

BASES DE L'IMAGERIE NUMERIQUE

1.1. INTRODUCTION

Une image est la représentation planaire d'une scène ou d'un objet situé en général dans un espace tridimensionnel (3D). Elle est acquise à l'aide de systèmes de production d'images (caméra CCD, radiographe, scanner, sonar,.....). Sa forme peut être analogique (ex: négatif, photographie, vidéo..), ou numérique (tableau de valeurs) suivant divers formats et dans ce cas un traitement par ordinateur est possible.

Ainsi donc, la transformation de l'espace 3D en un tableau de valeurs passe par trois étapes:

1. *formation* de l'image (projection) à travers un système optique;
2. *échantillonnage* temporel, spatial et spectral par le capteur;
3. *quantification* par le capteur et/ou le système de traitement.

1.2. ECHANTILLONNAGE TEMPOREL

1.2.1. CADENCE D'ACQUISITION DES IMAGES

L'information lumineuse ne peut être mesurée de façon continue dans le temps, elle est donc captée sous forme *d'échantillons*. On parle alors de mesure *périodique* de l'information lumineuse.

La *cadence* d'acquisition représente le nombre d'images acquises en une seconde (1s) et est l'inverse de la *période* d'acquisition.

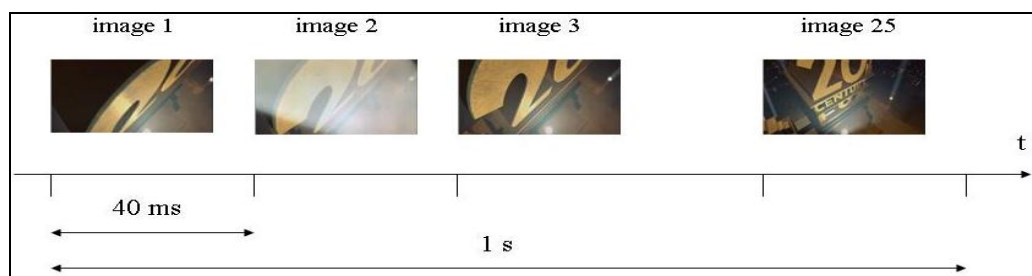


Fig.1 – Exemple d'une cadence vidéo de 25 images / s, période de 40 ms.

1.2.2. INTEGRATION DE L'INFORMATION LUMINEUSE

Théoriquement, un capteur élémentaire mesure instantanément le flux énergétique qu'il reçoit par unité de temps ($W = J \cdot s^{-1}$).

En pratique, un capteur mesure l'énergie totale (en **J**) qu'il reçoit pendant un intervalle de temps (en **s**) de durée non nulle, appelé temps *d'intégration*.

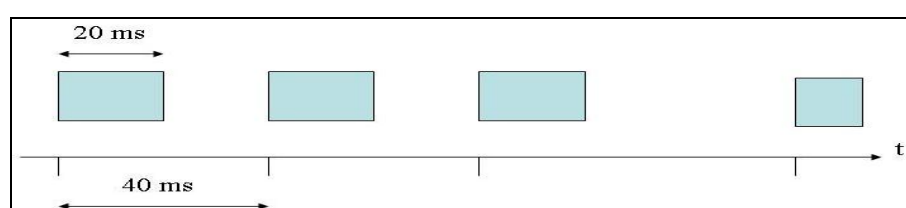


Fig.2 – Exemple d'une cadence d'acquisition de 25 images / s, durée d'intégration de 20 ms.

1.3. IMAGE CONTINUE

1.3.1. FONCTION IMAGE

Le terme image réfère à une fonction de brillance analogique continue, le plus souvent bidimensionnelle, notée $f(x,y)$. L'information dont elle est le support est caractéristique de l'intensité lumineuse (couleur ou niveaux de gris) au point (x,y) .

$$\begin{aligned} f: \mathbf{R}^2 &\rightarrow \mathbf{R}^n \\ (x,y) &\rightarrow f(x,y) \end{aligned} \quad (1.1)$$

n est le nombre de composantes de l'image ($n \geq 1$). Dans ce cas, l'espace \mathbf{R}^2 est appelé *plan image*.

1.3.2. SUPPORT DE L'IMAGE

C'est un sous ensemble de \mathbf{R}^2 de définition de $f(x,y)$, de surface *finie*, en général un *rectangle*, de côtés N_x et N_y .

1.3.3. INFORMATION IMAGE

La lumière étant une forme d'énergie, $f(x,y)$ est définie dans un domaine borné tel que

$$0 < f(x,y) < \infty \quad (1.2)$$

La nature de $f(x,y)$ est caractérisée par deux composantes essentielles:

1. la quantité de lumière incidente sur la scène observée, appelée *luminance* et notée $i(x,y)$;
2. la quantité de lumière réfléchiée par les objets contenus dans la scène, appelée *réflectance* et notée $r(x,y)$.

La fonction $f(x,y)$ est formée par le produit de $i(x,y)$ et $r(x,y)$ telle que:

$$f(x,y) = i(x,y) \cdot r(x,y) \quad (1.3)$$

où

$$0 < i(x,y) < \infty \quad (1.4)$$

et

$$0 < r(x,y) < 1 \quad (1.5)$$

Notons que (1.5) indique que la réflectance est bornée par **0** (*absorption totale*) et **1** (*réflectance totale*). La nature de $i(x,y)$ est déterminée par la source de lumière incidente, $r(x,y)$ est déterminée par les caractéristiques des objets contenus dans la scène.

1.4. IMAGE NUMERIQUE

Sous l'aspect décrit par l'équation (1.1) et (1.2), l'image est inexploitable pour un traitement par ordinateur, ce qui nécessite sa numérisation à la fois spatiale et en amplitude.

1.4.1. ECHANTILLONNAGE SPATIAL

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

Ainsi, l'échantillonnage en x et en y de la fonction f ne conserve que les valeurs pour les points $(x,y) = (i.\Delta x, j.\Delta y)$ avec i, j entiers. L'image numérique obtenue est un tableau de valeurs I tel que:

$$\begin{aligned} I: \mathbf{Z}^2 &\rightarrow \mathbf{R}^n \\ (i,j) &\rightarrow I(i.\Delta x, j.\Delta y) \end{aligned} \quad (1.6)$$

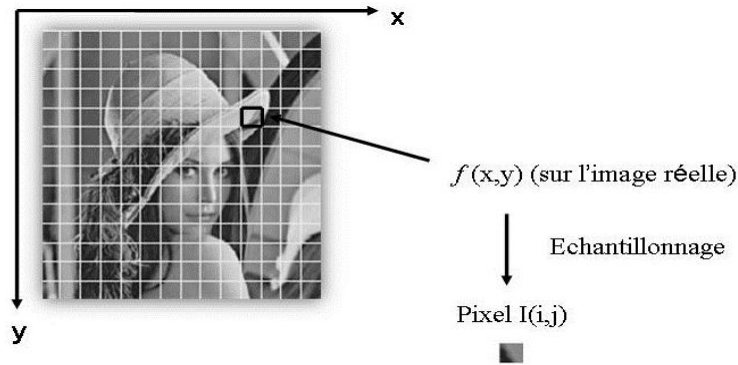


Fig.3 -- Opération d'échantillonnage avec maillage rectangulaire.

i. Pixel

Un *pixel* (contraction de *picture element*) est le nom associé à une unité de base de l'image qui correspond à un pas de discrétisation. Un pixel est caractérisé par sa position et sa valeur (son intensité). Le plus souvent, ce vocable est utilisé pour représenter indifféremment l'un ou l'autre de ces attributs.

ii. Maillage

Le *maillage* est l'arrangement géométrique des pixels dans l'image. On utilise généralement un des trois types suivants. Tout d'abord, le maillage carré qui correspond à la réalité physique du capteur CCD, et consiste à diviser uniformément une image en $M \times N$ surfaces élémentaires (pixels). On peut également utiliser un maillage hexagonal (maillage de référence pour le domaine de la morphologie mathématique). Le maillage triangulaire est plus confidentiel.

iii. Définition

La *définition* d'une image numérique correspond au nombre de pixels qui composent l'image en hauteur (axe vertical) et en largeur (axe horizontal) : *200 pixels par 450 pixels* par exemple, abrégé en « 200×450 »

iv. Résolution

La *résolution* d'une image est définie par le nombre de pixels par unité de longueur de la structure à numériser. La résolution est l'inverse du pas de la grille d'échantillonnage.

Classiquement, l'unité de mesure de la résolution est le Point Par Pouce (PPP) ou Pixel Par Pouce, utilisée principalement pour les écrans et les imprimantes (en anglais : Dots Per Inch (DPI) ou Pixels Per Inch (PPI). Note : un pouce (inch) vaut 2,54 cm.).

Ce paramètre est défini lors de la numérisation et dépend principalement des caractéristiques du matériel utilisé à cet effet. Plus le nombre de pixels par pouce est élevé, plus la résolution est élevée et plus la quantité d'information qui décrit l'image est importante. Ainsi donc, plus la résolution est élevée, meilleure est la restitution.

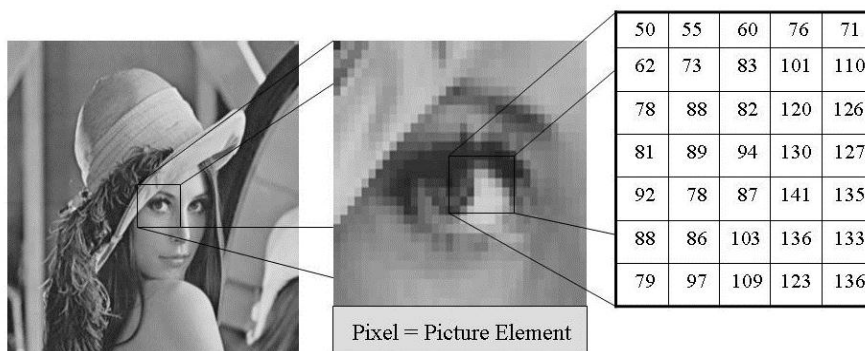


Fig.4 – Résolution spatiale.

Exemple

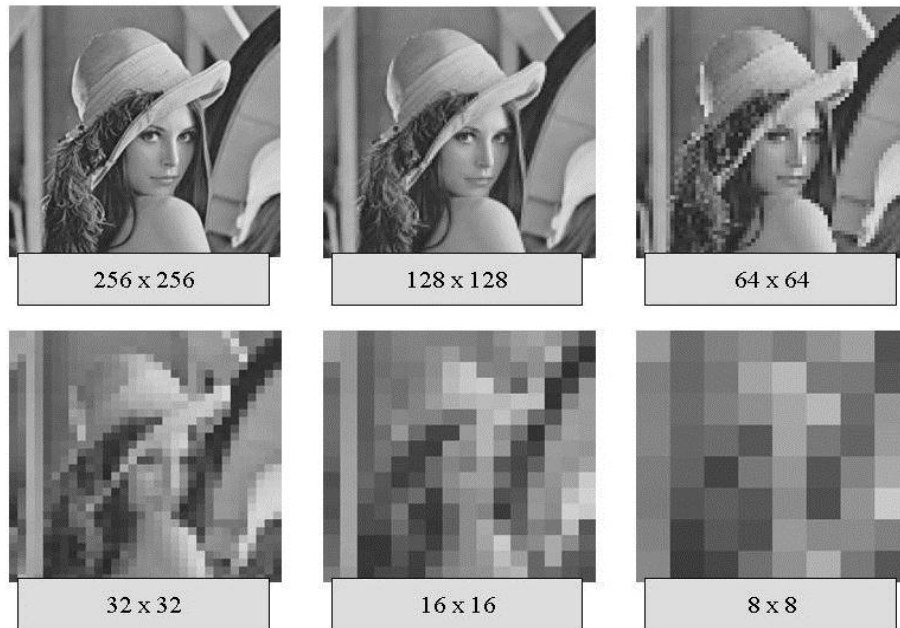


Fig.5 – Echantillonnage d'une image.

1.4.2. ECHANTILLONNAGE SPECTRAL

i. Image Monochrome: 1 échantillon

Pour chaque pixel, on mesure le *flux énergétique total*, c.a.d, pour toutes les longueurs d'onde du spectre. En pratique, un capteur monochromatique est caractérisé par une courbe de sensibilité spectrale.

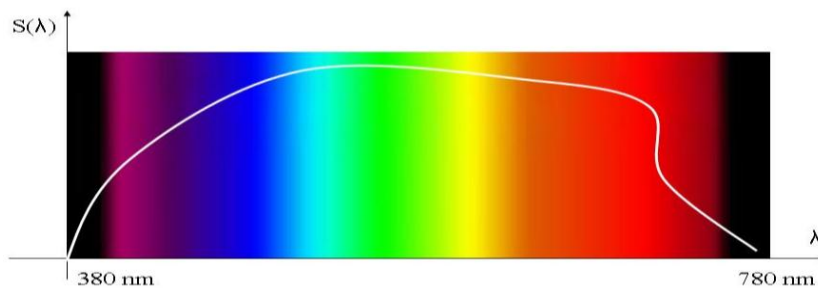


Fig.6 -- Courbe de sensibilité spectrale.

ii. Echantillonnage du Spectre en n Bandes

Pour acquérir une composante, on utilise un capteur qui a une sensibilité spectrale correspondant à une bande du spectre. Une composante est mesurée par l'intégrale du flux énergétique monochromatique, pondérée par la sensibilité spectrale.

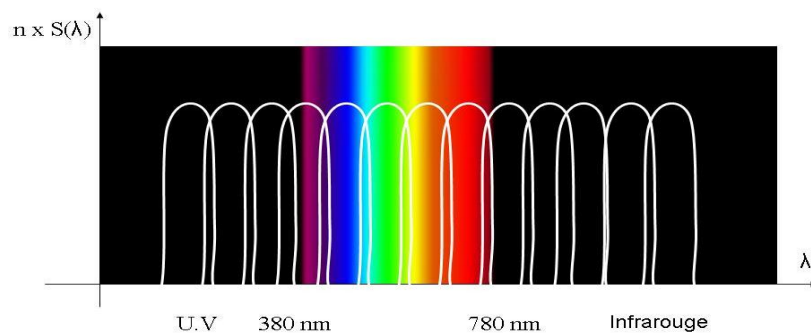


Fig.7 -- Echantillonnage du spectre en n bandes.

iii. Image Multi-Spectrale

Les bandes sont en nombre limité et sont choisies parce que leurs longueurs d'ondes ont des propriétés particulières. Par exemple, en imagerie satellitaire, la surface de la terre est observée dans le visible et l'infrarouge (végétation).

Exemple: satellite SPOT 5 (Pa, B1, B2, B3, MIR)



Fig.8 -- Exemple d'image satellite multi-spectrale.

iv. Image Couleur

On échantillonne le spectre visible dans trois bandes: Rouge, vert et Bleu. Celles-ci ont été sélectionnées afin de correspondre à la vision humaine standard.

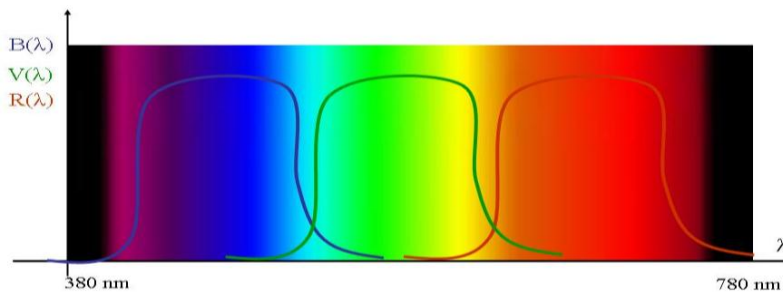
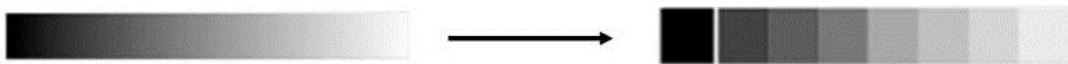


Fig.9 -- Spectre de l'image couleur.

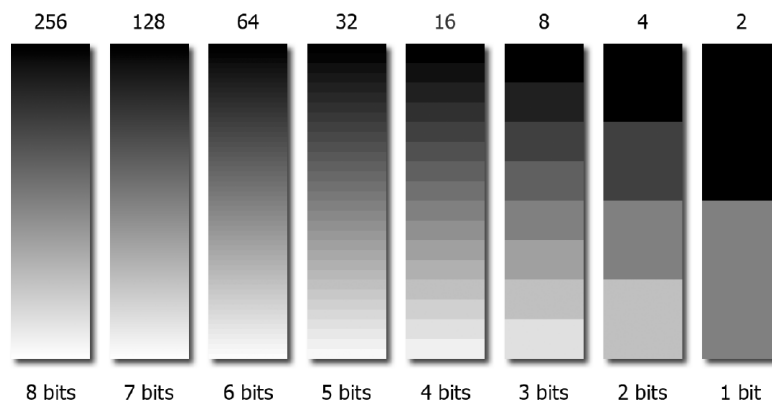
1.4.3. QUANTIFICATION

La quantification désigne la limitation du nombre de valeurs différentes que peut prendre $I(i, j)$.



i. Valeurs Numériques

Chaque valeur d'une composante de la fonction image est représentée par un mot binaire, codé sur un nombre fini de bits. Pour un mot de n bits, la valeur varie entre 0 et $2^n - 1$. Par exemple, sur 8 bits les composantes varient entre 0 et 255.



ii. Taille des Données Image

Dimension du support	Nombre de composantes	Nombre de bits de quantification
$N_x \times N_y$ pixels	m	n
D'où une taille en bits = $N_x \cdot N_y \cdot n \cdot m$		

Exemple

- Image monochrome binaire 256x256: 65536 bits = 8 Ko;
- Image spot 2048x2048, 4 canaux, 12 bits: 201.326.592 bits = 24 Mo

iii. Effet de la Quantification sur le Rendu Visuel

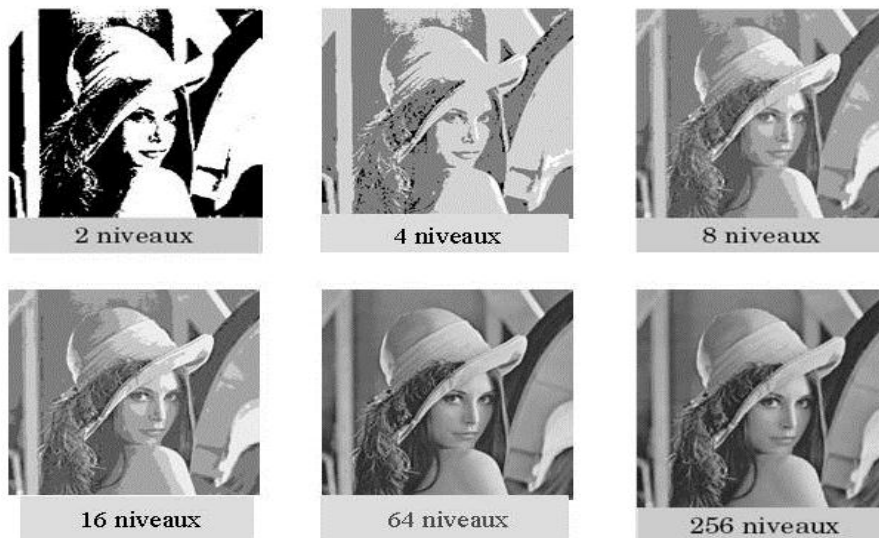


Fig.10 -- Quantification d'une image.

1.5 LA COULEUR

1.5.1. DEFINITIONS

i. Lumière Achromatique

Une lumière est dite achromatique lorsqu'elle contient toutes les longueurs d'onde de façon approximativement égales.

ii. Luminosité

C'est la quantité d'énergie du signal. Elle caractérise l'intensité de la lumière achromatique (permet de distinguer les niveaux de gris), et le niveau lumineux d'un stimulus de couleur.

iii. Teinte

C'est le nom de la couleur telle que rouge, vert, bleu, jaune, etc. Elle correspond à la *longueur d'onde dominante* d'un stimulus de couleur, notons que le blanc, le noir ou les gris sont des couleurs qui n'ont pas de teinte. Elles sont dites *neutres* ou *achromatiques*.

iv. Saturation

C'est une grandeur permettant d'estimer le *niveau de coloration* d'une teinte indépendamment de la luminosité. Elle représente l'inverse du degré de dilution de la couleur dans la lumière blanche. La saturation est un indice de *pureté* de la couleur perçue comme vive, pâle, terne, etc. Une couleur pure est saturée à 100%, le blanc et les gris le sont à 0%.

1.5.2. QUANTIFICATION

Les images couleur sont acquises sous la forme de trois composantes rouge (*R* pour *Red*), verte (*G* pour *Green*) et bleue (*B* pour *Blue*). Une image couleur numérique est en fait l'association de trois plans de niveau de gris, chacun d'eux étant une couleur de base: *R*, *G* et *B*.

En règle générale, chacune des composantes *R*, *G* et *B* est codée sur 8 bits et peut donc prendre 256 valeurs. Une couleur est donc codée sur 24 bits. Cependant, il est nécessaire, parfois, de coder chaque composante sur 12 bits afin de différencier toutes les couleurs que le système visuel humain peut percevoir.

R	G	B	Couleur
0	0	0	noir
0	0	1	nuance de noir
255	0	0	rouge
0	255	0	vert
0	0	255	bleu
128	128	128	gris
255	255	255	blanc

Fig.11 -- les couleurs d'une image

1.5.3. REPRESENTATION SUR TROIS COMPOSANTES DE COULEURS

Le principe de *trivariance visuelle*, connu sous le nom de *théorie trichromatique*, indique que trois *primaires* sont nécessaires et suffisantes pour reproduire toute couleur. Les quantités de chacune des primaires nécessaires à reproduire une couleur sont appelées les *composantes trichromatiques*. Il existe deux types de synthèse : la *synthèse additive* et la *synthèse soustractive*.

i. Synthèse Additive

La synthèse additive résulte de la juxtaposition de lumières colorées, correspondant chacune à une des trois primaires : Rouge, Vert et Bleu. Le mélange additif en quantités égales des trois primaires fournit le blanc. La synthèse additive est utilisée par la plupart des capteurs et des moniteurs.

On obtient des couleurs secondaires par ajout des couleurs primaires deux à deux. Ainsi donc, le mélange additif du rouge et du bleu produit le magenta ; celui du rouge et du vert le jaune ; celui du vert et du bleu le cyan.

ii. Synthèse Soustractive

La synthèse soustractive est utilisée en imprimerie et en peinture. Les primaires correspondant à cette synthèse sont le magenta, le cyan et le jaune. Si on soustrait la lumière bleue de la lumière blanche on obtient de la lumière jaune. Si on soustrait la lumière verte, on obtient de la lumière magenta, et si on soustrait la lumière rouge on obtient de la lumière cyan. Théoriquement, la synthèse soustractive de trois primaires doit permettre d'absorber une lumière blanche et donner ainsi le noir.

Remarque

Les couleurs primaires relatives aux deux types de synthèse sont dites *complémentaires*. La synthèse additive de deux couleurs complémentaires donne le blanc. Ainsi le bleu est complémentaire du jaune, le rouge celui du cyan et le vert celui du magenta.

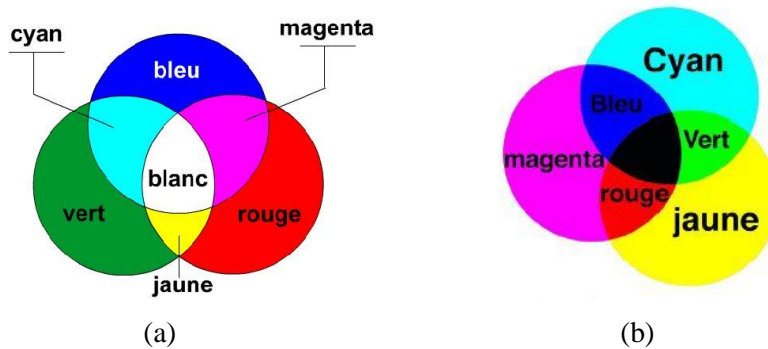


Fig.9 -- Synthèse de la couleur. (a) Additive. (b) Soustractive.

1.5.4. SYSTÈMES DE REPRÉSENTATION DE LA COULEUR

i. Espace RGB de la CIE, noté (R_C, G_C, B_C)

Le système (R_C, G_C, B_C) de la CIE a été défini en 1931. Chaque stimulus de couleur $[C]$ est représenté par un *vecteur couleur* dont les coordonnées sont les composantes trichromatiques R_C, G_C et B_C . La normalisation de ces composantes par rapport à la luminosité permet d'obtenir les composantes appelées *coordonnées trichromatiques, coordonnées réduites* ou encore *composantes normalisées*. Elles sont notées r_c, g_c et b_c et sont définies par :

$$\begin{cases} r_c = \frac{R_C}{R_C + G_C + B_C} \\ g_c = \frac{G_C}{R_C + G_C + B_C} \\ b_c = \frac{B_C}{R_C + G_C + B_C} \end{cases} \quad (1.7)$$

L'espace de représentation associé aux coordonnées trichromatiques est appelé l'espace (R_C, G_C, B_C) normalisé. Il est noté (r_c, g_c, b_c) . Comme $r_c + g_c + b_c = 1$, deux composantes suffisent à représenter la chrominance d'une couleur, et (r_c, g_c) est appelé diagramme de chromaticité.

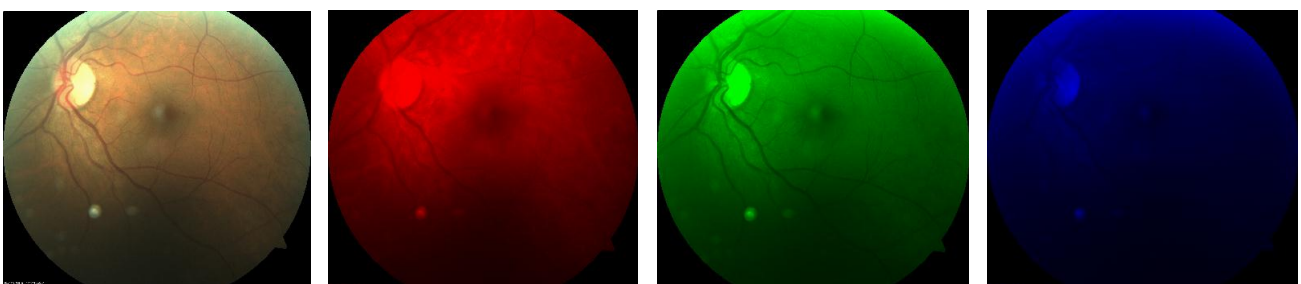


Fig.10 – Exemple d'image couleur RGB et ses composantes trichromatiques.

ii. Espace XYZ de la CIE

L'espace RGB présente quelques inconvénients majeurs dont :

- les coordonnées et les composantes trichromatiques peuvent prendre des valeurs négatives ;
- les valeurs des composantes trichromatiques sont liées à la luminance qui est une combinaison linéaire des composantes trichromatiques et non une composante elle-même ;
- il existe une multitude de systèmes $(R G B)$.

L'espace XYZ (CIE 1931) est défini à partir d'une transformation linéaire sur l'espace RGB. Les coordonnées XYZ des couleurs naturelles ont ainsi toujours des valeurs positives. La transformation d'un système RGB vers le système de référence colorimétrique (sans normalisation) s'exprime sous forme matricielle de la façon suivante :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{P} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad \text{avec} \quad \mathbf{P} = \begin{pmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{pmatrix} \quad (1.8)$$

Par exemple

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.430574 & 0.341550 & 0.178325 \\ 0.222015 & 0.706655 & 0.071330 \\ 0.020183 & 0.129553 & 0.939180 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1.9)$$

Le système (X, Y, Z) normalisé noté (x, y, z) est exprimé par :

$$x = \frac{X}{X + Y + Z} ; \quad y = \frac{Y}{X + Y + Z} ; \quad z = \frac{Z}{X + Y + Z} \quad (1.10)$$

Comme $x + y + z = 1$, deux composantes suffisent à représenter la chrominance d'une couleur, et (x, y) est appelé diagramme de chromaticité.

iii. Teinte – Saturation - Lumière

HSI (Hue-Saturation-Intensity), **HSV** (Hue-Saturation-Value) et **HSL** (Hue-Saturation-Lightness) sont une famille de systèmes basés sur les trois paramètres perceptifs de description d'une couleur, plus intuitive pour le système visuel humain (SVH): *Teinte*, *Saturation*, *Luminosité*. Ces espaces représentent les couleurs dans des systèmes de coordonnées cylindriques plus aisés à manipuler. Ils sont basés sur les mêmes principes, avec une mise en œuvre différente. Ils sont très utilisés en graphisme informatique et en infographie.

🚦 La teinte (**H**) est codée suivant l'angle qui lui correspond sur le cercle des couleurs.

- 0° ou 360° : **Rouge** ;
- 60° : **Jaune** ;
- 120° : **Vert** ;
- 180° : **Cyan** ;
- 240° : **Bleu** ;
- 300° : **Magenta**.



Fig.11 – Cercle chromatique avec graduation de l'angle pour définir la teinte.

🚦 La saturation (**S**) est l'intensité de la couleur :

- elle varie entre 0 et 100 % ;
- elle est parfois appelée *pureté* ;
- plus la saturation d'une couleur est faible, plus l'image sera *grisée* et plus elle apparaîtra fade

✚ Conversion RGB – HSV/HSI/HSL

La transformation d'un système **RGB** vers les systèmes **HSV/HSI/HSL** s'effectue conformément aux équations ci-dessous :

$$R' = R / 255; \quad G' = G / 255; \quad B' = B / 255; \quad (1.11)$$

$$C_{max} = \max(R', G', B'); \quad C_{min} = \min(R', G', B'); \quad \Delta = C_{max} - C_{min}; \quad (1.12)$$

$$H = \begin{cases} 0, & \Delta = 0 \\ 60 \left(\frac{G' - B'}{\Delta} \bmod 6 \right), & C_{max} = R' \\ 60 \left(\frac{B' - R'}{\Delta} + 2 \right), & C_{max} = G' \\ 60 \left(\frac{R' - G'}{\Delta} + 4 \right), & C_{max} = B' \end{cases} \quad (1.13)$$

$$V = C_{max}; \quad I = (R' + G' + B') / 3; \quad L = (C_{max} + C_{min}) / 2; \quad (1.14)$$

$$\text{Pour HSV: } S = \begin{cases} 0, & C_{max} = 0 \\ \frac{\Delta}{C_{max}}, & C_{max} \neq 0 \end{cases} \quad (1.15)$$

$$\text{Pour HSI: } S = \begin{cases} 0, & I = 0 \\ 1 - \frac{C_{min}}{I}, & I \neq 0 \end{cases}$$

$$\text{Pour HSL: } S = \begin{cases} 0, & \Delta = 0 \\ \frac{\Delta}{1 - |2L - 1|}, & \Delta \neq 0 \end{cases}$$

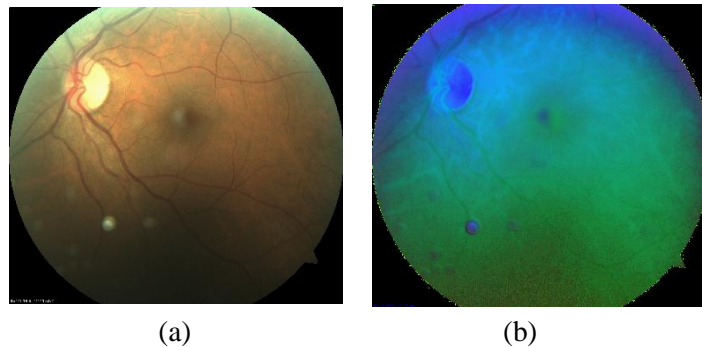


Fig.12 – Exemple de conversion RGB en (a) vers HSV en (b).

iv. Espaces YUV, YIQ et YCbCr

YUV, YIQ et YCbCr sont trois espaces semblables. Leur principe est de représenter les couleurs en utilisant une composante de luminance **Y**, et deux composantes de chrominance : (**U,V**) pour l'espace **YUV** ; (**I,Q**) pour l'espace **YIQ** et (**Cb,Cr**) pour l'espace **YCbCr**.

YUV est utilisé pour le codage des couleurs dans le système de télévision PAL, **YIQ** pour le système NTSC. Ces deux systèmes sont dépendants des appareils. De ce point de vue l'espace **YCbCr** est plus intéressant. **Y** correspond à la luminance, **Cb** et **Cr** aux compléments bleus et rouges. La dernière recommandation concernant le système **YCbCr** provient de l'ITU et définit les conversions **RGB / YCbCr** ainsi:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.2220 & 0.7067 & 0.0713 \\ -0.1195 & -0.3810 & 0.5000 \\ 0.5000 & -0.4542 & -0.0458 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1.16)$$

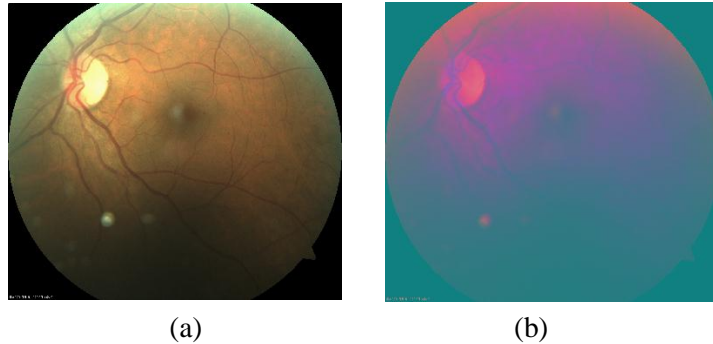


Fig.13 – Exemple de conversion RGB en (a) vers YCbCr en (b).

1.6 HISTOGRAMME

L'histogramme de l'image **I**, noté $H(g)$, traduit le nombre d'apparitions du niveau de gris g dans l'image **I**. L'image est associée à une modélisation probabiliste dont l'histogramme représente aussi la densité de probabilité de la variable niveau de gris:

$$p(g) = \frac{n_g}{N} \quad g = 0, 1, \dots, G-1 \quad (1.17)$$

où n_g représente le nombre d'occurrence du niveau de gris g , N est le nombre total des pixels de l'image et G est le nombre total des niveaux de gris. $p(g)$ est une estimation de la probabilité d'occurrence du niveau de gris g . La représentation de la densité de probabilité $p(g)$ en fonction de g fournit une description globale de l'apparence de l'image, comme dans le cas de la figure 14 qui illustre une image médicale en niveaux de gris avec son histogramme.

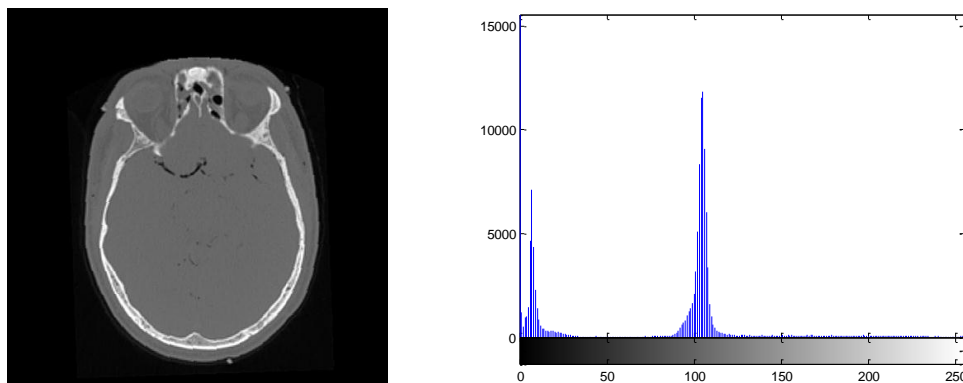


Fig.14 – Image CT du cerveau et son histogramme correspondant.

Pour une image couleur, plusieurs histogrammes peuvent être utilisés : a) l'histogramme des luminances ; b) les trois histogrammes correspondant à chacune des composantes R, G et B, comme dans l'exemple illustré par la figure 15.

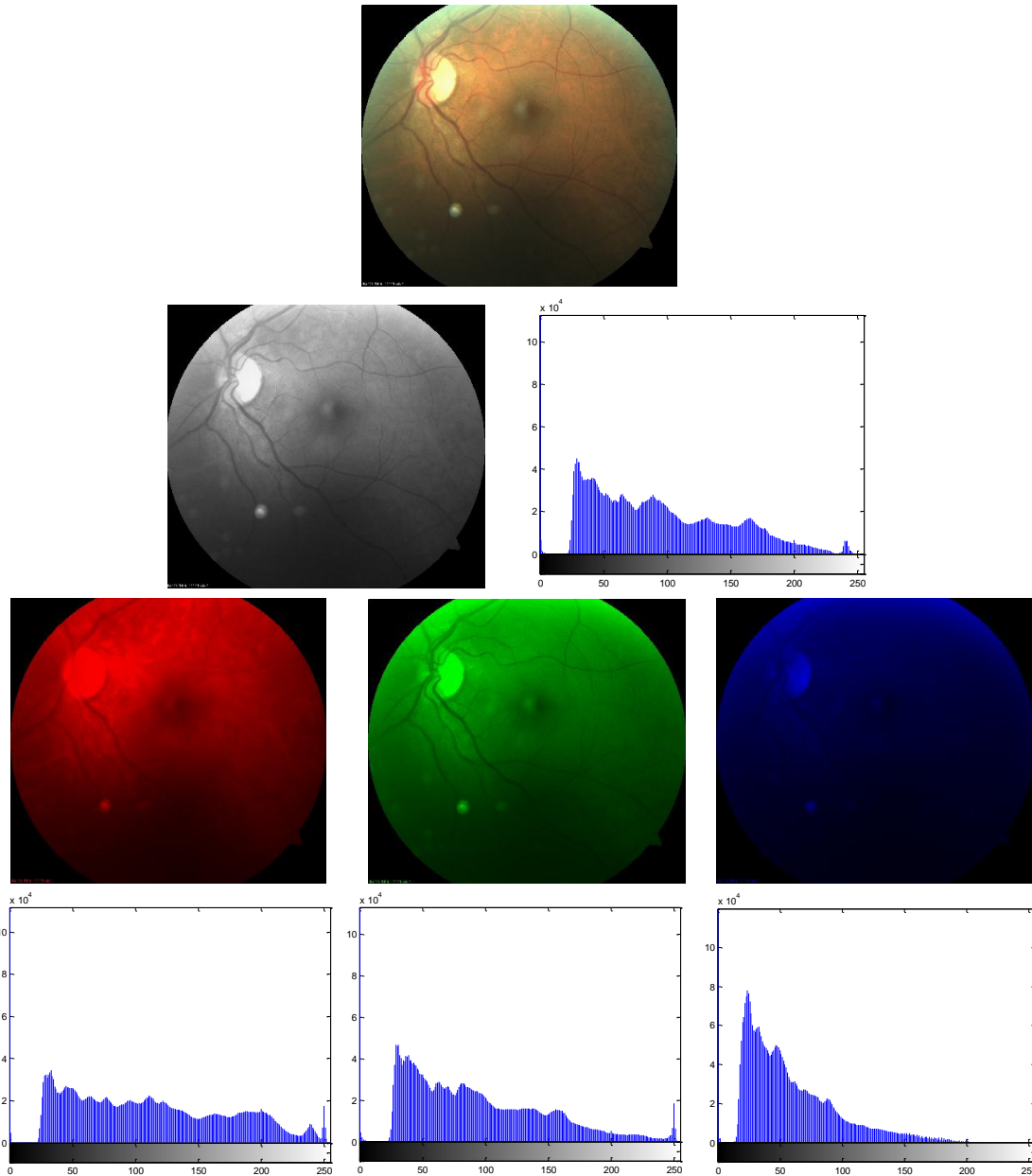


Fig.15 – Image CT du cerveau et son histogramme correspondant.

1.7 VOISINAGE

Un pixel P aux coordonnées (i, j) possède quatre voisins horizontaux et verticaux qui forment ce qu'on appelle le 4-voisinage de P . Leurs coordonnées sont:

$$(i, j-1), (i, j+1), (i-1, j), (i+1, j)$$

Si on considère un pixel comme un carré élémentaire, les pixels présentant un côté commun avec (i, j) sont appelés ses 4-voisins. Le pixel (i, j) possède aussi quatre voisins diagonaux. Ceux-ci possèdent un sommet commun avec (i, j) . Leurs coordonnées sont:

$$(i-1, j-1), (i-1, j+1), (i+1, j-1), (i+1, j+1)$$

L'ensemble des huit voisins du pixel (i, j) sont appelés les 8-voisins et forment ce qu'on appelle le 8-voisinage. Ces huit voisins forment la fenêtre 3x3 du pixel (i, j) . On dit aussi que ces pixels sont 8-adjacents de (i, j) .

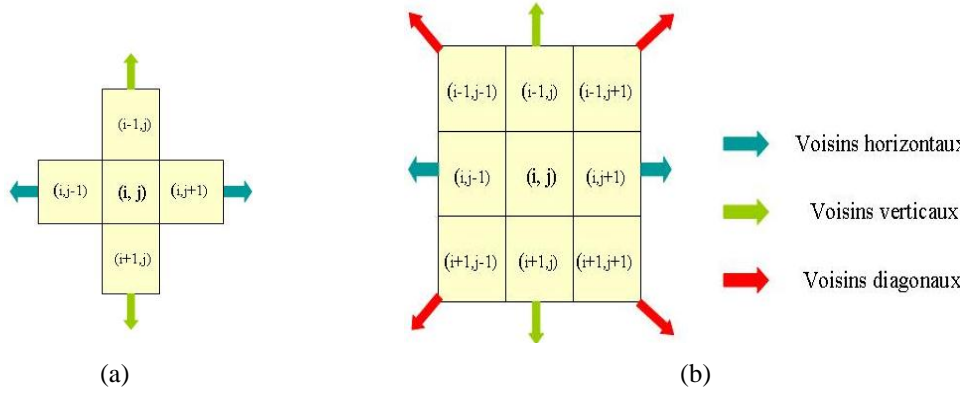


Fig.16 -- Systèmes de voisinage. (a) 4-voisinage. (b) 8-voisinage.

1.8 DISTANCE

Les distances les plus courantes entre deux pixels P et Q aux coordonnées respectives (x_p, y_p) et (x_q, y_q) :

- *Distance euclidienne*

$$d_1(P, Q) = \left[(x_p - x_q)^2 + (y_p - y_q)^2 \right]^{1/2} \quad (1.18)$$

Pour cette mesure de distance, les pixels ayant une distance inférieure ou égale à une valeur r à partir de (x_p, y_p) appartiennent au disque de rayon r centré en (x_p, y_p) .

- *Distance de Manhattan*

$$d_2(P, Q) = |x_p - x_q| + |y_p - y_q| \quad (1.19)$$

Pour cette mesure de distance, les pixels ayant une distance inférieure ou égale à une valeur r à partir de (x_p, y_p) forment un diamant centré en (x_p, y_p) .

- *Distance de l'échiquier*

$$d_{\text{inf}}(P, Q) = \text{Max}(|x_p - x_q|, |y_p - y_q|) \quad (1.20)$$

Pour cette mesure de distance, les pixels ayant une distance inférieure ou égale à une valeur r à partir de (x_p, y_p) forment un carré centré en (x_p, y_p) .

Ces distances sont reliées par la propriété :

$$d_{\text{inf}}(P, Q) \leq d_1(P, Q) \leq d_2(P, Q) \quad (1.21)$$