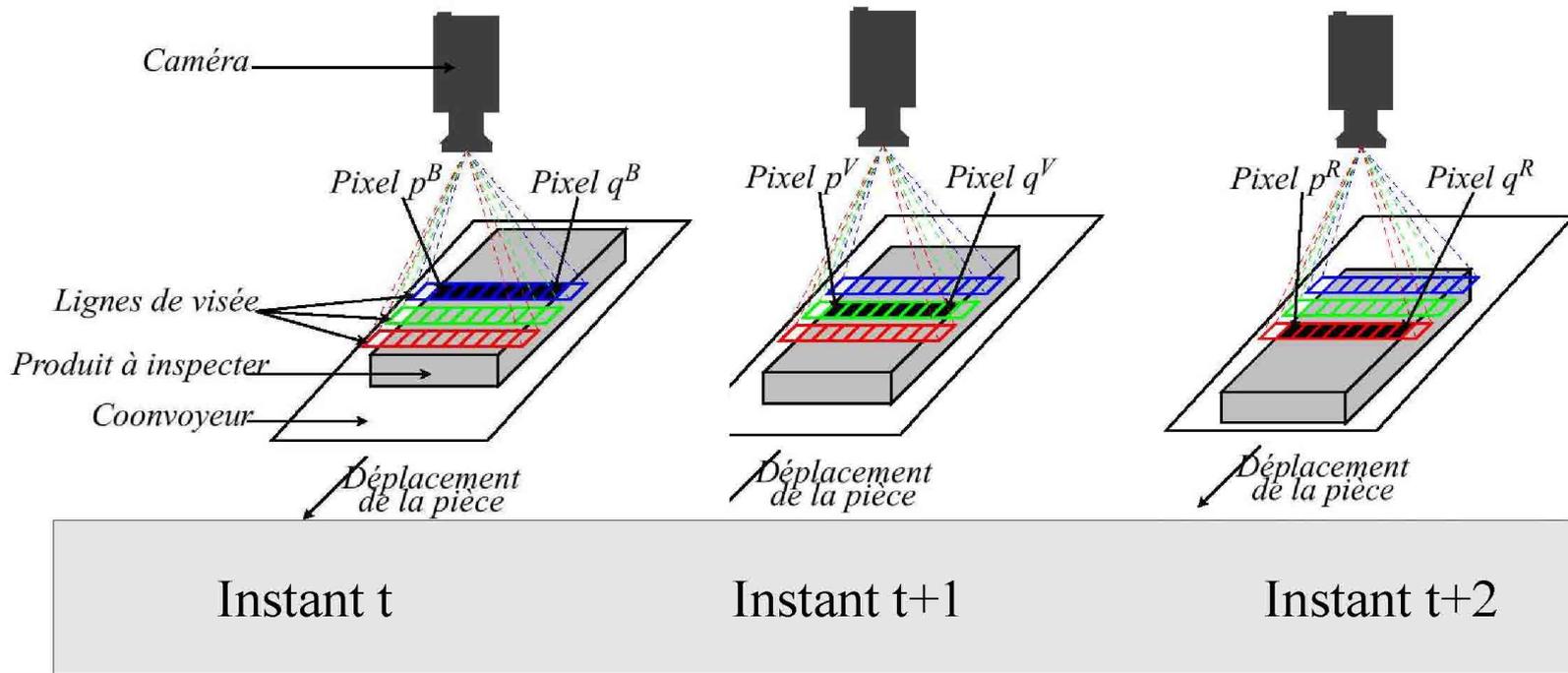


Caméra RGB



CIE 1931

Acquisition images-lignes couleur



Décalage Spatial



DECALAGE SPATIAL

The image shows the text "DECALAGE SPATIAL" in a bold, sans-serif font. Each letter is rendered with a multi-colored, rainbow-like border, creating a significant visual distortion and misalignment of the text.

(a) Image initiale.



DECALAGE SPATIAL

The image shows the text "DECALAGE SPATIAL" in a bold, black, sans-serif font. The text is sharp and clear, with no visible misalignment or distortion, representing the corrected version of the image.

(b) Image corrigée.

Simple comparaison



Image avec défaut

-



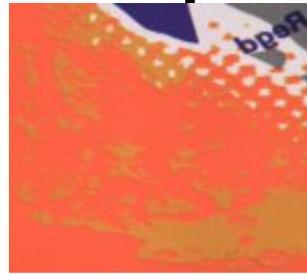
Image sans défaut

=



Image résultat

Apprentissage supervisé hors-ligne



ROI originale

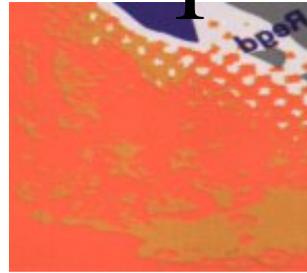
Contexte

Sélection d'espaces

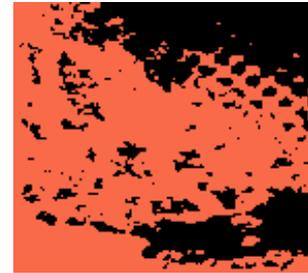
Application

Conclusion

Apprentissage supervisé hors-ligne



ROI originale



Classe 1

Contexte

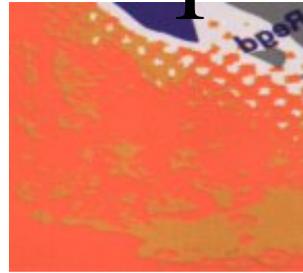
Sélection d'espaces

Application

Conclusion

Apprentissage supervisé hors-ligne

(A,C1,C2)



ROI originale



Classe 1

(A,C1,C2)

Contexte

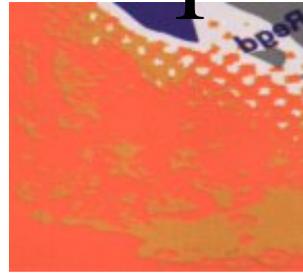
Sélection d'espaces

Application

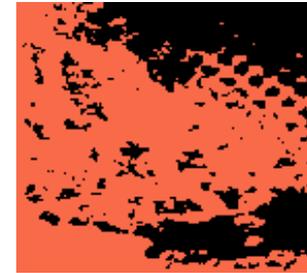
Conclusion

Apprentissage supervisé hors-ligne

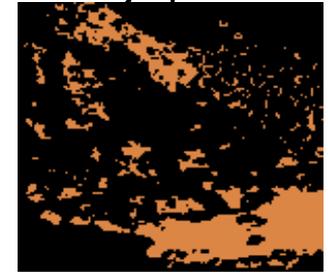
(X, Y, Z)



ROI originale



Classe 1



Classe 2

$(A, C1, C2)$

(X, Y, Z)

Contexte

Sélection d'espaces

Application

Conclusion

Apprentissage supervisé hors-ligne

(R,G,B)



ROI originale



Classe 1



Classe 2



Classe 3

(A,C1,C2)

(X,Y,Z)

(R,G,B)

Contexte

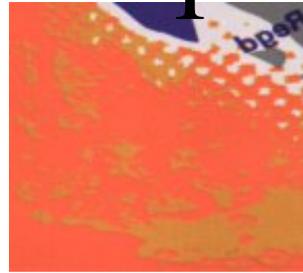
Sélection d'espaces

Application

Conclusion

Apprentissage supervisé hors-ligne

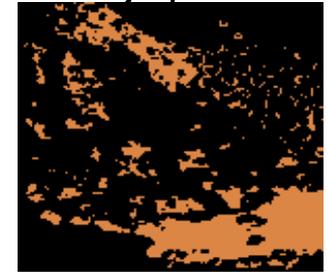
(I1,I2,I3)



ROI originale



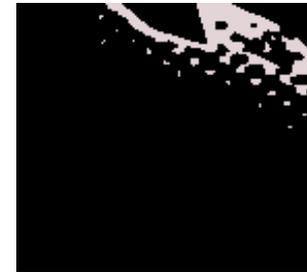
Classe 1



Classe 2



Classe 3



Classe 4

(A,C1,C2)

(X,Y,Z)

(R,G,B)

(I1,I2,I3)

Contexte

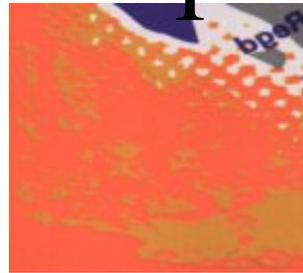
Sélection d'espaces

Application

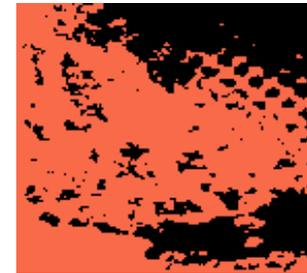
Conclusion

Apprentissage supervisé hors-ligne

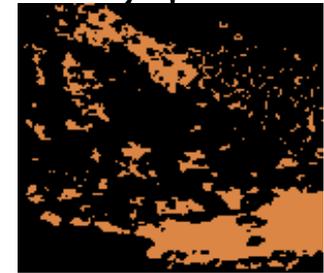
(X, Y, Z)



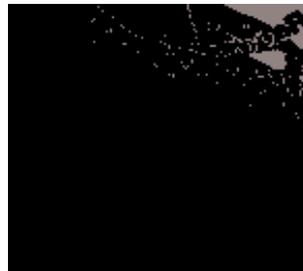
ROI originale



Classe 1



Classe 2



Classe 3



Classe 4



Classe 5

$(A, C1, C2)$

(X, Y, Z)

(R, G, B)

$(I1, I2, I3)$

(X, Y, Z)

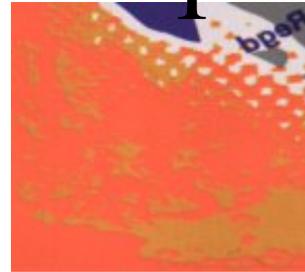
Contexte

Sélection d'espaces

Application

Conclusion

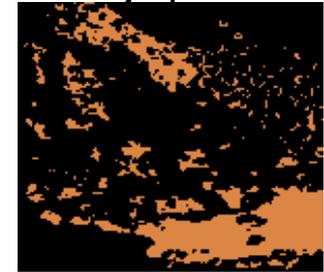
Apprentissage supervisé hors-ligne



ROI originale



Classe 1



Classe 2



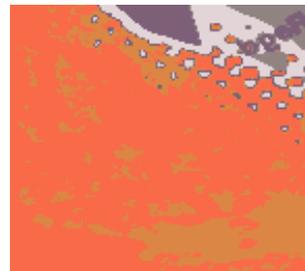
Classe 3



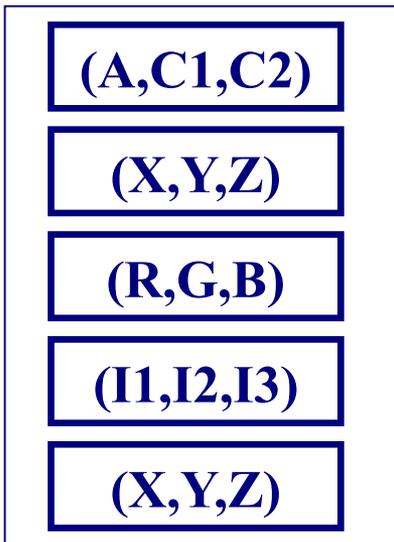
Classe 4



Classe 5



ROI
segmentée



Contexte

Sélection d'espaces

Application

Conclusion

Détection des défauts en-ligne



ROI à inspecter

Contexte

Sélection d'espaces

Application

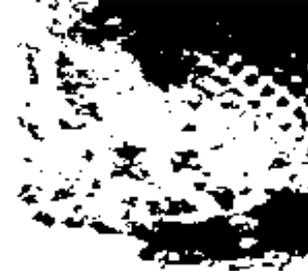
Conclusion

Détection des défauts en-ligne

(A,C1,C2)



ROI à inspecter



Classe 1

(A,C1,C2)

(X,Y,Z)

(R,G,B)

(I1,I2,I3)

(X,Y,Z)

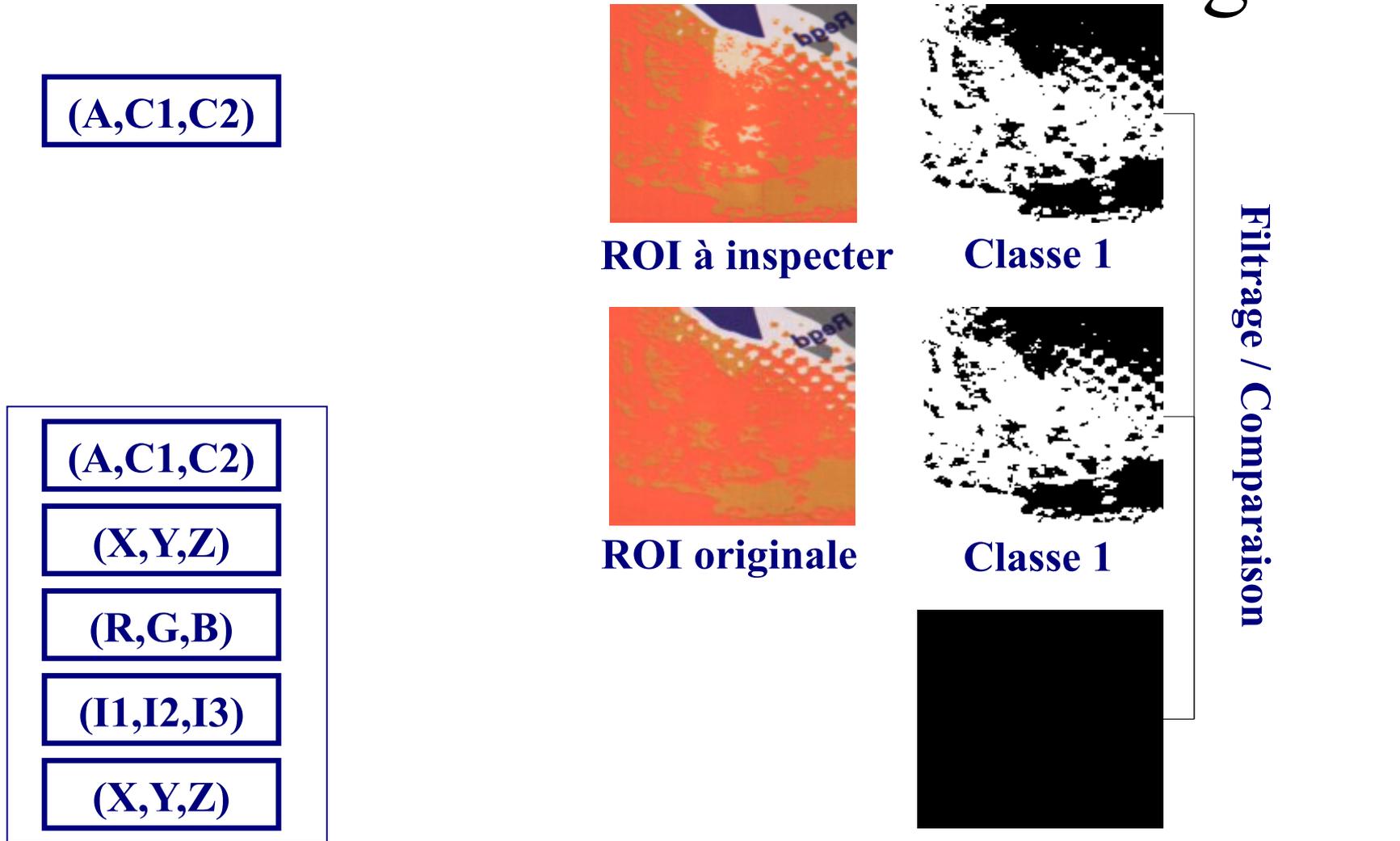
Contexte

Sélection d'espaces

Application

Conclusion

Détection des défauts en-ligne



Contexte

Sélection d'espaces

Application

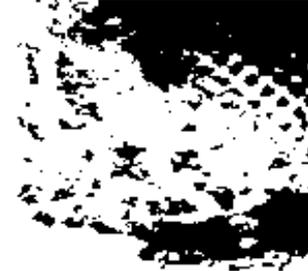
Conclusion

Détection des défauts en-ligne

(X,Y,Z)



ROI à inspecter



Classe 1



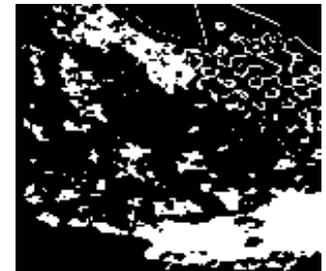
Classe 2



ROI originale

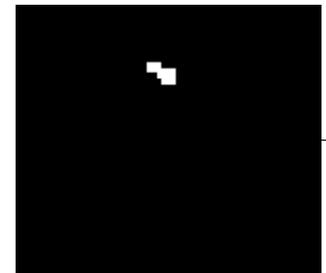


Classe 1



Classe 2

REBUS



(X,Y,Z)

(R,G,B)

(I1,I2,I3)

(X,Y,Z)

Contexte

Sélection d'espaces

Application

Conclusion

Apprentissage non supervisé

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

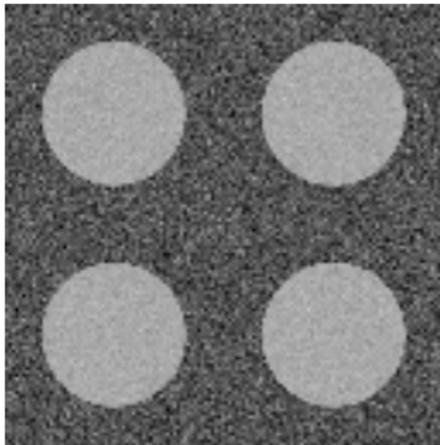
Un algo de
segmentation
non
supervisée :
k-means

Principes

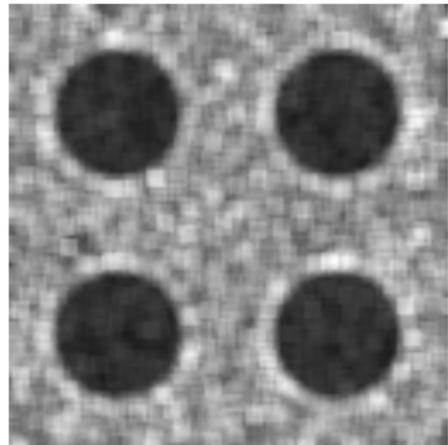
Algorithme

Exemples

- Classification des données sans connaissance a priori
- Mise en oeuvre de la méthode K-means
- Segmentation d'une image multi-variée



Niveaux de gris rdf-2-classes-
texture-1.png



Niveaux de texture rdf-2-
classes-texture-1-text.png

Apprentissage non supervisé

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

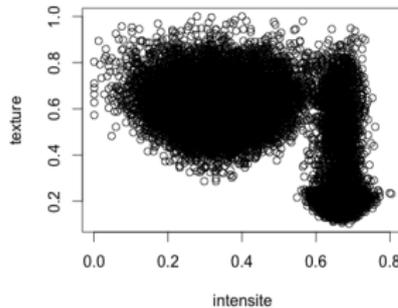
Un algo de
segmentation
non
supervisée :
k-means

Principes

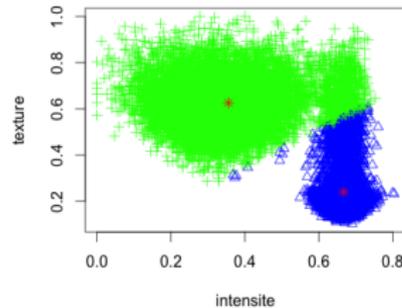
Algorithme

Exemples

- Projection des pixels dans l'espace des attributs
- Mise en oeuvre de la méthode K-means



Espace d'attributs



Observations classées en 2
classes

Apprentissage non supervisé

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

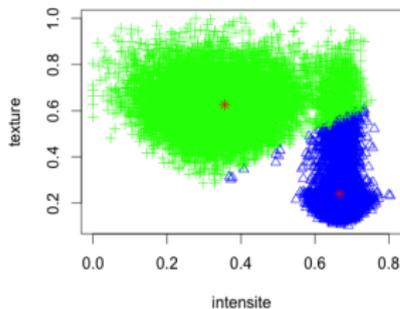
Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Classification des pixels par comparaison des distances observations-centres de gravité



Observations classées en 2 classes Image segmentée en 2 classes

Matrice des données discrètes \mathbf{X}

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes
Algorithmes
Exemples

- Soient N observations pour chacune des D attributs $X_i, i = 1, \dots, D$.
- On peut donc représenter les N observations de l'attribut x_i sous la forme d'un vecteur

$$\mathbf{x}_i = (x_{i,1}, \dots, x_{i,j}, \dots, x_{i,N})^T$$

- On peut alors rassembler les D vecteurs d'attributs \mathbf{x}_i dans une matrice \mathbf{X} de dimension $D \times N$.
- $x_{i,j}$ est donc le i eme attribut de la j eme observation.
- \mathbf{x} est une observation quelconque de dimension D .

Clustering par kmeans - Principes

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Soit K le nombre de classes C_k à retrouver, donné par l'utilisateur.
- L'algorithme kmeans va identifier les K centres de gravité $\hat{\mu}_k$ des classes.
- Ils minimisent la distance entre les points assignés à chaque classe et les centre de gravité associés :

$$\frac{1}{N} \sum_{\mathbf{x}} (\mathbf{x} - \mu_{\hat{\omega}(\mathbf{x})})^T \cdot (\mathbf{x} - \mu_{\hat{\omega}(\mathbf{x})})$$

- où $\hat{\omega}(\mathbf{x})$ est la classe d'assignation de la donnée \mathbf{x} .

Clustering par kmeans -Principes

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Minimisation de la distance entre le centre de gravité et les points assignés à chaque classe :

$$\frac{1}{N} \sum_{\mathbf{x}} (\mathbf{x} - \boldsymbol{\mu}_{\hat{\omega}(\mathbf{x})})^T \cdot (\mathbf{x} - \boldsymbol{\mu}_{\hat{\omega}(\mathbf{x})})$$

- où $\hat{\omega}(\mathbf{x})$ est la classe d'assignation de la donnée \mathbf{x} .
- $\hat{\omega}(\mathbf{x})$ est la classe dont le centre de gravité est le plus proche de \mathbf{x} :

$$\hat{\omega}(\mathbf{x}) = \operatorname{argmin}_k (\mathbf{x} - \boldsymbol{\mu}_k)^T \cdot (\mathbf{x} - \boldsymbol{\mu}_k)$$

Clustering par kmeans - Algorithme

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Données d'entrée : K et \mathbf{X} .
- Soient les positions initiales des centres de gravité des classes $\hat{\boldsymbol{\mu}}_k^{(0)}$, $k = 1, \dots, K$.
- $t = 1$
- Tant que critère d'arrêt non satisfait
 - Assignation des points aux K classes
$$\hat{\omega}(\mathbf{x})^{(t)} = \underset{k}{\operatorname{argmin}} (\mathbf{x} - \boldsymbol{\mu}_k^{(t-1)})^T \cdot (\mathbf{x} - \boldsymbol{\mu}_k^{(t-1)})$$
 - Soit $S_k^{(t)}$ l'ensemble des points assignés à la classe k
$$S_k^{(t)} = \{\mathbf{x}.tq.\hat{\omega}(\mathbf{x})^{(t)} = k\}$$
 - Mise à jour des centres de gravité des K classes
$$\boldsymbol{\mu}_k(t) = \frac{1}{|S_k^{(t)}|} \sum_{\mathbf{x} \in S_k^{(t)}} \mathbf{x}$$
 - $t = t + 1$

Kmeans- Exemple 1

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

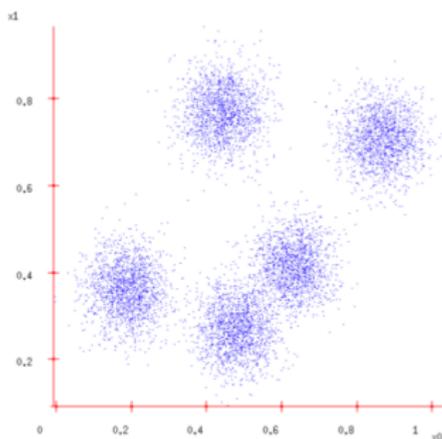
Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

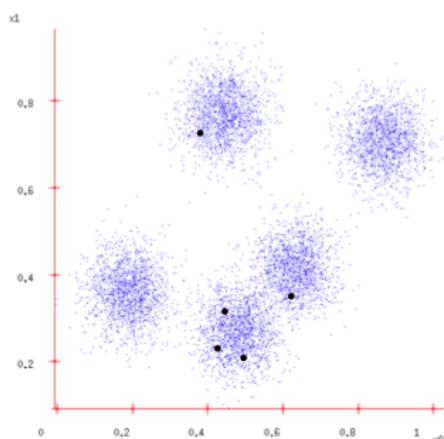
Principes

Algorithme

Exemples



Espace d'atributs



$\hat{\mu}_k^{(0)}$

Kmeans- Exemple 1

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

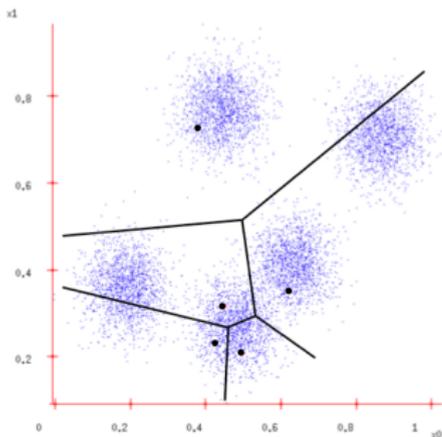
Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

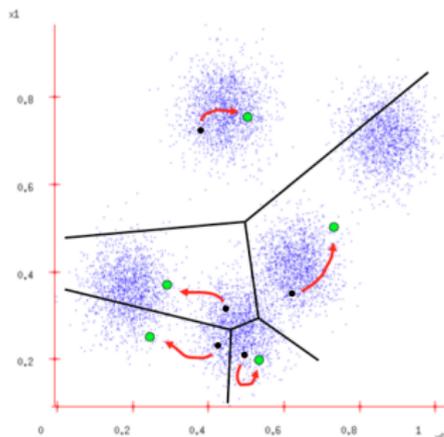
Principes

Algorithme

Exemples



Assigination



Mise à jour des centres $\hat{\mu}_k^{(1)}$

Kmeans- Exemple 2

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

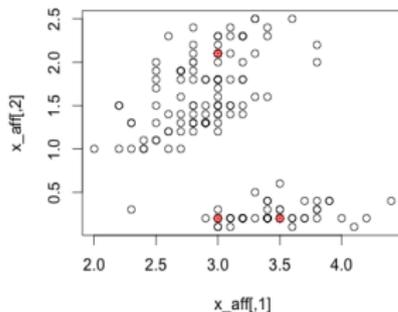
Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

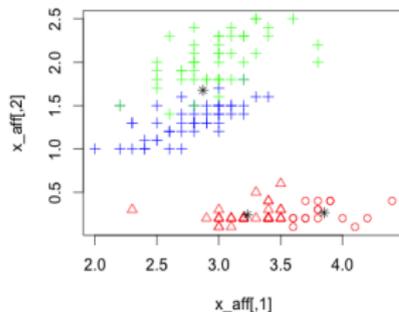
Principes

Algorithme

Exemples



Espace d'attributs et $\hat{\mu}_k^{(0)}$



t=1

Kmeans- Exemple 2

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

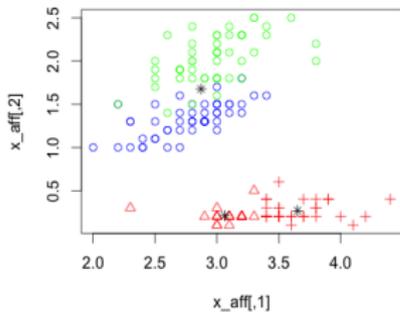
Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

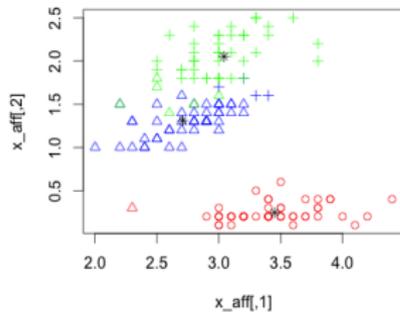
Principes

Algorithme

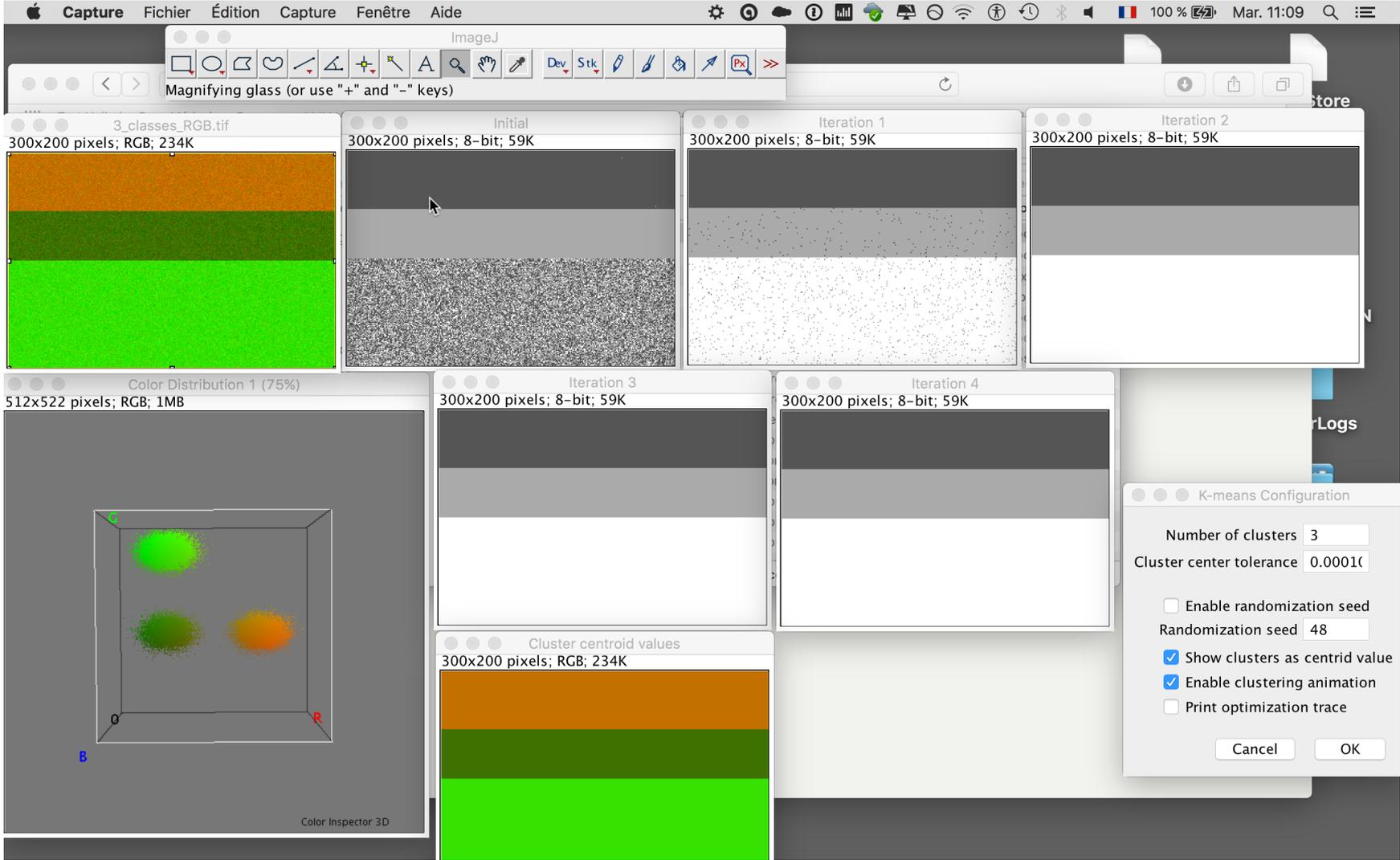
Exemples

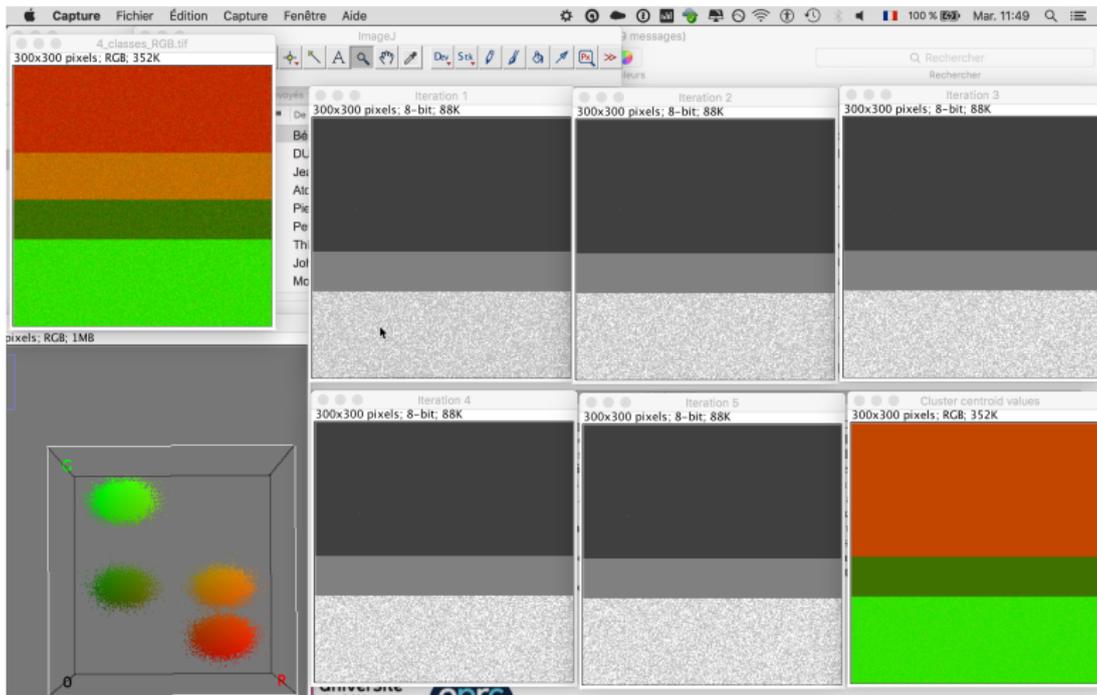


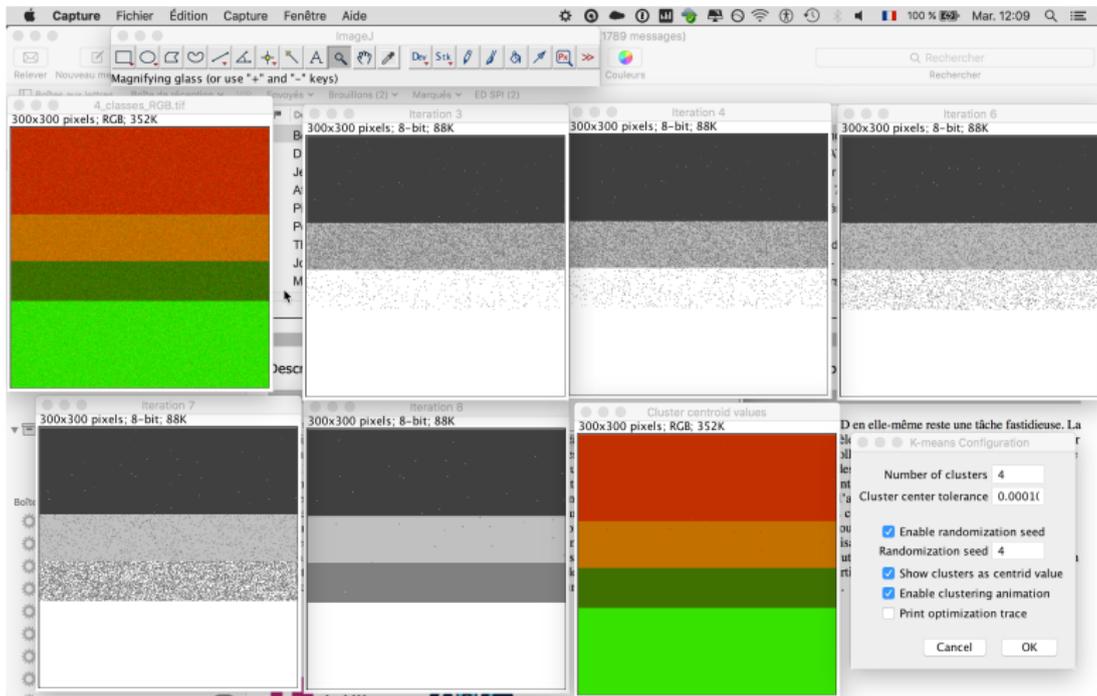
t=2



t=3







Limites de la méthode kmeans

Segmentation
par kmeans

L. Macaire -
M1 IVI - RDF
- Cours 8

Une
application
industrielle

Un algo de
segmentation
non
supervisée :
k-means

Principes

Algorithme

Exemples

- Sensible aux positions initiales des centres de gravité
- Nécessite de préciser le nombre K des classes
- Critères d'arrêt :
 - Nombre max d'itérations
 - seuil entre 2 itérations : $\sum_k (\boldsymbol{\mu}_k^{(t)} - \boldsymbol{\mu}_k^{(t-1)})^T (\boldsymbol{\mu}_k^{(t)} - \boldsymbol{\mu}_k^{(t-1)})$
- Adaptée aux nuages sphériques