

Université Echahid Hamma Lakhder El Oued

Intitulé du Master : Système Distribué et l'Intelligence Artificielle (SDIA)

Intitulé de la matière : Image, son, vidéo codage et transmission

Objectifs de la matière :

Le multimédia numérique modifie profondément le paysage médiatique, fournit de nouveaux moyens de communications interpersonnels, des moyens de surveillance et de contrôle biométrique. Cette matière présente des fondements scientifiques et technologiques qui ont permis ces évolutions. A cet effet, nos objectifs seront les suivants :

- 1- Traitement de la représentation, les traitements élémentaires, l'analyse, la compression et la transmission des données multimédia numériques audiovisuelles.
- 2- Montrer, à l'aide d'outils algorithmiques, statistiques et probabilistes, algébriques, géométriques, de modèles psycho-visuels ou psycho-acoustiques, des caractéristiques des transmissions sur réseaux, les mécanismes principaux sur lesquels s'appuient les technologies de compression d'image, de vidéo, d'audio et de parole, de transmission du multimédia sur réseau avec des contraintes de temps-réel, d'amélioration visuelle des contenus.
- 3- Discuter le lien entre les aspects scientifiques, technologiques et industriels (norme, propriété intellectuelle,...).

Sommaire

Préambule.

- 1- Représentations numériques, couleur.
- 2- Analyse de Fourier, convolution, filtrage.
- 3- Éléments de théorie de l'information (entropie), MICD.
- 4- Compression de type JPEG/JPEG2000, psycho visuel.
- 5- Compression de vidéo (estimation/compensation de mouvement), aspects industriels normatifs (MPEG, H26x).
- 6- Scalabilité, codage robuste, RTP/RTCP, gestion des erreurs de transmission.
- 7- Compression de la parole : production de la parole, codage de type CELP (Applications : voix sur IP, mobiles).
- 8- Compression audio : bancs de filtrage et masquage (codage de type MP3/Ogg).

Préambule

Un environnement multimédia est un ensemble de matériels et de logiciels permettant:

- De créer
- De stocker et d'organiser
- De consulter et modifier en temps réel

– De transmettre

Des documents structurés de façon homogène comportant des informations textuelles, de la voix, des images fixes ou animées, codées numériquement.

L'information qui transite sur les réseaux de télécommunication consiste en messages de types divers : textes, sons, images fixes ou animées, vidéo, etc. La forme que revêt cette information est commode pour une communication directe et classique (conversation, échange sur papier, ...) lorsque les interlocuteurs sont en présence. Quand ils sont distants l'un de l'autre, l'emploi des réseaux de télécommunication est une manière moderne de résoudre la transmission d'informations. Toutefois, pour les nécessités du transport, la transmission d'un message nécessite un encodage en signaux de type électrique ou électromagnétique : l'émetteur et le récepteur sont, de nos jours, des ordinateurs. La voie de transmission peut être une simple liaison directe entre émetteur et récepteur ou beaucoup plus complexe dans le cadre d'un ou plusieurs réseaux de télécommunications. Les signaux sont les véhicules de transport de l'information.

Nous distinguons divers application du multimédia dont nous pouvons citer :

- Compression d'images
- Amélioration visuel (Télévision)
- Imagerie Médicale
- Identification empreinte digitale & Reconnaissance de visage
- Authentification
- Aide à la conduite
- Imagerie satellitaire
- Indexation
- Production audio visuelle numérique (CD, film, vidéo)
- Edition Hors ligne (CD/DVD)
- Edition En ligne (site web, télévision interactive)
- Les dispositifs technique du spectacle vivant (danse, concert, théâtre, installations interactives)

Chapitre I : Représentations numériques, couleur

L'information existe sous des formes diverses. Pour la manipuler et, en particulier, la transporter, nous sommes amenés à la coder. De nos jours, l'information est souvent présentée dans des documents composites, comme une page Web, où simultanément peuvent être présentés : un texte, une image fixe, un clip vidéo. L'information est, en effet, présentée sous forme multimédia. Chaque type d'information possède son système de codage, mais le résultat est le même : une suite de 0 et de 1. Le transport de l'information consiste alors à transmettre des bits, quelque soit la signification du train de bits transmis.

Nous pouvons affirmer que l'information existe sous des formes diverses, citons :

- **Image fixe**

Système : télécopie (appareil électronique convertissant l'image de documents en impulsions électriques pour les transmettre à un destinataire au travers d'une ligne téléphonique. À la réception, nous utilisons un appareil similaire à celui de l'émission pour faire la conversion inverse et réaliser l'impression d'un document identique à l'original)

Codeur : scanner

Décodeur : interpréteur de fichier

Transmission : signaux analogiques et numériques

- **Parole**

Système : téléphone

Codeur : microphone

Décodeur : écouteur

Transmission : signaux analogiques et Numériques

- **Télévision**

Système : télévision hertzienne (correspondant à une diffusion par l'intermédiaire d'ondes électromagnétiques, en effet, les signaux de télévision ou Radio (FM) transitent dans l'espace, sans support matériel, puis seront reçus par une antenne appropriée)

Codeur : caméra

Décodeur : récepteur TV + antenne

Transmission : signaux analogiques

- **Données informatiques**

Système : réseaux de téléinformatique (pour effectuer la transmission des informations d'un ordinateur à un autre ordinateur via le réseau téléphonique, le signal sort tout d'abord de la carte réseau de l'ordinateur, pénètre dans un modem, puis du modem vers le réseau téléphonique, puis du réseau téléphonique vers un autre modem lié à la carte réseau de l'autre ordinateur)

Codeur : contrôleur de communication + ETCD (un ETCD permet d'adapter le flux des données aux conditions de la ligne ainsi que de réaliser la transformation analogique/ numérique ou bien numérique/ analogique)

Décodeur : contrôleur de communication + ETCD

Transmission : signaux analogiques ou numériques

1- Traitement d'image et intelligence artificielle

Dans le domaine de la vision artificielle, nous distinguons le traitement d'images juste après les étapes d'acquisition et de numérisation dans le but d'assurer les transformations d'images ainsi que la phase de calcul qui permet la réalisation d'une interprétation des images traitées. En effet, l'appel à l'intelligence artificielle est nécessaire pour la manipulation des connaissances sur les informations dont nous disposons à propos de ce que représentent les images traitées.

Dans la conception moderne de l'Intelligence Artificielle, la machine agit sur le monde extérieur, éventuellement se déplace, et aussi perçoit son environnement pour pouvoir s'y adapter. En effet, le traitement d'image ainsi que la vision par ordinateur se basent sur des connaissances et des techniques d'intelligence artificielle dans la gestion et la prise de décision.

Le traitement d'image consiste en à partir des informations initiales, de chercher souvent à extraire des informations plus pertinentes, et à les interpréter. Nous distinguons ainsi des prétraitements, des traitements « bas-niveau » (appliqués aux pixels, sans leur donner de signification) et « haut-niveau » : analyse et interprétation.

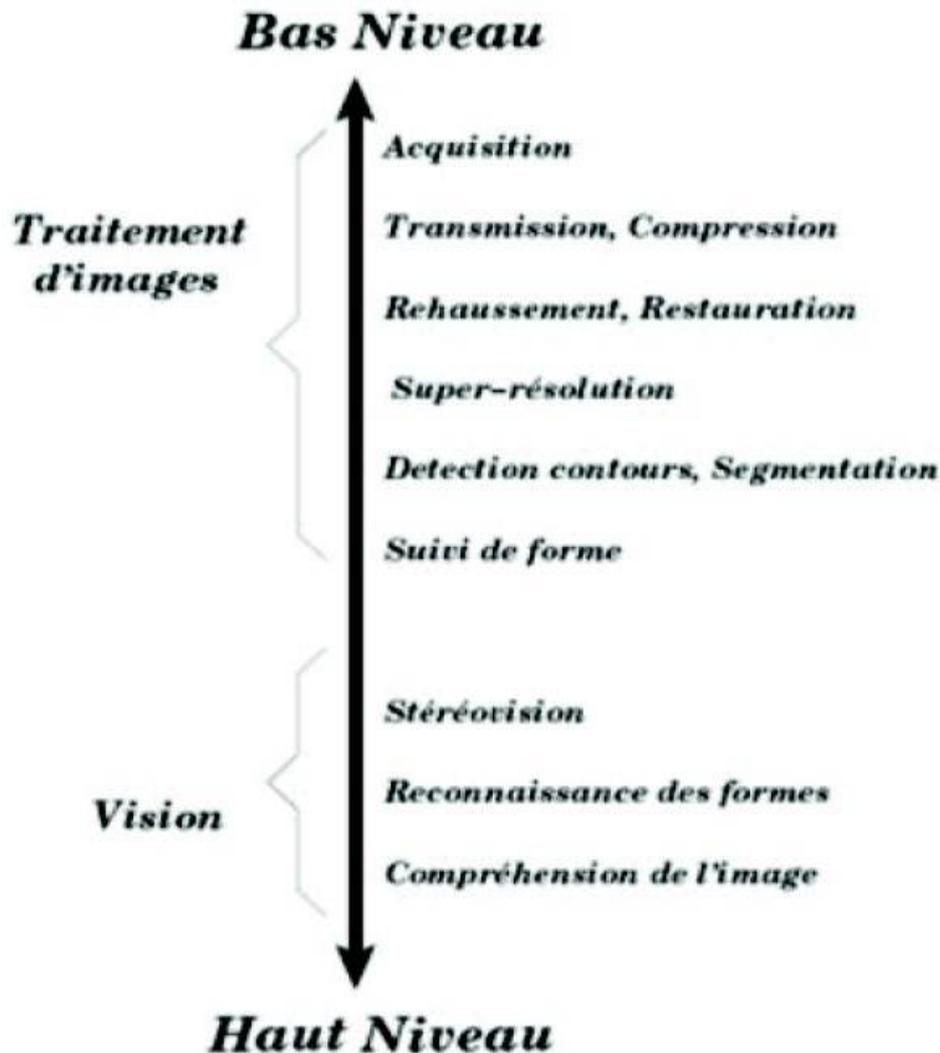


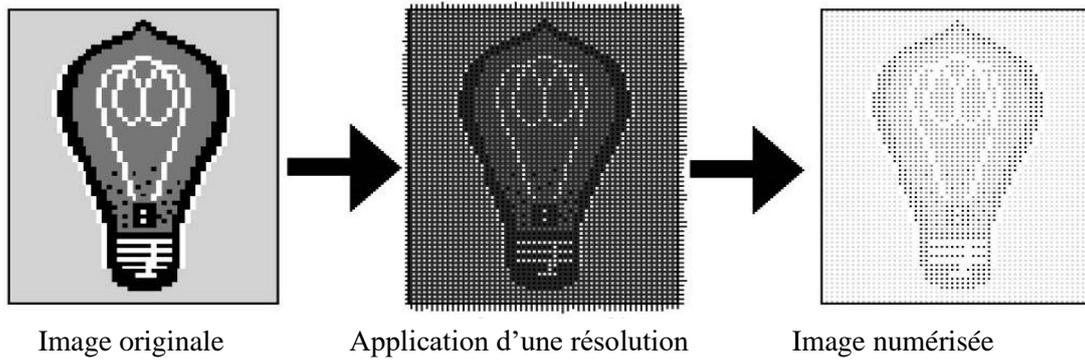
Figure I.1 : Les niveaux du traitement d'images

2- L'image numérique

2.1 Acquisition

Une image est une représentation planaire d'une scène ou d'un objet situé dans un espace à 3 dimensions (+ la dimension temporelle). L'acquisition, le traitement et la restitution d'une image s'apparentent à une chaîne de traitement (traitement d'un signal) avec toutes ses problématiques notamment les capteurs: capteurs (numériseur).

Une image numérique est constituée d'une grille rectangulaire d'échantillonnage dont les constituants sont des pixels portant des informations sur l'intensité lumineuse des différents lieux au sein de l'image. De ce fait, Une image numérique est représentée par une matrice bidimensionnelle, dont les éléments sont des nombres naturels correspondant à des niveaux de quantification dans l'échelle de l'intensité lumineuse. Tout d'abord une image d'entrée est capturée par une caméra de télévision en (2D) et numérisée, étant décrite par une fonction d'image $f(x, y)$ dont la valeur est l'intensité lumineuse en fonction de deux paramètres x, y , représentant les coordonnées de l'emplacement dans l'image.



Une image n'est qu'une représentation imparfaite d'une scène et élaborer une image correspond à la volonté de proposer une entité observable par l'œil humain.

3- Composition et caractéristique de l'image numérique

3.1 Le pixel

Une image numérique est constituée d'un ensemble de points appelés pixels (abréviation de PICTURE Element) pour former une image. Le pixel représente ainsi le plus petit élément constitutif d'une image numérique. L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image comme mentionné sur la figure I.2.

Nous pouvons considérer qu'un pixel n'a pas de dimension car si nous regardons une photo, un pixel d'un objet situé près de la caméra correspond à une taille plus petite qu'un pixel d'un objet situé loin de la caméra, même si ces deux pixels font partie de la même image.

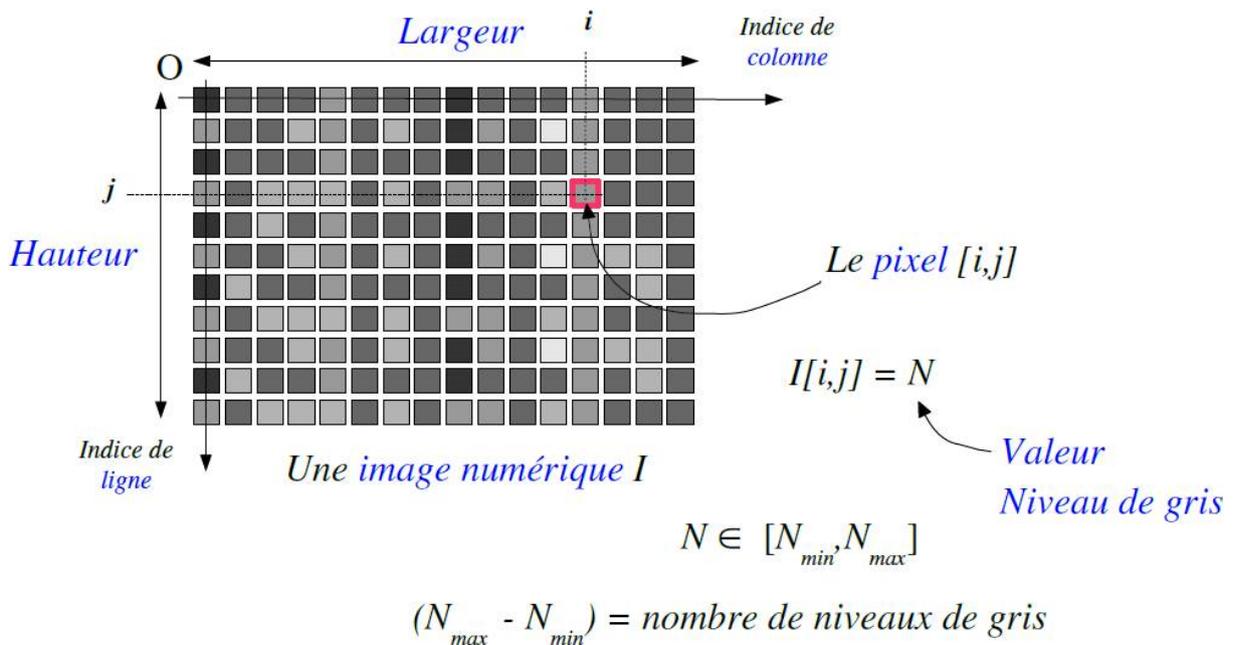


Figure I.2 : Image numérique

3.2 La définition

Nous appelons définition le nombre de points (pixels) constituant une image: c'est le nombre de colonnes de l'image que multiplie son nombre de lignes. Une image possédant 10 colonnes et 11 lignes aura une définition de 10 x11.

- **Les Principaux formats d'affichage:**

L'image s'affiche sur un écran (appelé aussi moniteur), il s'agit d'un périphérique de sortie permettant de fournir une représentation visuelle. Ces informations proviennent de l'ordinateur, mais de façon " indirecte ". En effet le processeur n'envoie pas directement les informations au moniteur, mais traite les informations provenant de sa mémoire vive (RAM), puis les envoie à une carte graphique qui est chargée de convertir les informations en impulsions électriques qu'elle envoie au moniteur.

Pour afficher ces images, des formats d'affichages standards ont été définis en fonction de l'évolution des capacités matérielles des cartes graphiques et des écrans. Voici les principaux:

CGA (320 x 200) 4 couleurs.

VGA (640 x 480) 16 couleurs.

SVGA (800x600) 256 couleurs.

XGA (1024 x 768) 256 couleurs.

SXGA (1280 x 1024) en 16 millions de couleurs.

- **Calcul du nombre total des pixels dans une image:**

Nombre total des pixels = colonnes x lignes.

Ex: $10 \times 11 = 110$ pixels au total pour l'image ci-dessus.

3.3 La résolution

C'est le nombre de pixels que l'on peut loger sur une longueur donnée (en pouce). Elle est exprimée en points par pouce (PPP, en anglais: DPI pour Dots Per Inch). Prenons l'exemple d'une image de résolution 20 pixels par cm : la taille de chaque pixel est telle que sur une ligne de 1cm nous pouvons en loger 20 (la taille des pixels étant le plus souvent carré, nous logons le même nombre de pixels sur les deux axes). Un deuxième exemple concerne une image de résolution 40 pixels par cm, il y a donc deux fois plus de pixels pour chaque centimètre linéaire et donc chaque pixels est deux fois plus petit. A cet effet, il est aisé de comprendre que plus le pixel est petit meilleure sera la qualité de l'image.

Un pouce mesure 2.54 cm, c'est une unité de mesure britannique. La résolution permet ainsi d'établir le rapport entre la définition en pixels d'une image et la dimension réelle de sa représentation sur un support comme mentionné sur la figure I.3.

Plus il y a de pixels par pouce et plus il y aura d'information dans l'image car la résolution d'une image définit la netteté et la qualité d'une image.

A titre de comparaison, une image à 300 PPP est seize fois plus lourde que la même image à 75 PPP.

Il est utile de mentionner qu'un pixel n'a pas de taille prédéfinie, donc si nous affichons une image à l'écran, chaque pixel constitutif de l'image occupera un pixel de l'écran.

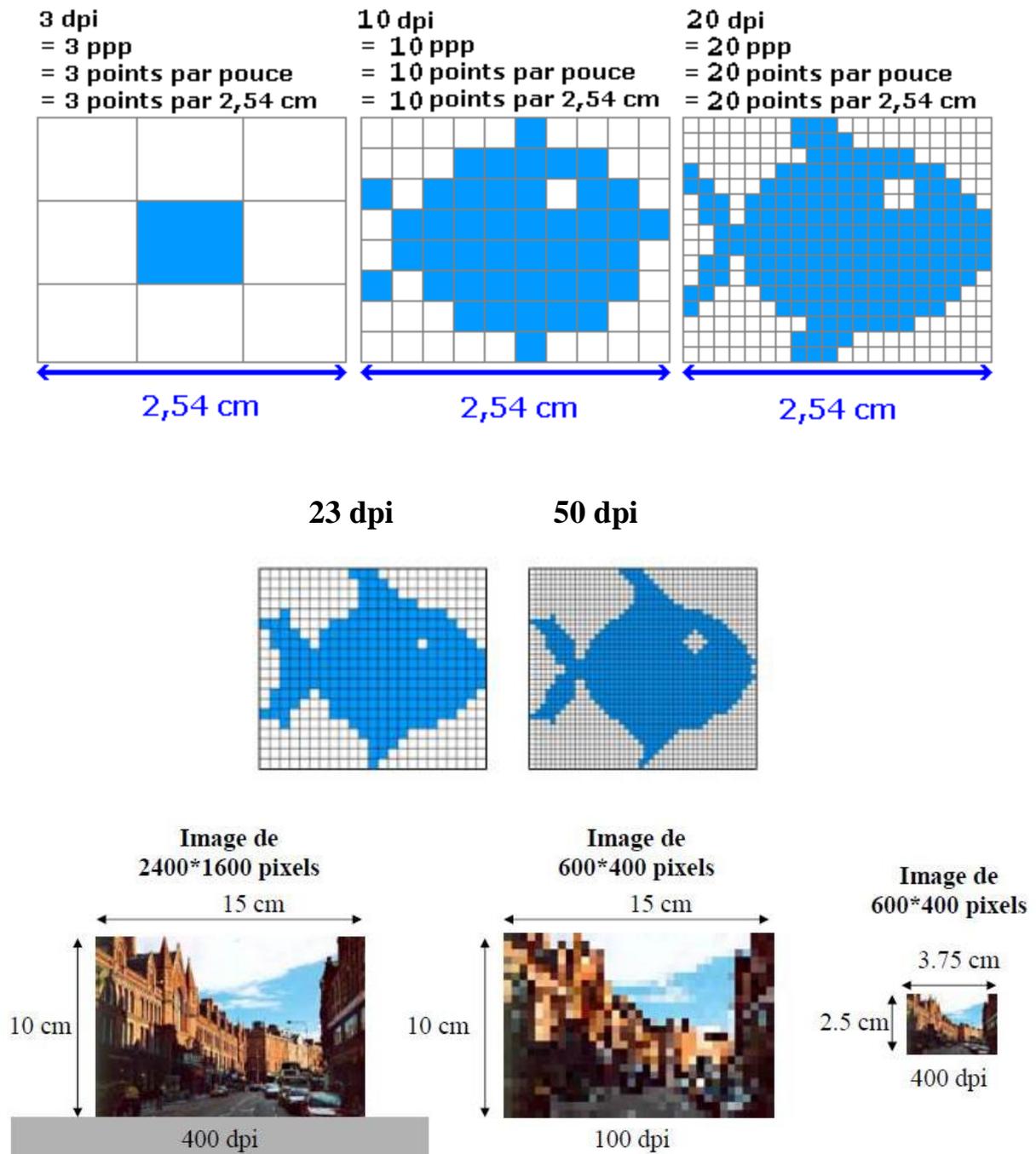


Figure I.3 : Diverses résolutions d'image

La résolution est souvent confondu avec la " définition ", détermine par contre le nombre de points par unité de surface ; un pouce représentant 2.54 cm. La résolution permet ainsi d'établir le rapport entre le nombre de pixels d'une image et la taille réelle de sa représentation sur un support physique. Une résolution de 300 dpi signifie donc 300 colonnes et 300 rangées de pixels sur un pouce carré ce qui donne donc 90000 pixels sur un pouce carré.

- **Calculer la résolution à partir de la définition et de la dimension**

Exemple :

Quelle serait la définition en pixel d'une feuille scannée d'une largeur de 8,5 pouces sur une hauteur de 11 pouces en 300dpi?

Réponse:

$300 \times 8,5 = 2550$ pixels

$300 \times 11 = 3300$ pixels

La définition de l'image serait donc de 2550 X 3300 pixels.

- **Quelques exemples de résolutions fréquemment utilisées**

Ecrans d'ordinateur: 72 dpi. C'est aussi dans cette résolution que sont les images sur Internet. Résolution non adapté à l'impression!

- fax: en générale en 200 dpi.

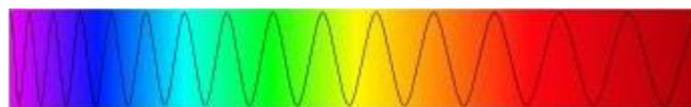
- Imprimantes grand public: entre 360 dpi et 1400 dpi. Cela permet d'obtenir une qualité tout à fait honorable pour tous les travaux courants (courriers, rapports, etc...)

- Scanners grand public: 300, 600 ou 1200 dpi.

- Matériel d'impression professionnel: aux minimum 4800 dpi (impression de qualité et grandes tailles pour les affiches).

4- La couleur

Les couleurs trouvent leur origine dans la séparation de la lumière blanche naturelle en composantes absorbées et composantes réfléchies. Toute source lumineuse visible est composée d'un mélange d'ondes électromagnétiques cohérentes (i.e. couleurs pures), dont la longueur d'onde est comprise entre 0,4 μm (violet) et 0,7 μm (rouge) :



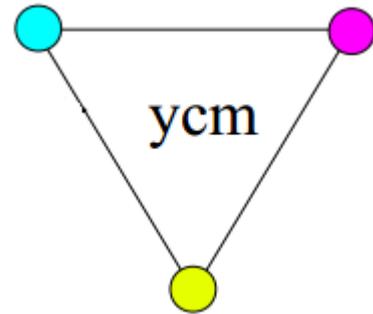
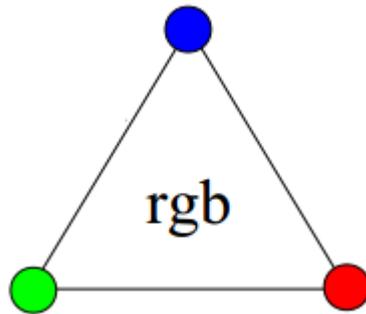
Spectre de la lumière blanche

La question principale que nous allons considérer dans la suite est : « Comment représenter ces couleurs dans un espace qui soit à la fois facile à manipuler et qui soit pertinent en termes d'analyse d'images couleur ? »

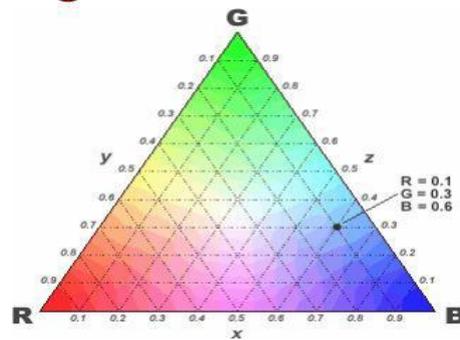
5.1 La Trichromie

Le principe de la trichromie réside dans la découverte qu'il suffit d'un triplet de couleurs pures pour restituer, par mélange de ces 3 composantes que nous pouvons qualifier de primaires, l'intégralité des couleurs.

Le cercle de Newton peut ainsi se « résumer » en un triangle de primaires.



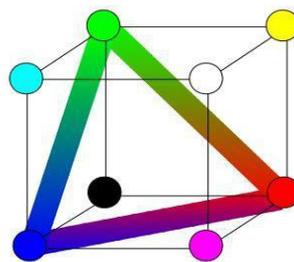
En remplissant le triangle de primaires selon la loi des aires, nous obtenons le triangle de Maxwell :



La trichromie permet donc de voir l'ensemble des couleurs comme un espace vectoriel de dimension 3, où 3 sources lumineuses primaires (r,g,b) varient en intensité entre 0 et 100%. Ce principe de synthèse additive de la couleur se retrouve dans la plupart des dispositifs lumineux de restitution de la couleur : CRT, LCD, Plasma.

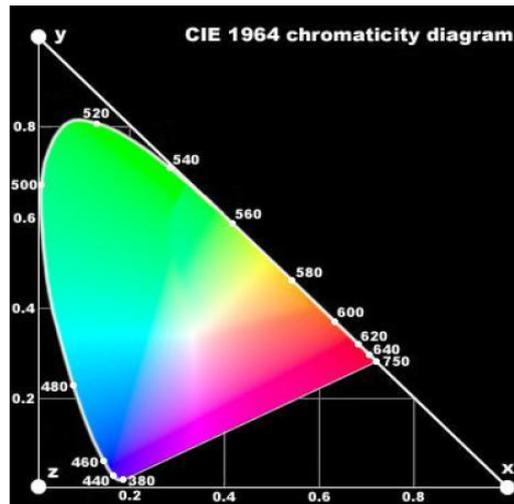
- **L'espace RGB**

L'espace RGB est l'espace vectoriel engendré par les 3 composantes primaires (Rouge, Vert, Bleu). L'ensemble des couleurs produites se représente comme l'intérieur d'un cube tel qu'il est mentionné sur le triangle chromatique suivant :



- **L'espace XYZ**

L'espace XYZ (CIE 1931) est défini à partir d'une transformation linéaire sur l'espace RGB telle que toutes les couleurs du spectre visible soient contenues dans le triangle xyz. Les coordonnées XYZ des couleurs naturelles ont ainsi toujours des valeurs positives.



- **L'espace HLS (HSV)**

La perception humaine de la couleur est une réaction subjective qui peut être caractérisée en termes de luminosité, de teinte et de saturation.

- La luminosité correspond à une sensation traduite par des vocables comme clair, foncé, lumineux, sombre...

- La teinte ou tonalité chromatique correspond aux dénominations telles que rouge, vert, bleu, ... Elle correspond à la longueur d'onde dominante d'une couleur. Le blanc, le noir ou les gris sont dites neutres ou achromatiques.

-La saturation est une grandeur estimant le niveau de coloration d'une teinte indépendamment de la luminosité. Elle représente la pureté de la couleur perçue comme vive, pâle.

Le passage de RGB à HSV se fait par une transformation non linéaire. Plusieurs opérateurs ont été proposés pour la conversion.

5- Codage des images

Un pixel est codé suivant la qualité de l'image :

5.1 L'image binaire

Dans une image en noir et blanc (image binaire), un seul bit suffit pour coder le point (0 pour noir, 1 pour blanc) comme mentionné sur la figure suivante :

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

Figure I.4 : Image binaire

5.2 L'image en 256 nuances de gris

Dans une image en 256 nuances de gris, chaque pixel est représenté par un octet (8 bits). Nous ne code ici que le niveau de l'intensité lumineuse, généralement sur un octet ($2^8 = 256$ valeurs). Par convention, la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale) :

Nous présentons sur la figure I.5, les niveaux de gris relatifs aux pixels d'une partie d'une image.

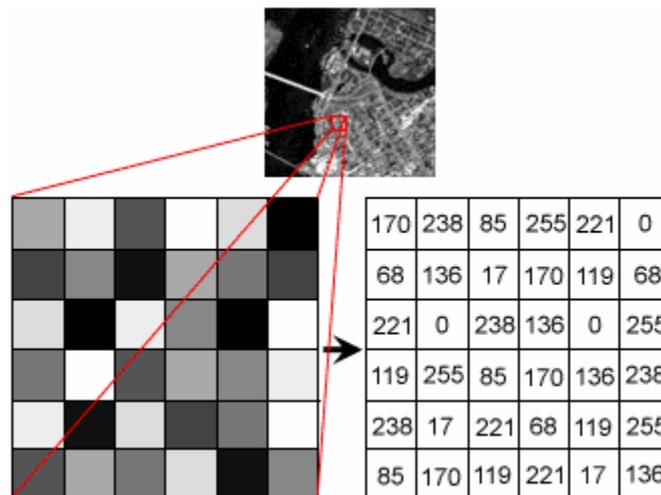


Figure I.5 : Niveaux de gris relatifs aux pixels d'une partie d'une image

Donc, l'image numérique se représente sous la forme d'un tableau à deux dimensions contenant des valeurs entières pour les images en niveaux de gris ou des triplets de valeurs entières pour les images couleurs.

C'est ce tableau de valeurs entières qui est le "code génétique" de l'image.

Quant à la figure ci-dessous, nous présentons un exemple de photo en 8 bpp.

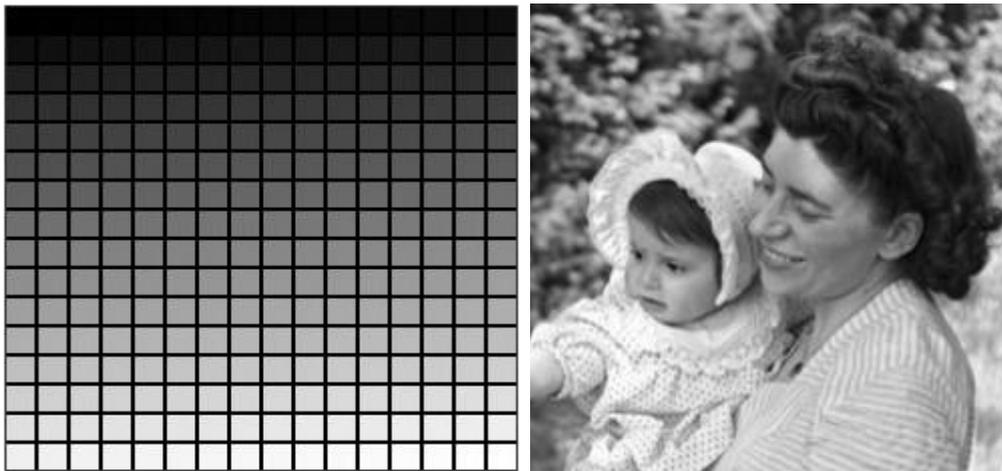


Figure I.6 : Nuances de 256 gris (à gauche), exemple de photo possible en 8 bpp (à droite)

Il est utile d’avoir une idée sur les 256 nuances de gris comme mentionné sur la figure I.7.

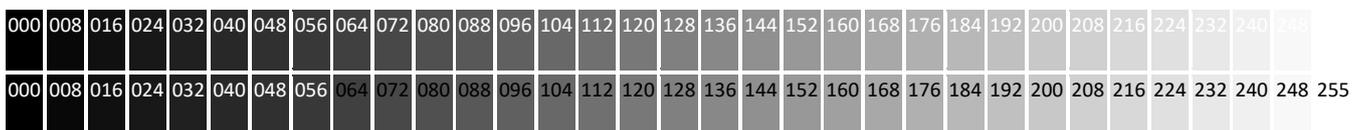


Figure I.7 : Nuances de 256 gris

Ce procédé est fréquemment utilisé pour reproduire des photos en noir et blanc ou du texte.

5.3 L’image couleur

Dans une image couleur, nous montrons que la couleur peut être exprimée comme une combinaison linéaire de trois couleurs de base, par exemple Rouge(R), Vert(V), Bleu(B). Ainsi une couleur quelconque x est exprimée comme $x = aR + bV + cB$. Où a, b, c sont des doses de couleurs de base. Chacun de ces éléments dispose de nuances allant de 0 à 255 : 256 couleurs. Pour avoir 256 couleurs, il faut 8 bits, donc 1 octet. Comme il y a 3 éléments différents RVB, il nous faut donc 3 octets pour rendre bien compte de toutes les nuances.

Le codage RVB (ou RGB en anglais) est celui utilisé dans de nombreux périphériques numériques (Scanners, appareils photos, écrans...). Le principe est simple, un octet sert à coder un nombre compris entre 0 et 255 (d’où 256 valeurs) qui correspond à la valeur de la composante rouge, verte ou bleue du pixel (0 étant l’absence de cette composante, 255 la saturation). Le mélange de ces trois composantes donne la couleur finale.

Les 24 bits d’une couleur se décomposent en 3 fois 8 bits, c’est-à-dire 3 octets :

- 8 bits sont consacrés à la teinte primaire rouge ;
- 8 bits sont consacrés à la teinte primaire vert ;
- 8 bits sont consacrés à la teinte primaire bleue.

Usuellement, une bonne image correspond à des doses allant de 0 à 255. Par suite une image couleur de ce type peut être représentée par 3 matrices (une par couleur de base) dont chacune d’elle possède des éléments sur 8 bits, ce qui au total fait 24 bits par pixel. Nous nous rendons vite compte du volume atteint pour des images importantes et de bonne définition. Une image 640x480 en couleur (24 bits)

occupe un volume de 921 600 octets. Nous sommes alors amenés à utiliser des techniques de compression pour réduire la taille des fichiers d'images.

L'une des premières normes d'affichage, la norme CGA, permettait d'afficher des pixels dans une palette de 4 couleurs différentes. Pour cela, il fallait que chaque pixel soit codé par 2 bits.

Par exemple :

Numéro (décimal)	Numéro (binaire)	Couleur
0	00	Noir
1	01	Vert
2	10	Rouge
3	11	Jaune

Ensuite, le nombre de couleurs que nous pouvions représenter à la fois a continué à augmenter. Nous pouvons à présent avoir des images de 16 couleurs. Et puisque $16 = 2^4$, il suffit pour cela de coder chaque pixel par une suite de 4 bits. Voici un exemple de palette courante :

N° décimal	N° binaire	Couleur	N° décimal	N° binaire	Couleur
0	0000	Noir	8	1000	Gris foncé
1	0001	Bleu	9	1001	Bleu clair
2	0010	Vert	10	1010	Vert clair
3	0011	Cyan	11	1011	Cyan clair
4	0100	Rouge	12	1100	Rouge clair
5	0101	Magenta	13	1101	Magenta clair
6	0110	Marron	14	1110	Jaune
7	0111	Gris	15	1111	Blanc

Nous pouvons continuer ainsi pour découvrir des images de plus en plus nette grâce à leur grand nombre de couleurs : avec 8 bits par pixels, nous obtenons 256 couleurs (car $2^8 = 256$) ; avec 16 bits, 65536 couleurs ; avec 24 bits, plus de 16 millions de couleurs ; avec 32 bits, plus de 4 milliards. À partir de 24 bits, nous parlons d'ailleurs d'affichage en vraies couleurs tant le réalisme est parfait. A cet effet :

- 1 bit : 2^1 possibilités : image noir et blanc
- 2 bits : 2^2 possibilités : 4 couleurs
- 4 bits : 2^4 possibilités : 16 couleurs
- 8 bits : 2^8 possibilités : 256 couleurs
- 24 bits : 2^{24} : 16 777 216 millions de couleurs

5.4 L'image couleur indexée

Permet d'obtenir jusqu'à 256 couleurs fixes, définies à l'avance dans une palette. Il n'utilise qu'une seule couche.

- Codage en 8 bits par pixel (bpp) : $2^8 = 256$ possibilités
- Chaque pixel peut avoir jusque 256 couleurs fixes possibles

Comme nous pouvons le constater sur la figure I.8, avec 256 couleurs, certains dégradés de cette image apparaîtront tramés, la qualité est proche de la photo mais il est possible de faire beaucoup mieux.



Figure 1.8 : Exemple d'une image couleur indexée

Nous pouvons constater que la manière la plus naturelle de représenter une image couleur consiste à utiliser trois matrices. Toutefois, pour de grandes images, ceci conduit à l'occupation d'un espace mémoire important. A cet effet, Matlab a proposé la représentation indexée vu qu'elle est plus économique pour la représentation des images couleur. En contrepartie, la représentation indexée ne permet de représenter qu'un nombre limité de couleurs.

Nous allons démontrer plus précisément le principe de cette représentation comme indiqué sur la figure 1.8, en effet, ces couleurs sont mémorisées dans une table de couleurs (colormap) qui est une matrice $n \times 3$ (où n est le nombre de couleurs). L'image est alors une matrice contenant des nombres entiers compris entre 1 et n , chaque entier jouant le rôle d'un index relatif à la table de couleurs.

Exemple :

L'élément représenté sur la figure I.9 contient l'entier 18, ce qui ramène ou pointe vers la ligne 18 de la table de couleurs pour connaître la couleur du pixel correspondant.

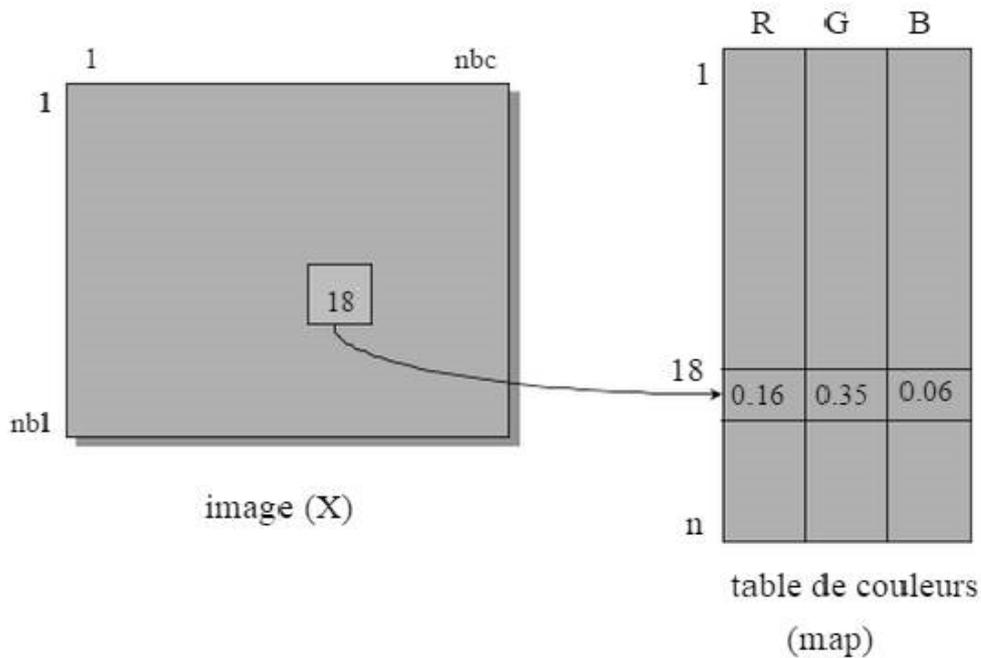


Figure I.9 : table de couleurs (colormap)

6- Représentation des images

6.1 Numérisation des images

La numérisation d'une image est obtenue par l'intermédiaire d'un capteur et d'un numériseur, qui transforment un signal optique en un signal numérique. Le signal optique peut être obtenu directement ou par l'intermédiaire d'un système d'agrandissement optique : lentille simple, loupe ou microscope photonique. Le capteur est constitué par un ensemble de capteurs élémentaires, une barrette de CCD (Charge Couple Device), ça peut être une caméra, noir et blanc ou couleur, un appareil photo numérique ou un scanner. Le signal électrique est repris par un convertisseur analogique-digital (carte de numérisation intégrée ou non au capteur) qui transforme les données continues en données numériques codées sur 1, 8, 16 ou 24 bits. Le codage utilisé définit le type d'images (noir et blanc, niveaux de gris ou couleur) et sa " profondeur ".

L'échantillonnage et la quantification sont deux opérations permettant de numériser une image :

- L'échantillonnage, visant à prélever régulièrement des échantillons sur le signal. Il est limité par la capacité du capteur, donc le nombre de pixels disponible.
- La quantification, visant à approcher le signal continu par un ensemble fini de valeurs. Elle est limitée par la quantité de tons (de gris) définie dans l'intervalle.

Nous présentons sur la figure I.10, une schématisation de l'échantillonnage ainsi que de la quantification.

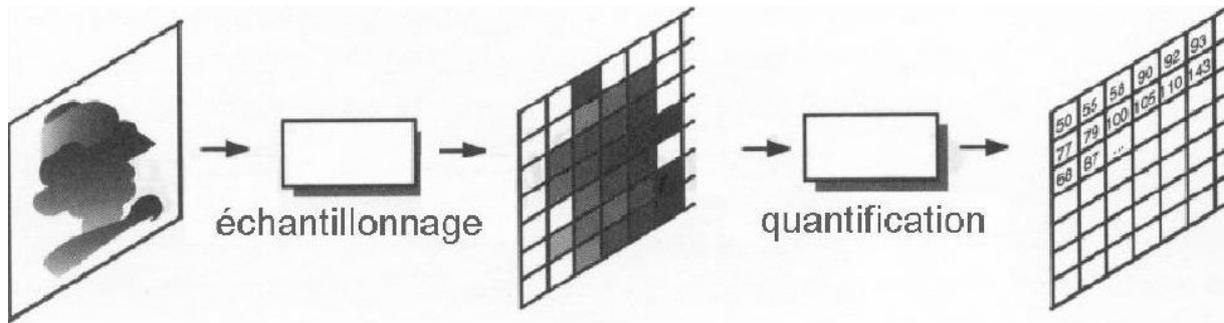


Figure I.10 : Schématisation des processus de l'échantillonnage et de la quantification

Les buts de la numérisation des images consistent en : les visualiser sur un moniteur, les imprimer, les traiter sur un ordinateur ; les stocker sur des supports informatiques ou les transmettre sur des réseaux informatiques.

De l'analogique au numérique :

Image **analogique** : image formée sur la rétine de l'œil, image obtenue par la photo argentique classique.

Image **numérique** : signal numérique composé d'unités élémentaires (pixels) représentant chacun une portion de l'image. Contrairement aux images obtenues à l'aide d'un appareil photo (analogique), ou dessinées sur du papier, les images manipulées par un ordinateur sont numériques (représentées par une série de bits).

6.2 Définition

L'image numérique : $f: [1, M] \times [1, N] \rightarrow I$, où MN est le nombre de pixels et I est un ensemble discret de valeurs.

- Cas où $I = \{0, 1\} \rightarrow$ c'est une image binaire.
- Cas où $I = \{0, \dots, P - 1\}$ (souvent $P = 2^8 = 256$: codage sur 8 bits) \rightarrow c'est une image en niveaux de gris.
- Cas où $I = \{0, \dots, P - 1\} \times \{0, \dots, P - 1\} \times \{0, \dots, P - 1\} \rightarrow$ c'est une image couleur.

Rouge

Vert

Bleu

6.3 L'échantillonnage

Comment passer d'un signal continu à un signal discret pouvant être enregistré, analysé et/ou traité ?

Rappel :

- Image continue 2D (ou perçue comme telle) :

$$D \in \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$$

Où \mathbb{C} est un espace de couleurs.

- Image discrète 2D :

$$D \in \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{C}$$

Où \mathbb{C} est un ensemble fini de couleurs

- Résolution verticale : nombre de lignes dans l'image.
- Résolution horizontale : nombre de colonne dans l'image.
- Résolution spatiale : résolution verticale x résolution horizontale.

- Densité de résolution : nombre de pixels par unité de longueur. S'exprime en ppi (pixels per inch) ou dpi (dots per inch).

La réponse technologique pour les images consiste à utiliser un appareil numérique (ce sont alors les capteurs CCD qui numérisent le signal) ou un scanner pour numériser des photos. La réponse théorique à cette question est la théorie de l'échantillonnage. En effet, l'échantillonnage est une étape fondamentale qui doit tenir compte du contenu informationnel pertinent de l'image à analyser.

But : transformer un signal analogique en un signal numérique. Pour ceci, des échantillons ponctuels sont prélevés régulièrement sur le signal, évidemment, tout le reste est perdu.

La fréquence à laquelle on prélève les échantillons (fréquence d'échantillonnage) doit être bien choisie:

- Assez grande pour pouvoir bien restituer le signal.
- Pas trop grande pour limiter l'espace de stockage nécessaire.

Lors de la numérisation, on cherche à conserver une qualité maximale (un nombre de pixels maximal, un nombre de couleurs maximal) et à obtenir des données les moins volumineuses possibles. Le problème est que ces deux besoins sont antagonistes. Il conviendra alors de chercher la résolution et le nombre des niveaux de couleurs satisfaisants les besoins. Par exemple, si le but est uniquement la visualisation sur un moniteur donné, il est inutile d'échantillonner l'image à une résolution supérieure à celle du moniteur.

Nous pouvons affirmer alors que l'échantillonnage désigne le procédé de discrétisation spatiale d'une image consistant à associer à chaque zone rectangulaire $R(x,y)$ d'une image continue une unique valeur $I(x,y)$ comme mentionné sur les figures I.11. Notons que la taille d'un pixel est liée à l'échantillonnage.

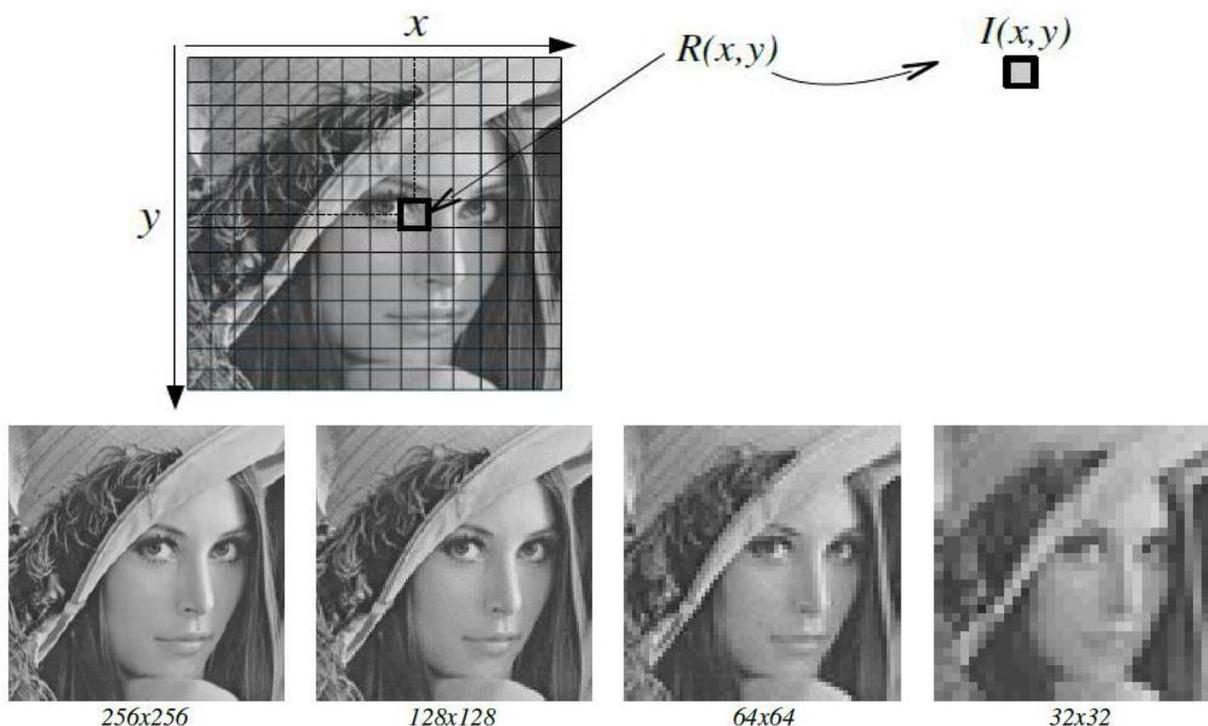


Figure I.11 : Échantillonnage d'une image

6.4 La quantification

La quantification désigne le choix du bon nombre de bits pour coder les images numériques comme nous pouvons le constater sur la figure I.12.

But : affecter une valeur numérique à chaque échantillon prélevé sur le signal (après échantillonnage) tout en réduisant le nombre de bits nécessaires au codage.

L'une dépend du capteur, et de sa capacité effective à observer des signaux de valeurs différentes : le rapport signal sur bruit.

Le rapport signal sur bruit est défini à partir du rapport entre l'amplitude des niveaux de gris mesurables par le capteur ($n_{\max} - n_{\min}$) et le niveau du bruit, en gros l'écart type s_n de la perturbation aléatoire qui affecte les niveaux de gris. En prenant le logarithme, on a le nombre de bits utile au capteur pour coder les images. Outre les capacités du capteur, le nombre de bits réellement nécessaires pour coder une image varie d'une image à l'autre, en fonction de leur contenu informationnel.

Ce nombre dépend de l'entropie, définie à partir de la distribution des niveaux de gris de l'image.

$$E = \sum_{i \leq N} -p_i \log_2(p_i)$$

Où N est le nombre de niveaux de gris présents, p_i est la proportion ($0 < p_i < 1$) de points de l'image ayant pour niveau de gris i . Cette grandeur représente le nombre moyen de bits par pixel nécessaires pour coder toute l'information présente. Elle est utilisée dans les techniques de compression sans perte pour adapter le volume de donnée des images à leur contenu informationnel.



Figure I.12 : Quantification d'une image

L'échantillonnage et la quantification ne sont pas limités à la numérisation d'une image. Ce sont des techniques que l'on utilise également sur des images déjà numérisées, afin de modifier la résolution (on ré échantillonne l'image) ou le nombre de couleurs utilisés (quantification). Cependant, il est dans ce cas fait appel à des techniques de reconstruction d'images, c'est-à-dire de retrouver un signal 2D continu en espace et/ou en couleur.

Nous présentons sur la figure ci-dessous l'échantillonnage et la quantification d'un signal 2d continu selon différentes résolutions spatiales et colorimétriques.

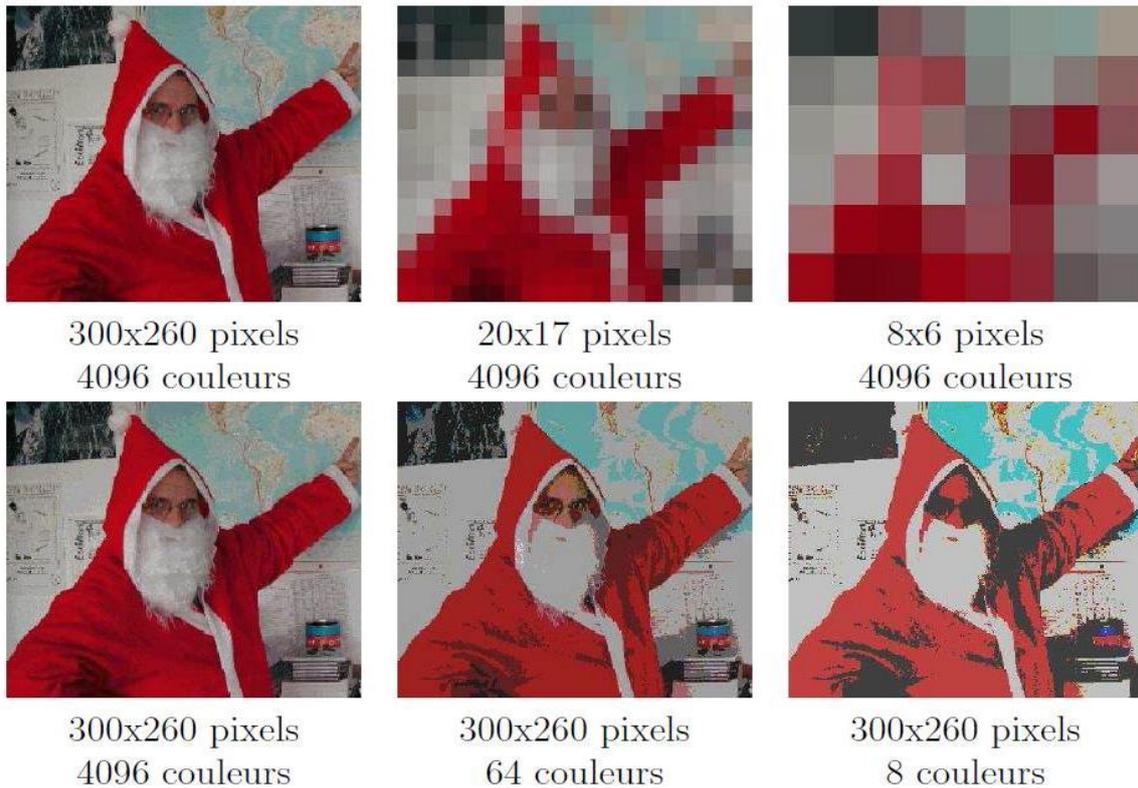


Figure I.13 : l'échantillonnage et la quantification d'un signal 2d continu selon différentes résolutions spatiales et colorimétriques

Donc, nous pouvons affirmer que la résolution spatiale est obtenue après échantillonnage (Sampling), déterminant le plus petit détail perceptible dans l'image, en effet, la taille des CCD et le pas d'échantillonnage de l'image au niveau de la carte de numérisation définissent la résolution spatiale de l'image. Le format de l'image est défini par l'entête du fichier et peut être précisé par son extension .bmp, .tif, .gif, .jpeg etc..., il renseigne sur le mode de présentation des données et leur degré de compression. Tandis que la résolution en niveaux de gris est obtenue après quantification déterminant le plus petit changement de niveau de gris discernable dans l'image.

6.5 Stockage des images

Pour adapter le tableau de valeurs entières désignant le code génétique de l'image aux supports et moyens de transmission informatiques, nous devons le retranscrire en binaire. Les infos qui vont être stockées sont la largeur et la hauteur ainsi que les valeurs des pixels. Nous pouvons stocker davantage d'informations telles que le type de données, l'auteur, la date, les conditions d'acquisition, ... etc.

Nous allons ainsi stocker les informations concernant l'image dans un en-tête puis les données. Les données sont souvent stockées dans l'ordre des pixels de gauche à droite et de haut en bas en remplaçant chaque valeur entière de niveau de gris par son code en binaire, i.e. sa valeur en base 2 comme mentionné sur la figure I.14.

La première ligne P1 signifie qu'il s'agit d'un format pbm et que les données seront stockées en ASCII. La deuxième ligne, commençant par le caractère # est une ligne de commentaire. La ligne suivante 640 480 précise les dimensions de l'image. Viennent ensuite les données, ligne par ligne et pour chaque ligne de gauche à droite.