

1

Introduction à l'image

1.1 Introduction

L'objectif de ce chapitre est d'introduire les notions de base dans le domaine de traitement d'image. La section 1.2 résume les mécanismes de système de vision humaine ainsi que la formation d'image dans l'œil et ses capacités pour l'adaptation et la discrimination de la luminosité. La section 1.3 discute la lumière, le spectre électromagnétique, et ses caractéristiques d'imagerie. La section 1.6 introduit les concepts d'échantillonnage et quantification ainsi que d'autres sujets supplémentaires tels que les concepts de la résolution spatiale et tonale. Finalement, la section 1.7 traite les différentes relations entre pixels.

1.2 Eléments de la perception visuelle

1.2.1 Anatomie et physiologie de l'œil humain

La Fig. 1.1 présente une coupe transversale horizontale simplifiée de l'œil humaine. L'œil est proche d'une sphère, avec un diamètre moyen approximatif de 20 mm. L'œil humain est constitué d'un globe oculaire comportant :

- sur sa partie antérieure, la cornée, qui est une calotte sphérique transparente ;
- sur le reste du globe, la sclère, ou sclérotique, qui forme le "blanc" de l'œil.

Le globe oculaire mesure environ 2,5 cm de diamètre et a une masse de 8 grammes. Il est formé de 3 enveloppes, ou tuniques, entourant une substance gélatineuse appelée le corps vitré. Les trois tuniques s'appellent la tunique externe, la tunique moyenne, et la tunique interne ; le corps vitré est principalement constitué d'eau et sert à maintenir la forme de l'œil.

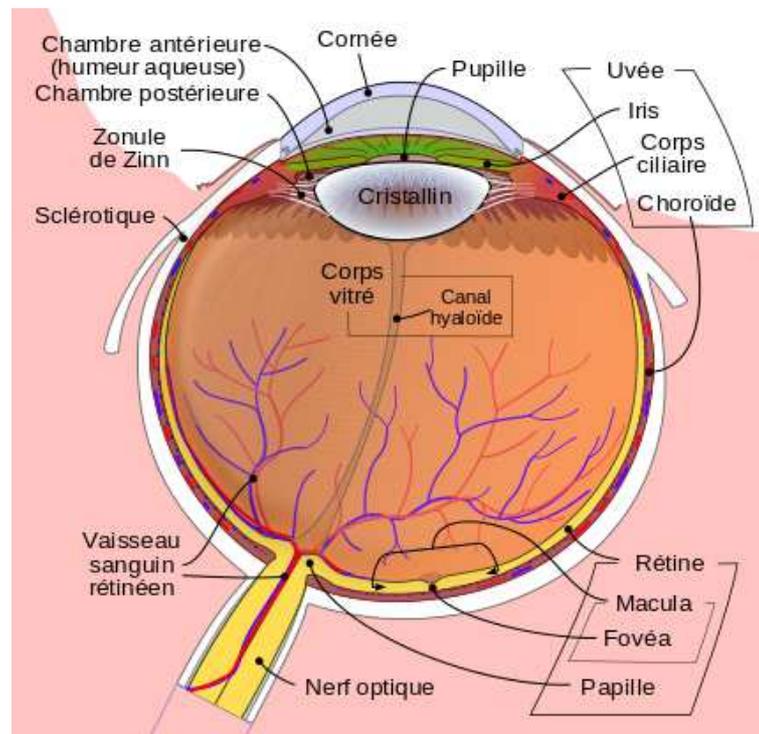


FIG. 1.1: Graphique simplifié d'une coupe transversale de l'œil humaine.

L'œil s'adapte en premier lieu à la lumière ambiante. L'être humain peut ainsi percevoir avec une sensibilité équivalente en plein soleil ou sous la lumière de la pleine lune, soit avec une intensité lumineuse 10 000 fois moindre. Une première adaptation provient de l'écartement de l'iris qui, en mode nocturne, peut atteindre une ouverture maximale de 7 mm pour des jeunes gens (maximum qui décroît à 4 mm avec l'âge).

1.2.1.1 Tunique externe

- La sclérotique est la plus résistante des tuniques de l'œil ; elle le protège des dégâts mécaniques et soutient sa structure ; elle est percée en avant d'un orifice pour la cornée.
- La conjonctive est une muqueuse transparente qui recouvre la partie antérieure de la sclérotique et qui produit un mucus lubrifiant.
- La cornée est une membrane transparente circulaire et bombée vers l'avant qui permet le passage des rayons lumineux, située au centre de la partie antérieure de la tunique de l'œil ; elle est en continuité avec la sclérotique et la conjonctive autour d'elle au niveau du limbe cornéen.

1.2.1.2 Tunique moyenne

- La choroïde est une membrane vascularisée qui assure la nutrition de la rétine. Les cellules de cette tunique renferment un pigment, la mélanine, qui lui donne une couleur

brun-foncé, afin que les rayons ne pénètrent que par la pupille. La choroïde forme en avant l'iris.

- L'iris donne la couleur à l'œil. Il est percé en son centre par une ouverture circulaire, la pupille, qui se dilate ou se contracte selon l'intensité de la lumière, grâce à l'action des muscles lisses de l'iris.
- La pupille laisse passer la lumière. Le diamètre de son ouverture s'adapte automatiquement à l'intensité lumineuse perçue.
- Le corps ciliaire sécrète l'humeur aqueuse. Il contient un réseau de muscles qui permettent de modifier la courbure du cristallin afin de rendre la vision nette.
- Le cristallin est un petit disque fibreux, transparent et flexible qui permet de focaliser l'image sur la rétine en fonction de la distance.

1.2.1.3 Tunique interne

- La rétine est la tunique sensible de l'œil. Elle est formée de cellules sensorielles, les cônes (vision diurne=jour) et les bâtonnets (vision nocturne=nuit), et de cellules nerveuses, les neurones.
- La macula sorte de tache jaune, assure une motilité visuelle maximale, car elle est formée de nombreuses cellules visuelles.
- La tache aveugle, ou papille, est la zone où les fibres se réunissent pour former le nerf optique, qui ne renferme aucune cellule photosensible.
- La fovéa est une petite zone de la rétine sensible aux couleurs et sert pour la précision de la vision.
- Le nerf optique est formé du regroupement des fibres nerveuses de la rétine et conduit l'information visuelle au cerveau.

1.2.1.4 Rétine

La rétine possède deux types de photorécepteurs :

- Les bâtonnets (~ 120 millions)
 - Distribués en périphérie de la rétine.
 - Captent la luminosité et incapables de dissocier des couleurs.
 - Sensibles aux faibles intensités lumineuses.
 - Associés à la détection de mouvement dans le cortex.
- Les cônes (~ 5 millions)

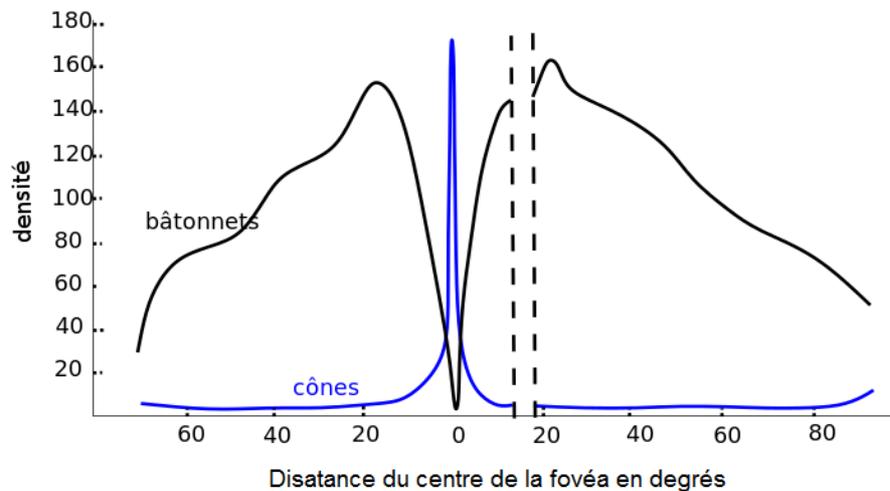


FIG. 1.2: Distribution des bâtonnets et des cônes dans la rétine.

- 3 types de cônes,
- Ceux sensibles aux basses fréquences (les rouges).
- Ceux sensibles aux fréquences moyennes (les verts).
- Ceux sensibles aux hautes fréquences (les bleus).
- Insensibles aux faibles intensités.
- Distribués près de la fovéa, là où la vision des détails est la meilleure.

La Fig. 1.2 montre la densité des bâtonnets et des cônes pour une coupe transversale de l'œil en passant par l'endroit où le nerf optique quitte l'œil. L'absence de photorécepteurs dans cette zone entraîne la présence d'un *point aveugle* (voir la Fig. 1.1). La distribution des photorécepteurs est symétrique par rapport à la fovéa. La densité des photorécepteurs est mesurée en degrés à partir de la fovéa (c'est-à-dire en degrés hors axe, mesurée par l'angle formé par l'axe visuel et une ligne passant par le centre de la lentille et coupant la rétine). Noter la grande densité des cônes au centre de la rétine (dans la zone centrale de la fovéa). Noter également que la densité des bâtonnets augmentent au centre vers environ 20° hors axe, et elle se diminue jusqu'à la périphérie extrême de la rétine.

1.2.2 Formation d'image dans l'œil

Dans une caméra ordinaire, la lentille a une longueur focale fixe, et la focalisation sur les différentes distances est achevée par la variation de la distance entre la lentille et le plan d'image. Dans l'œil humaine, c'est l'inverse qui est vrai; la distance entre la lentille et le plan d'image (la rétine) est fixe, et la longueur focale demandée pour avoir un focus est obtenue par la variation de la forme de lentille.

La géométrie dans la Fig. 1.3 montre comment obtenir les dimensions d'une image formée sur la rétine. Par exemple, supposons qu'une personne regarde un arbre de hauteur de 15m

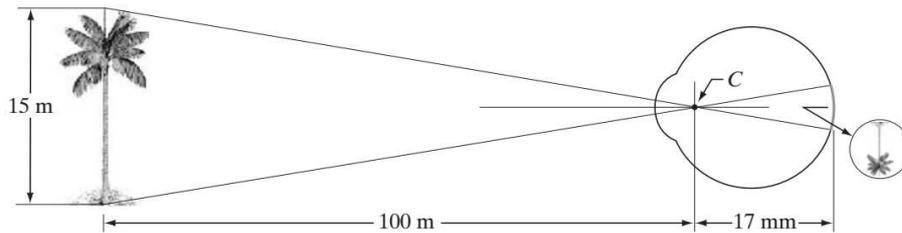


FIG. 1.3: Représentation graphique de l'œil regardant un palmier. Le point C est le centre optique de l'objectif.

à partir d'une distance de $100m$. Soit h la hauteur de cet objet dans l'image rétinienne, la géométrie de la Fig. 1.3 donne $15/100 = h/17$ d'où $h = 2.55mm$. L'image rétinienne est axée sur la région de la fovéa. La perception se fait par l'excitation relative des photorécepteurs de lumière, qui transforment l'énergie de la lumière en impulsions électriques qui vont être finalement décodées par le cerveau.

1.3 Lumière et spectre électromagnétique

En 1666, Sir Isaac Newton a découvert que lorsqu'un rayon de lumière solaire passe par un prisme en verre, le rayon de lumière n'est pas blanc mais consiste à un spectre continu de couleurs allant du violet au rouge. La Fig. 1.4 montre que le rayon électromagnétique accessible à la perception humaine visuelle est la très étroite portion du spectre électromagnétique.

La première extrémité du spectre représente les ondes radio avec des longueurs d'ondes des milliards de fois plus longues que celles de la lumière visible. À l'autre extrémité du spectre, les rayons gamma ont des longueurs d'ondes, des millions de fois inférieures à celles du spectre visible. Le spectre électromagnétique peut être exprimé en termes de longueur d'onde, de fréquence ou d'énergie. La longueur d'onde (λ) et la fréquence (ν) sont liées par l'expression suivante:

$$\lambda = \frac{c}{\nu} \quad (1.1)$$

Où c est la vitesse de la lumière ($2.998 \times 10^8 m/s$). L'énergie des différentes composantes du spectre électromagnétique est donnée par l'expression:

$$E = h\nu \quad (1.2)$$

où h est la constante de Planck. Les unités de longueur d'onde sont les *mètres*, les *micromètres* (noté $\mu m = 10^{-6}m$) et les *nanomètres* (noté $nm = 10^{-9}m$). La fréquence est mesurée en Hertz (Hz), où 1 Hertz est égal à une période d'une onde sinusoïdale par seconde. L'unité utilisée généralement pour mesurer l'énergie est appelée *electron-volt*.

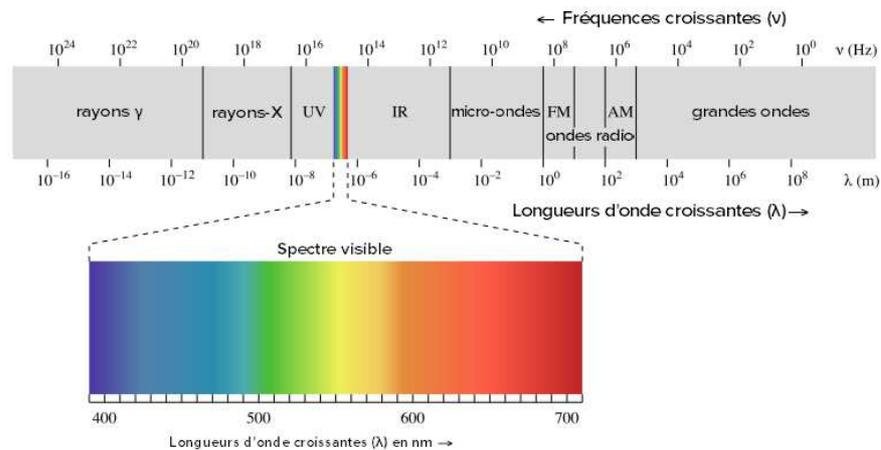


FIG. 1.4: Spectre des fréquences de couleurs visibles.

1.4 Acquisition d'image à l'aide des matrices de capteurs

En général, dans les appareils photo numériques les capteurs sont arrangés sous forme de matrice. Un capteur typique pour ces caméras est le capteur CCD, qui peut être fabriqué avec une large gamme de propriétés de détection et peut être emballé dans des matrices de 4000×4000 éléments ou plus. Les capteurs CCD sont largement utilisés dans les appareils photo numériques ainsi que dans d'autres instruments de détection de lumière. La réponse de chaque capteur est proportionnelle à l'intégrale de l'énergie lumineuse projetée sur la surface du capteur, cette propriété est utilisée dans les applications astronomiques ou toute autre application nécessitant un prétraitement de réduction de bruit. La réduction du bruit est obtenue en laissant le capteur intégrer le signal lumineux en entrée sur des minutes ou même des heures.

La façon principale pour utiliser les matrices de capteurs est illustrée à la Fig. 1.5. Cette figure montre l'énergie provenant d'une source d'illumination réfléchie à partir d'un élément de la scène. La première fonction réalisée par le système d'imagerie dans la Fig. 1.5(c) consiste à collecter l'énergie entrante et de la focaliser sur un plan d'image. L'avant du système est une lentille optique qui projette la scène vue sur le plan focal de la lentille, tel que le montre la Fig. 1.5(d). La matrice de capteurs, qui coïncide avec le plan focal, produit des sorties proportionnelles à l'intégrale de la lumière reçue à chaque capteur. Les circuits numériques et analogiques balayent ces sorties et les convertissent en un signal analogique, qui est ensuite numérisé par une autre composante du système d'imagerie. La sortie est une image numérique, comme le montre schématiquement la Fig. 1.5(e).

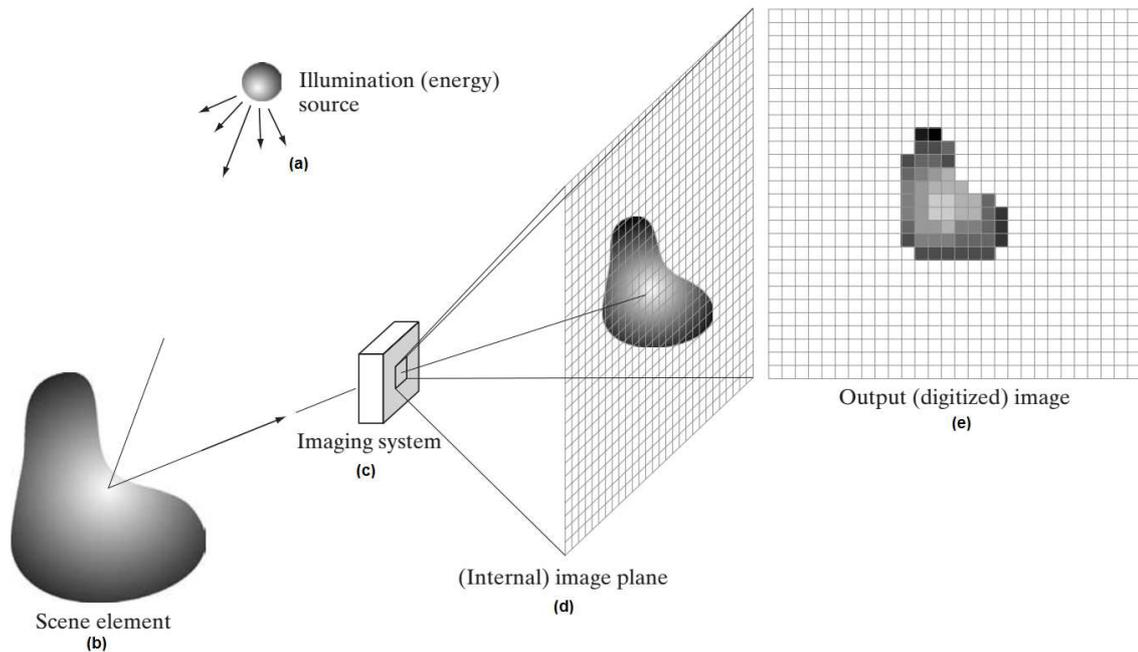


FIG. 1.5: Un exemple du processus d'acquisition d'image numérique. (a) Source d'énergie ("illumination"). (b) Un élément d'une scène. (c) Système d'imagerie. (d) Projection de la scène sur le plan de l'image. (e) Image numérisée.

1.5 Un Modèle simple de formation d'image

Les images sont définies par les fonctions bidimensionnelles de la forme $f(x, y)$. La valeur ou l'amplitude de f aux coordonnées spatiales (x, y) est une quantité scalaire positive dont la signification physique est déterminée par la source de l'image. Lorsqu'une image est générée à partir d'un processus physique, ses valeurs d'intensité sont proportionnelles à l'énergie générée par la source physique (par exemple, des ondes électromagnétiques). Par conséquent, la fonction $f(x, y)$ doit être non nulle et finie, c-à-d:

$$0 < f(x, y) < \infty \quad (1.3)$$

La fonction $f(x, y)$ peut être caractérisée par deux composantes: (1) *l'illumination* notée $i(x, y)$ et (2) *la réflectance* notée $r(x, y)$. Les deux fonctions $i(x, y)$ et $r(x, y)$ se combinent pour former la fonction $f(x, y)$, comme suit:

$$f(x, y) = i(x, y)r(x, y) \quad (1.4)$$

où

$$0 < i(x, y) < \infty \quad (1.5)$$

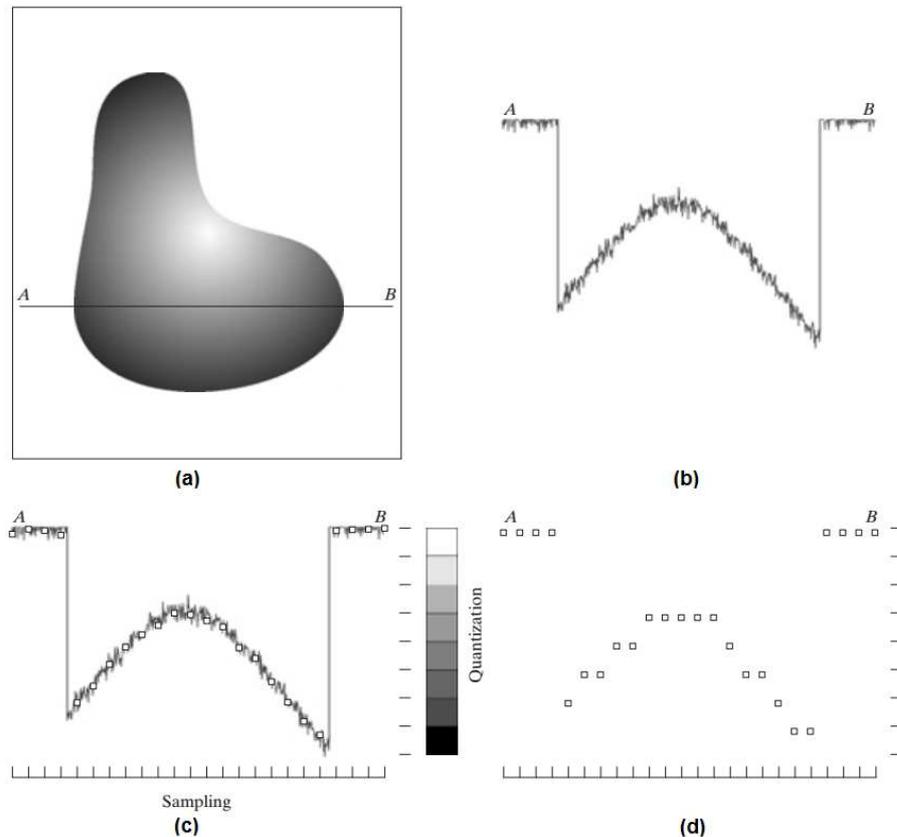


FIG. 1.6: Génération d'une image numérique. (a) Image continue. (b) Un profil d'intensité de A à B dans l'image continue, utilisé pour illustrer les concepts d'échantillonnage et de quantification. (c) L'échantillonnage et la quantification. (d) Un profil d'intensité numérique.

et

$$0 < r(x, y) < 1 \quad (1.6)$$

1.6 Échantillonnage et quantification d'image

Pour créer une image numérique, on doit convertir les données capturées sous forme continue (signal analogique) vers une forme numérique. Cela implique deux processus : *l'échantillonnage* et *la quantification*.

1.6.1 Concepts de base d'échantillonnage et quantification

L'idée de base derrière l'échantillonnage et la quantification est illustrée sur la Fig. 1.6. La Fig. 1.6(a) montre une image continue, $f(x, y)$, que nous voulons convertir en forme numérique. Une image peut être continue par rapport à l'amplitude et les coordonnées

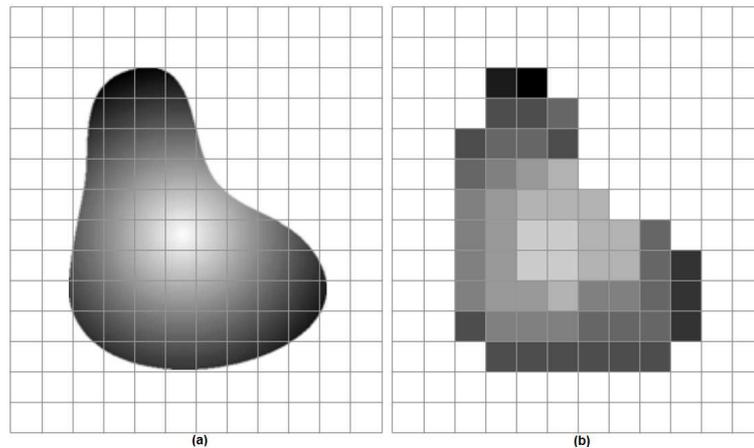


FIG. 1.7: (a) Image continue projetée sur un réseau de capteurs. (b) Résultat de l'échantillonnage et la quantification d'image.

x, y . Pour la convertir en forme numérique, on doit l'échantillonner dans les coordonnées et l'amplitude. La numérisation des valeurs des coordonnées est appelée *échantillonnage* et la numérisation des valeurs d'amplitude est appelée *quantification*.

Lorsqu'une matrice de capteurs est utilisée pour l'acquisition de l'image, le nombre de capteurs dans la matrice limite l'échantillonnage dans les deux directions x et y . La Fig. 1.7 illustre ce concept. La Fig. 1.7(a) montre une image continue projetée sur le plan d'une matrice de capteurs. La Fig. 1.7(b) montre le résultat de l'image après l'échantillonnage et la quantification. Il est évident que la qualité d'une image numérique est fortement déterminée par le nombre d'échantillons et les niveaux de gris discrets utilisés lors de l'échantillonnage et la quantification.

1.6.2 Représentation des images numériques

Les vecteurs numériques sont utilisés pour le développement d'algorithmes de traitement d'image. Mathématiquement, une image numérique de taille $M \times N$ est représentée sous la forme matricielle suivante :

$$f(x, y) = \begin{bmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ f(M,1) & f(M,2) & \dots & f(M,N) \end{bmatrix} \quad (1.7)$$

Le côté droit de l'Eq. (1.7) est par définition une image numérique. Chaque élément de cette matrice est appelé *élément d'image* ou *pixel*. Les termes *image* et *pixel* seront utilisés tout au long de ce polycopié pour désigner une image numérique et ses éléments.

La seule condition à vérifier par les nombres M et N est d'être des entiers positifs. Alors que, pour le nombre de niveaux de gris L , il est généralement choisi comme puissance

N/k	1(L=2)	2(L=4)	3(L=8)	4(L=16)	5(L=32)	6(L=64)	7(L=128)	8(L=256)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

TABLE 1.1: Nombre de bits de stockage pour différentes valeurs de N et k .

entière de 2. Ceci est dû à plusieurs considérations tel que le temps de calcul, le stockage et l'échantillonnage,

$$L = 2^k \quad (1.8)$$

On suppose que les niveaux de gris sont des entiers dans l'intervalle $[0, L-1]$. L'intervalle des valeurs des niveaux de gris est appelé *dynamique de l'image*. Une image dont les valeurs des niveaux des gris couvrent une partie importante de l'intervalle précédent est dite ayant une *grande dynamique*. Lorsqu'un grand nombre de pixels vérifient cette propriété, l'image aura un *contraste élevé*. Par contre, une image à *faible dynamique* aura une tendance à avoir un aspect gris terne.

Le nombre, b , de bits requis pour stocker une image numérisée est

$$b = M \times N \times k \quad (1.9)$$

Lorsque $M = N$, cette équation devient

$$b = N^2 k. \quad (1.10)$$

Le Tab. 1.1 montre le nombre de bits requis pour stocker des images carrées avec différentes valeurs de N et k . Le nombre de niveaux de gris correspondant à chaque valeur de k est indiqué entre parenthèses. Lorsqu'une image peut avoir des niveaux de gris de 2^k , On dit que l'image est une "image k -bits". Par exemple, une image avec 256 valeurs de niveaux de gris possibles est appelée image 8-bits.

1.6.3 Résolution spatiale et tonale

L'échantillonnage est le principal facteur déterminant la *résolution spatiale* d'une image. En principe, la résolution spatiale est le plus petit détail discernable dans une image. Cette résolution spatiale qui indique la "densité de pixels" est couramment simplement nommée "résolution". La résolution spatiale est le nombre de pixels sur une distance exprimée le plus souvent en pouces (1 pouce = 2.54 cm). Par exemple, la résolution d'une image numérique



FIG. 1.8: Une image 1024×1024 , 8 bits sous-échantillonnée jusqu'à 32×32 pixels. Le nombre de niveaux de gris a été maintenu à 256.

s'exprime en "ppi" (pixels per inch). Alors que, la résolution d'impression d'une imprimante ou de capture d'un scanner s'exprime en "DPI" (dots per inch).

De la même façon, la *résolution tonale* (de tons de gris) peut être définie par le plus petit changement discernable. Le nombre de niveaux de gris est généralement une puissance entière de 2, comme on a déjà mentionné dans la section précédente. Le nombre le plus courant est 8 bits, 16 bits est utilisé aussi dans certaines applications où l'augmentation du nombre des niveaux de gris est nécessaire.

1.6.3.1 Exemple: Effets typiques de la variation du nombre d'échantillons dans une image numérique

La Fig. 1.8 montre une image de taille 1024×1024 pixels dont les niveaux de gris sont représentés par 8 bits. Les autres images de cette figure représentent les résultats du sous-échantillonnage de l'image 1024×1024 . Le sous-échantillonnage est réalisé en supprimant un ensemble de lignes et de colonnes de l'image originale. Par exemple, l'image 512×512 a été obtenue en supprimant 512 lignes et colonnes de l'image 1024×1024 , et ainsi de suite. Le nombre de niveaux de gris est maintenu à 256.

Ces images montrent les proportions entre les différentes densités d'échantillonnage. A fin de bien comparer les effets de sous-échantillonnage, les images sous-échantillonnées sont amenées jusqu'à la taille de 1024×1024 . Les résultats sont montrés sur les Figs. 1.9(b) à (f). La Fig. 1.9(a) est la même image de 1024×1024 , 256 niveaux montrée sur la Fig. 1.8.

Noter qu'il est pratiquement impossible de distinguer les deux images de la Fig. 1.9(a) et celle de la Fig. 1.9(b). Le niveau de détail perdu est tout simplement trop fin pour être vu sur la page imprimée. Ensuite, l'image 256×256 de la Fig. 1.9(c) montre un très léger motif de damier dans les bordures entre les pétales de la fleur et le fond noir. Egalement, Un grain commence légèrement à apparaître. Ces effets sont beaucoup plus visibles dans

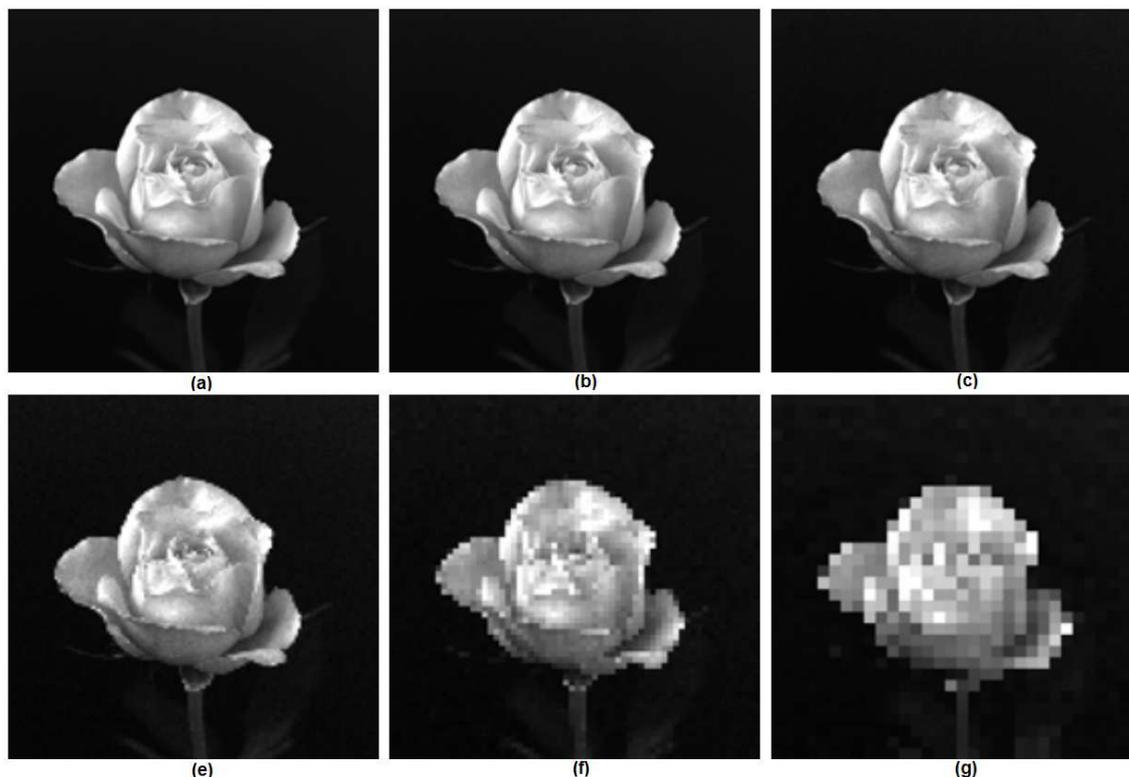


FIG. 1.9: (a) 1024×1024 , image 8 bits. (b) 512×512 image rééchantillonnée en 1024×1024 pixels par duplication de ligne et de colonne. (c) à (f) 256×256 , 128×128 , 64×64 et 32×32 images rééchantillonnées en 1024×1024 pixels.

l'image 128×128 de la Fig. 1.9(d), et ils se manifestent très bien dans les images 64×64 et 32×32 des Fig. 1.9(e) et (f), respectivement.

1.6.3.2 Exemple: Effets typiques de la variation du nombre de niveaux de gris dans une image numérique

Dans cet exemple, le nombre d'échantillons est fixé, alors que le nombre de niveaux de gris est réduit de 256 à 2, avec des puissances entières de 2. La Fig. 1.10(a) est une image radiographique de taille de 452×374 , affichée avec $k = 8$ ($L = 256$ niveaux de gris).

Les Figs. 1.10(b) à (h) ont été obtenues en réduisant le nombre de bits de $k = 7$ à $k = 1$ tout en maintenant la résolution spatiale constante à 452×374 pixels. Les images à 256, 128 et 64 niveaux sont visuellement identiques, alors que l'image à 32 niveaux montrée sur la Fig. 1.10(d) contient un ensemble de zones de niveaux de gris lisses (en particulier dans le crâne). Cet effet, causé par l'utilisation d'un nombre insuffisant de niveaux de gris dans les zones lisses d'une image, est appelé les *faux contours*. Les faux contours sont généralement bien visibles dans les images affichées en utilisant 16 ou moins de niveaux de gris, tel que montré sur les images 1.10(e) à (h).

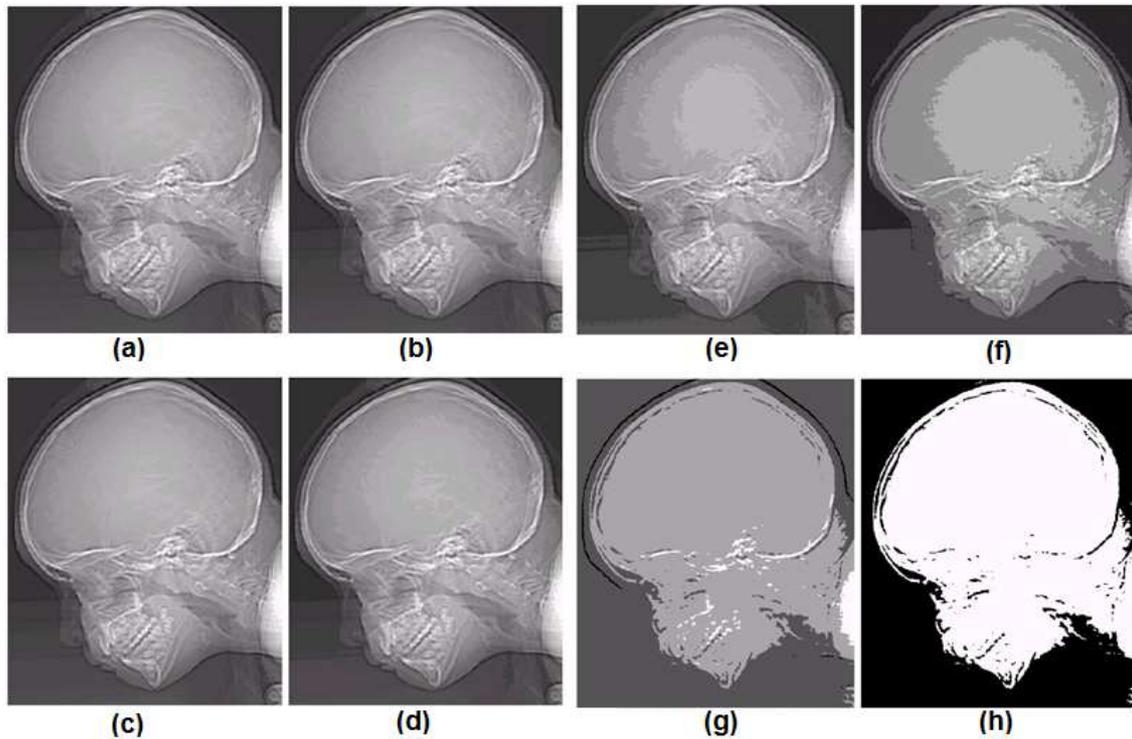


FIG. 1.10: (a) 452×374 , image de 256 niveaux. (b) - (h) Image affichée en 128, 64, 32, 16, 8, 4 et 2 niveaux de gris, tout en maintenant la résolution spatiale constante.

1.7 Quelques relations de base entre pixels

Dans cette section, nous allons étudier les relations basiques entre les pixels d'une image numérique. Une image est notée $f(x, y)$ et les lettres minuscules p et q désignent un pixel particulier.

1.7.1 Voisinage d'un pixel

Un pixel p à la position (x, y) a quatre voisins horizontaux et verticaux dont les coordonnées sont:

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1)$$

Cet ensemble de pixels, noté $N_4(p)$, est appelé le *4-voisinage* de p . La distance entre chaque pixel voisin et le pixel central est égale à une unité. Si le point (x, y) est sur la bordure de l'image, certains voisins de p se situent à l'extérieur de l'image.

Les quatre voisins diagonaux de p , notés $N_D(p)$, sont définis par les coordonnées:

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1)$$

Ces points, en plus des points du *4-voisinage*, constituent le *8-voisinage* de p noté $N_8(p)$.

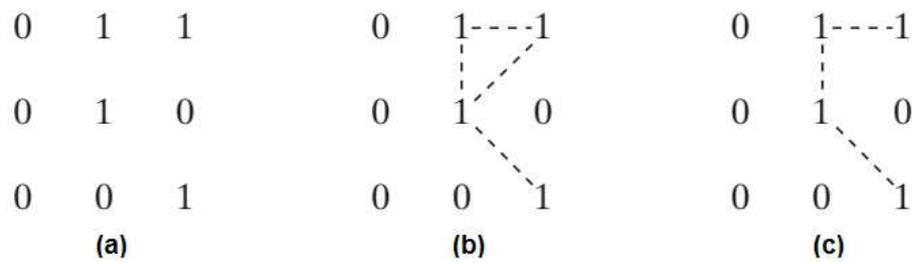


FIG. 1.11: (a) Disposition de pixels; (b) des pixels qui sont 8-adjacents (montrés en tirets) au pixel central; (c) m-adjacence.

1.7.2 Adjacence, connectivité, régions, et frontière

Soit V l'ensemble des valeurs d'intensité utilisé pour définir l'adjacence. Dans une image binaire, $V = \{1\}$ s'il s'agit de l'adjacence des pixels avec la valeur 1. Dans une image en niveaux de gris, l'ensemble V contient plus d'éléments. Par exemple, dans l'intervalle $[0, 255]$, l'ensemble V peut être un sous-ensemble de ces 256 valeurs. Nous considérons trois types d'adjacence:

- a. *4-adjacence*. Deux pixels p et q avec des valeurs dans V sont *4-adjacents* si $q \in N_4(p)$.
- b. *8-adjacence*. Deux pixels p et q avec des valeurs dans V sont *8-adjacents* si $q \in N_8(p)$.
- c. *m-adjacence* (adjacence mixte). Deux pixels p et q avec des valeurs dans V sont *m-adjacents* si
 - (i) $q \in N_4(p)$, ou
 - (ii) $q \in N_D(p)$ et l'ensemble $N_4(p) \cap N_4(q)$ n'a aucun pixel dont la valeur est dans V .

L'adjacence mixte est une modification de la 8-adjacence. Elle permet d'éliminer les ambiguïtés générées par l'utilisation de la 8-adjacence. Par exemple, soit l'image binaire de la Fig. 1.11(a) pour $V = \{1\}$. Les trois pixels en haut de la Fig. 1.11(b) constituent une 8-adjacence multiple ambiguë. Cette ambiguïté est supprimée en utilisant la m-adjacence, tel que montré dans la Fig. 1.11(c).

Un *chemin* (ou *courbe*) du pixel p ayant les coordonnées (x, y) au pixel q ayant les coordonnées (s, t) est la séquence des pixels distincts définis par la suite des coordonnées:

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

où $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, sachant que les pixels (x_i, y_i) et (x_{i-1}, y_{i-1}) sont adjacents pour $1 \leq i \leq n$, où n est la *longueur* du chemin. Si $(x_0, y_0) = (x_n, y_n)$ alors le chemin est *fermé*. On peut définir 4-, 8- ou *m-chemins* en fonction du type d'adjacence utilisé. Par exemple, les chemins représentés sur la Fig. 1.11(b) sont des *8-chemins*, alors que le chemin de la Fig. 1.11(c) est un *m-chemin*.

Soit S un sous-ensemble de pixels dans une image. Les deux pixels p et q sont *connectés* dans S , s'il existe un chemin entre p et q constitué uniquement des pixels de S . Pour chaque pixel p de S , l'ensemble des pixels connectés à p dans S est appelé une *composante connexe* de S . S'il n'y a qu'une seule composante connexe dans S , alors S est appelé *ensemble connexe*.

Soit R un sous-ensemble de pixels dans une image. R est une *région* de l'image si R est un *ensemble connexe*. La *frontière* (appelée aussi *bordure* ou *contour*) d'une région R est l'ensemble des pixels dans la région qui ont des voisins hors la région R .

Il existe une différence fondamentale entre les concepts *frontière* et *contour*. La frontière d'une région forme un chemin fermé, donc c'est un concept "global". Alors que, les *contours* sont formés à partir des pixels dont les valeurs des dérivées dépassent un seuil prédéfini, donc c'est un concept "local". Une exception dans laquelle les contours et les frontières se coïncident est dans les images binaires. Par conséquent, il est très important de considérer les contours comme des discontinuités d'intensité et les frontières comme des chemins fermés.

1.7.3 Distances et mesures

Soient p , q et z trois pixels ayant les coordonnées (x, y) , (s, t) et (v, w) , respectivement. On dit que D est une *fonction de distance* ou *métrique* si

- a. $D(p, q) \geq 0$ ($D(p, q) = 0$ ssi $p = q$)
- b. $D(p, q) = D(q, p)$, et
- c. $D(p, z) \leq D(p, q) + D(q, z)$.

La *distance Euclidienne* entre p et q est définie comme

$$D_e(p, q) = [(x - s)^2 + (y - t)^2]^{\frac{1}{2}} \quad (1.11)$$

Pour la distance Euclidienne, les pixels ayant une distance inférieure ou égale à r de (x, y) sont les pixels appartenant au disque défini par le rayon r et centré au point (x, y) .

La *distance D_4* (appelée aussi *distance de Manhattan*) entre le pixel p et q est définie par

$$D_4(p, q) = |x - s| + |y - t|. \quad (1.12)$$

Dans ce cas, les pixels ayant une distance $D_4(x, y)$ inférieure ou égale à r forment un diamant centré au point (x, y) . Par exemple, les pixels ayant une *distance $D_4 \leq 2$* du point (x, y) forment les contours isolignes suivants:

$$\begin{array}{ccccccc} & & & & 2 & & & & \\ & & & & 2 & 1 & 2 & & \\ & & & 2 & 1 & 0 & 1 & 2 & \\ & & & 2 & 1 & 2 & & & \\ & & & & 2 & & & & \end{array}$$

Noter que les pixels avec $D_4 = 1$ sont les 4-voisinages du pixel (x, y) .

La *distance* D_8 (appelée aussi *distance de Tchebychev*) entre p et q est définie par

$$D_8(p, q) = \max(|x - s|, |y - t|) \quad (1.13)$$

Dans ce cas, les pixels dont la distance D_8 du point (x, y) est inférieure ou égale à une certaine valeur r forment un carré centré en (x, y) . Par exemple, les pixels dont la *distance* $D_8 \leq 2$ du point (x, y) (le point central) forment les contours isolignes suivants:

$$\begin{array}{ccccc} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{array}$$

Noter que les pixels avec $D_8 = 1$ forment les 8-voisinages du pixel (x, y) .

1.8 Opérations linéaires vs nonlinéaires

Soit H un opérateur qui produit l'image $g(x, y)$ en sortie, ayant l'image $f(x, y)$ en entrée:

$$H[f(x, y)] = g(x, y) \quad (1.14)$$

H est un *opérateur linéaire* si, pour deux images f_i et f_j et deux scalaires a_i et a_j , on a

$$\begin{aligned} H[a_i f_i(x, y) + a_j f_j(x, y)] &= a_i H[f_i(x, y)] + a_j H[f_j(x, y)] \\ &= a_i g_i(x, y) + a_j g_j(x, y) \end{aligned} \quad (1.15)$$

Par exemple, un opérateur qui calcule la somme de K images est un opérateur linéaire. Par contre, un opérateur qui calcule la valeur absolue de la différence de deux images ne l'est pas. Par définition, un opérateur qui ne réussit pas le test de l'Eq. (1.15) est *non linéaire*.

Les opérations linéaires sont faciles à comprendre et reposent sur une multitude de résultats théoriques et pratiques. Bien que, les systèmes non linéaires ne sont pas toujours faciles à comprendre théoriquement, ils ont été utilisés dans différentes applications car ils dépassent (dans la plupart des cas) les opérations linéaires en termes de performance.

1.9 Conclusion

Dans ce chapitre, nous avons présenté les informations de base pour le reste du cours, tel que le système visuel humain, la lumière et le spectre électromagnétique. Les idées d'échantillonnage sont à la base de certains phénomènes de numérisation rencontrés en pratique. Les concepts d'échantillonnage et quantification sont aussi très importants dans le contexte de l'acquisition de l'image. Les méthodes de traitement des voisinages sont au cœur de nombreuses procédures d'amélioration et de restauration d'image. Ces méthodes sont privilégiées dans les applications commerciales de traitement d'images. Ceci est dû à leur rapidité d'exécution d'une part et de leur simplicité d'implémentation de l'autre part.

1.10 Travaux pratiques N° 1

1.10.1 Objectifs

- Introduire l'environnement de programmation Matlab.
- Introduire les commandes de base disponibles avec le Toolbox Image Processing.

1.10.2 Énoncé

Dans ce projet, nous allons introduire les instructions de base et les fondements de l'environnement de programmation Matlab, ainsi que quelques opérations de base sur les images numériques telles que la lecture d'une image, les traitements simples, l'affichage et l'écriture sur disque.

1) Boucle "for", et opérations sur les vecteurs

Donner les fonctions Matlab qui permettent de calculer les deux sommes suivantes en utilisant deux méthodes différentes: (1) la boucle "for", (2) opérations sur les vecteurs.

$$S = \frac{1^2}{2^2} + \frac{2^2}{3^2} + \cdots + \frac{999^2}{1000^2} \quad (1.16)$$

$$S = 1 - \frac{1}{4} + \frac{1}{7} - \frac{1}{10} + \cdots - \frac{1}{1000} \quad (1.17)$$

2) Opérations de base sur les images numériques

Donner les commandes nécessaires pour réaliser les opérations suivantes en utilisant l'environnement MATLAB:

- Lire l'image contenue dans le fichier 'rose_1024.tif' et afficher sa taille.
- Afficher l'image et parcourir son contenu (valeurs des pixels).
- Sauvegarder cette image avec un format différent (par exemple png).

3) Traitements simple sur les images numériques

Soit l'image f en niveaux de gris contenue dans le fichier 'rose_1024.tif' de taille 1024×1024 , implémenter la série d'instructions suivante sur l'image f :

- Lire l'image f .
- Flipper l'image f verticalement.
- Cropper la région de l'image f délimitée par le rectangle: [257:768, 257:768].
- Rééchantillonner l'image f en divisant sa taille sur 2.
- Afficher le profil de la ligne horizontale du milieu (la ligne 512).