

## Chapitre 1

# Les images numériques

## 1.1 Notions de bases - Vocabulaire

Une image numérique est une fonction à support discret et borné, et à valeurs discrètes. Le support est multidimensionnel, en général 2d ou 3d. Les valeurs peuvent être scalaires (images en niveaux de gris), ou bien vectorielles (imagerie multicomposante, imagerie couleur).

La gamme de valeurs possibles varie en fonction du type d'images considéré ; le tableau 1.1 présente les types les plus courants, les grandeurs physiques associées et les capteurs utilisés.

<i>Phénomène physique</i>	<i>Grandeur mesurée</i>	<i>Capteur</i>
Lumière visible	Flux photonique émis ou réfléchi	CCD, CMOS, Scanner,...
Rayonnement Infra-rouge	Luminance Infra-rouge (Chaleur)	Bolomètres,...
Echo ultrasonore	Distance, densité de tissus,...	Echographie, Sonar,...
Résonance magnétique	Présence d'un corps chimique,...	IRM, RMN,...
Echo électromagnétique	Distance, spécularité de surfaces,...	Radar, SAR,...
Absorption des rayons X	Densité de tissus,...	Radiographie,...

TABLE 1.1 – Quelques types d'images.

Une image numérique est associée à un *pavage* de l'espace, en général rectangulaire. Chaque élément du pavage, appelé *pixel*, est désigné par ses coordonnées entières : Figure 1.1(1).

L'*échantillonnage* est le procédé de discrétisation spatiale d'une image consistant à associer à chaque pixel une unique valeur : Figure 1.1(2). On parle de *sous-échantillonnage* lorsque l'image est déjà discrétisée et qu'on diminue le nombre de pixels.

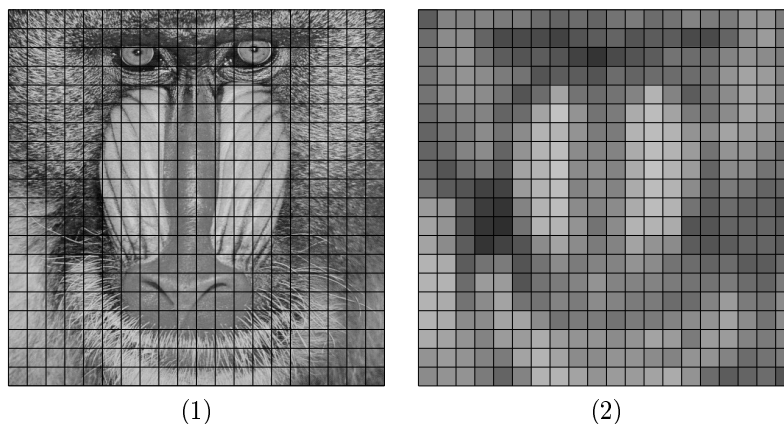


FIGURE 1.1 – (1) Pavage (2) Echantillonnage

La *quantification* désigne la discrétisation tonale correspondant à la limitation du nombre de valeurs différentes que peut prendre chaque pixel.

Une image numérique est donc une image *échantillonnée* et *quantifiée*.

Une image numérique 2d est représentée par un tableau  $I$  de  $h$  lignes et  $w$  colonnes. Le pixel est désignée par un couple  $(i, j)$  où  $i$  est l'indice de colonne  $i \in \{0, w - 1\}$ , et  $j$  l'indice de ligne  $j \in \{0, h - 1\}$ .  $w$  est la *largeur*,  $h$  la *hauteur* de l'image  $I$ . Par convention, le pixel origine  $(0, 0)$  est en général en haut à gauche (voir Figure 1.2). Le nombre  $I(i, j)$  est la *valeur* (ou le *niveau de gris*)

du pixel  $(i, j)$ ,  $I(i, j) \in \{0, N_{max} - 1\}$ .  $N_{max}$  est le nombre de niveaux de gris. On appelle *dynamique* de l'image le logarithme en base 2 de  $N_{max}$ , i.e. le nombre de bits utilisés pour coder l'ensemble des valeurs possibles.

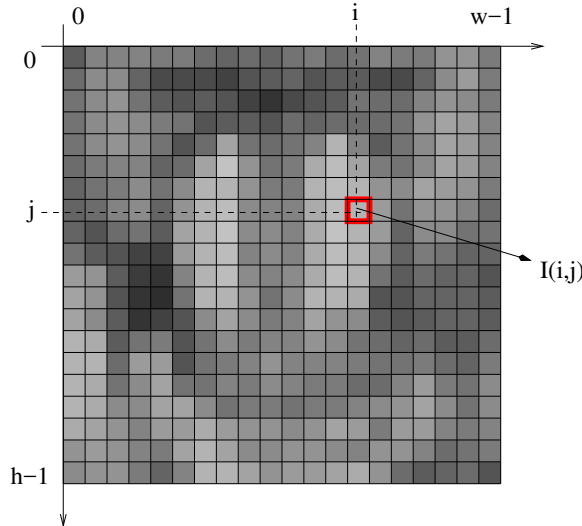


FIGURE 1.2 – Conventions de notation

Une image numérique ne constitue donc qu'une version approchée de l'«image réelle» formée par l'image (au sens mathématique) de la projection de la scène 3d sur la portion de plan correspondant à la surface photosensible du capteur. La qualité de l'approximation dépend de la quantité d'information portée par l'image numérique, en particulier du nombre de pixels utilisés : la *résolution spatiale*. La figure 1.3 montre un exemple d'une même image acquise à des résolutions différentes.

Quelle est l'information conservée (ou perdue) lorsqu'on diminue la résolution ? Ou en d'autres termes, quelle information peut-on reconstituer de l'image réelle à partir d'une image numérique ? C'est la principale question qui fait l'objet des sections 1.2 et 1.3.

Le changement de résolution tonale, correspondant à la quantification, fait aussi apparaître une perte d'information dans les images : voir Figure 1.4, pour la représentation d'une même image avec différentes dynamiques. La bonne dynamique dépend de la qualité des éléments photosensibles du capteur, mais aussi de la richesse du contenu informationnel de l'image, qui est lié à la distribution statistique de ses valeurs. C'est de ce problème que traite la section 1.4.

## 1.2 Echantillonnage

L'échantillonnage est une étape fondamentale qui doit prendre en compte le contenu de l'image à analyser. Intuitivement, on conçoit bien qu'une «structure fine», c'est-à-dire une partie de l'image comportant des oscillations avec de petites périodes spatiales, nécessitera plus de pixels qu'une partie présentant moins de variation.

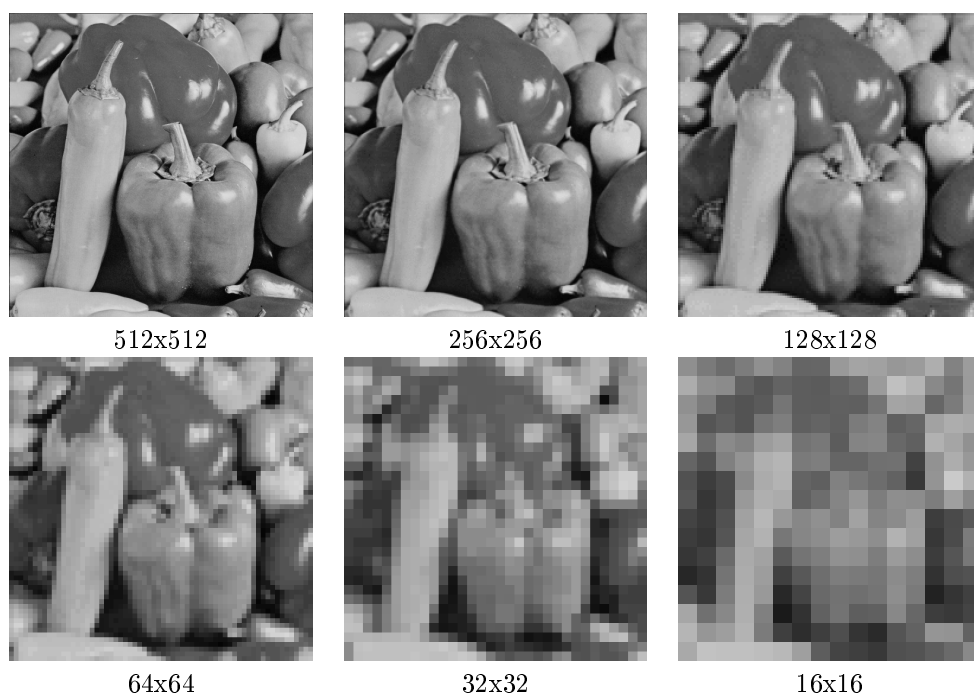


FIGURE 1.3 – Résolution spatiale : échantillonnage

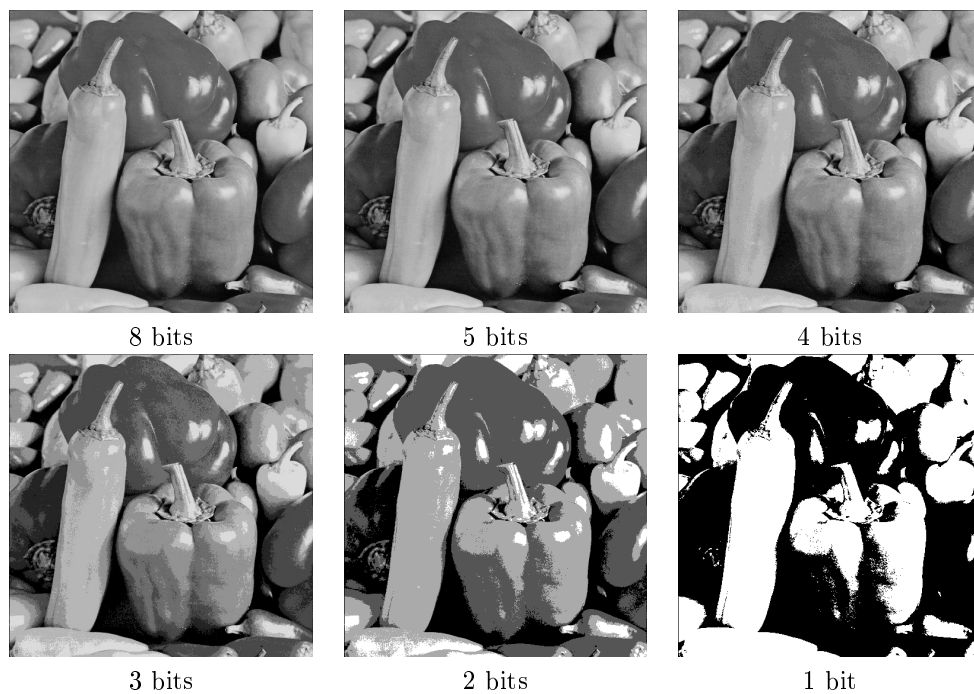


FIGURE 1.4 – Résolution tonale : Quantification

Mais si l'on ne dispose pas d'assez de pixels pour représenter la dite structure fine? Dans ce cas, il est souhaitable que les quelques pixels utilisés pour représenter la structure rendent compte d'un «comportement moyen» de la région concernée. Cependant, si l'on se contente de sous-échantillonner une région trop complexe, ce n'est pas du tout ce qui se produit : en général, de nouvelles structures apparaissent, qui ne correspondent pas à l'image réelle.

Ainsi, sur la figure 1.5, en 1d, on a un signal sinusoïdal qui est échantillonné à une fréquence nettement inférieure à la fréquence du signal. On peut voir que le signal discret résultant est à peu près le même que celui qu'on aurait obtenu en échantillonnant (suffisamment cette fois!) une sinusoïde de fréquence 8 fois plus faible.

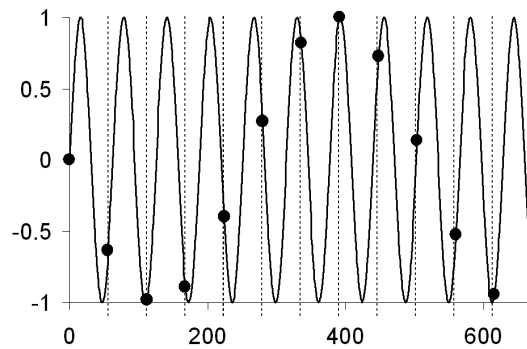


FIGURE 1.5 – Phénomène d'aliasing sur un signal 1d

Ce phénomène est appelé *aliasing*. Sur les signaux 2d que sont les images, il est dans une certaine mesure «encore pire», car il affecte la fréquence et la direction des structures périodiques. Ainsi, sur la figure 1.6, on souhaite échantillonner une image correspondant à des bandes noires et blanches (1). Avec un échantillonnage adapté (2a), on obtient des structures conformes à l'image continue. Mais en considérant seulement 1 échantillon sur 2 (2b), une structure différente apparaît, dont l'analyse (ici des bandes verticales, plus épaisses) ne sera pas conforme à la réalité de l'objet.

Si l'on observe la figure 1.6 de près, on voit que, dans le premier cas, on a utilisé 2 pixels pour représenter une période verticale (correspondant à la succession d'une bande noire et d'une bande blanche), tandis que dans le deuxième cas, on n'a plus qu'1 pixel par période verticale. Ceci est une illustration du théorème de Shannon-Nyquist [33] qui dit qu'on peut reconstruire le signal continu original à partir d'une version échantillonnée de ce signal, si et seulement si la fréquence d'échantillonnage est *au moins le double* de la plus grande fréquence contenue dans le signal original.

Cette notion de reconstruction est liée à la décomposition fréquentielle des fonctions périodiques dans la théorie de Fourier. Une fonction périodique peut ainsi se décomposer en somme de fonctions périodiques élémentaires (les sinusoïdes). L'équation de synthèse suivante, dite *transformée de Fourier inverse* fournit une telle décomposition sous forme de sinusoïdes complexes :

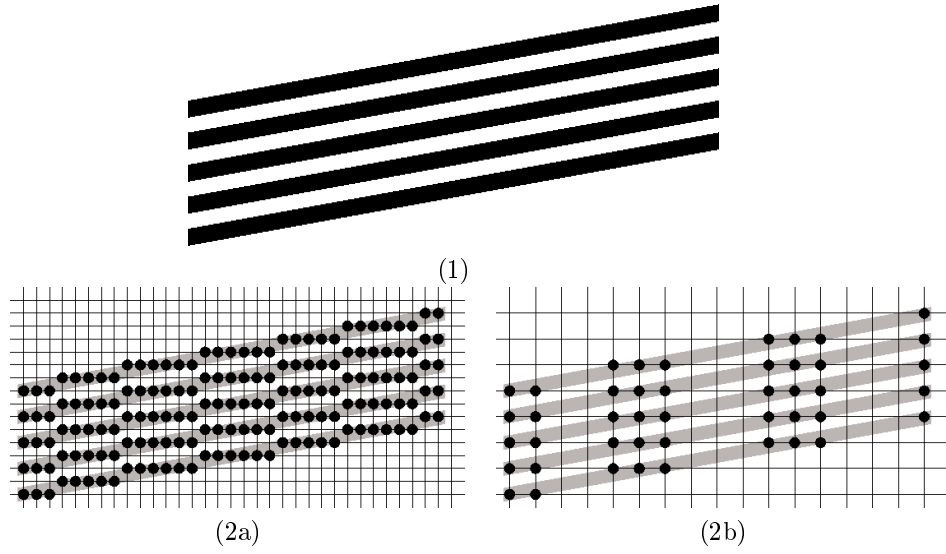


FIGURE 1.6 – Aliasing sur un signal 2d

$$f(t) = \int_{-\infty}^{+\infty} F(u) e^{i2\pi ut} du \quad (1.1)$$

$f$  est ici un signal 1d,  $t$  est le paramètre spatial (ou temporel),  $u$  le paramètre fréquentiel, les fonctions  $\{e^{i2\pi ut}\}_{u \in \mathbb{R}}$  représentent les vecteurs de base de la décomposition (sinusoïdes complexes).  $F(u)$  représente le coefficient attaché à la fréquence  $u$ . Il est calculé par l'équation d'analyse, dite *transformée de Fourier* :

$$F(u) = \int_{-\infty}^{+\infty} f(t) e^{-i2\pi ut} dt \quad (1.2)$$

Supposons maintenant qu'on souhaite échantillonner régulièrement la fonction  $f$  avec une période  $T$ . La fonction  $f_T$  résultante peut être définie comme suit :

$$f_T(t) = f(t) \cdot \Delta_T(t) \quad (1.3)$$

Avec  $\Delta_T$  le peigne de Dirac de période  $T$  :

$$\Delta_T(t) = \sum_{k=-\infty}^{k=+\infty} \delta(t - kT) \quad (1.4)$$

$\Delta_T$ , en tant que fonction périodique de période  $T$ , se décompose en série de Fourier sous la forme :

$$\Delta_T(t) = \sum_{n=-\infty}^{n=+\infty} c_n e^{i2\pi nt/T} \quad (1.5)$$

Où les coefficients  $c_n$ , sont définis, pour une origine  $t_0 \in \mathbb{R}$  quelconque :

$$c_n = \frac{1}{T} \int_{t_0}^{t_0+T} \Delta_T(t) e^{-i2\pi nt/T} dt \quad (1.6)$$

$$= \frac{1}{T} \int_{-T/2}^{T/2} \Delta_T(t) e^{-i2\pi nt/T} dt \quad (1.7)$$

$$= \frac{1}{T} \quad (1.8)$$

$$(1.9)$$

Finalement :

$$\Delta_T(t) = \frac{1}{T} \sum_{n=-\infty}^{n=+\infty} e^{i2\pi nt/T} \quad (1.10)$$

Si l'on calcule à présent  $F_T$ , la transformée de Fourier de  $f_T$  :

$$F_T(u) = \int_{-\infty}^{+\infty} f_T(t) e^{-i2\pi ut} dt \quad (1.11)$$

$$= \frac{1}{T} \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{n=+\infty} e^{i2\pi nt/T} f(t) e^{-i2\pi ut} dt \quad (1.12)$$

$$= \frac{1}{T} \sum_{n=-\infty}^{n=+\infty} \int_{-\infty}^{+\infty} e^{-i2\pi(u-n/T)t} f(t) dt \quad (1.13)$$

$$(1.14)$$

On obtient finalement :

$$F_T(u) = \frac{1}{T} \sum_{n=-\infty}^{n=+\infty} F(u - n/T) \quad (1.15)$$

Cette identité constitue une preuve du théorème de Shannon-Nyquist évoqué plus haut. En effet, si la fréquence maximum contenue dans la fonction  $f$  est inférieure à  $1/2T$ , c'est-à-dire si le support de  $F$  est contenu dans  $] -1/2T, 1/2T[$  (et donc, si la fréquence d'échantillonnage est au moins le double de la plus grande fréquence contenue dans  $f$ ), dans ce cas la transformée de Fourier de la fonction  $f$  échantillonnée, est égale, à un facteur multiplicatif près, à la réplication de la transformée de Fourier de la fonction originale  $f$ , régulièrement espacée d'une longueur  $1/T$ .

Cela signifie que, du point de vue de l'information contenue, le signal échantillonné  $f_T$  est équivalent au signal initial  $f$ , puisqu'on peut retrouver ce dernier en restreignant la transformée de Fourier de  $f_T$  sur  $] -1/2T, 1/2T[$ , puis en calculant la transformée de Fourier inverse (voir Figure 1.7(1)). Cette opération, comme nous le verrons dans le chapitre suivant, équivaut à un *filtrage linéaire*.

En revanche, si la condition de Shannon-Nyquist n'est pas respectée, c'est-à-dire si la fréquence d'échantillonnage est inférieure au double de la plus grande fréquence contenue dans  $f$ , l'équation 1.15 ne correspond plus à une simple réplication car le support des différentes versions de  $F(u - n/T)$  s'intersectent,

et les valeurs s'ajoutent. On ne peut plus retrouver le signal original, et le signal sous-échantillonné présente de nouvelles structures correspondant aux régions fréquentielles qui se recoupent (voir Figure 1.7(2)). On parle de *recouvrement de spectre* ou *aliasing*.

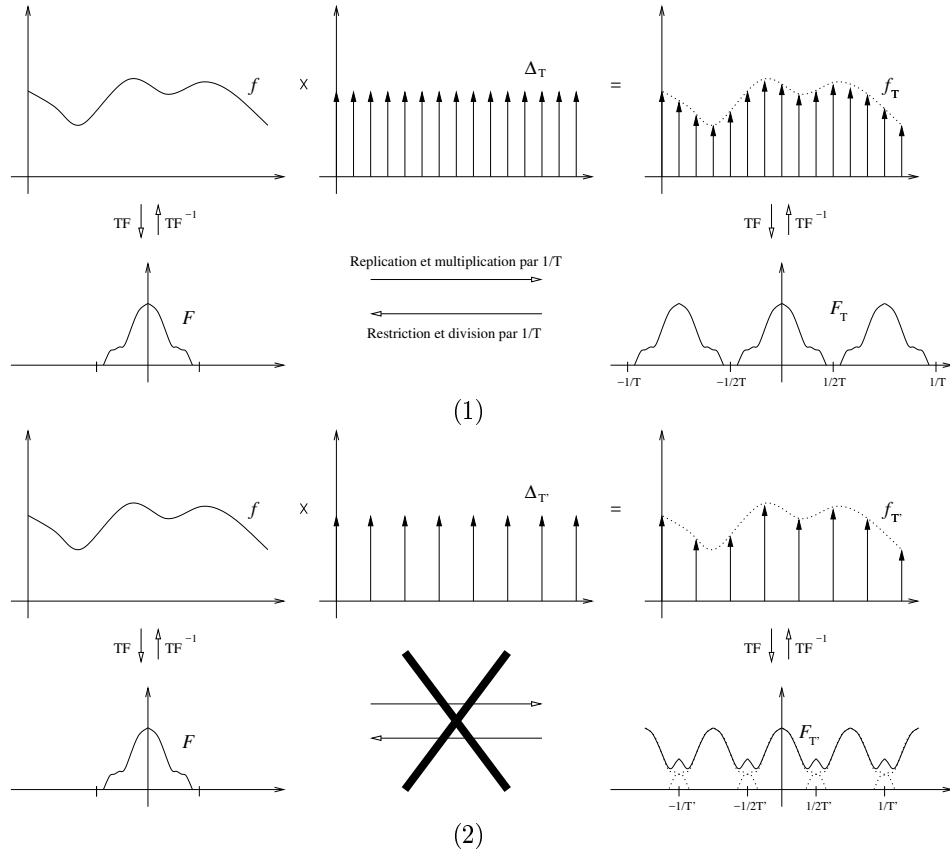


FIGURE 1.7 – Illustration du théorème de l'échantillonnage de Shannon-Nyquist (1) Critère de Shannon respecté, et (2) Critère de Shannon non respecté

Le phénomène d'aliasing va naturellement se retrouver dès lors qu'on sous-échantillonne une image numérique (supposée bien échantillonnée). L'image sous-échantillonnée fait alors apparaître des structures qui n'appartenaient pas à l'image naturelle (et donc pas à l'image numérique initiale). Pour sous-échantillonner sans faire apparaître des structures parasites, il faut donc réduire l'intervalle de fréquences de l'image numérique initiale (c'est-à-dire la filtrer), avant de la sous-échantillonner. La figure 1.8 illustre ce principe.

La transformée de Fourier que nous avons utilisée dans cette partie nous a permis de modéliser les images et de comprendre les règles d'échantillonnage qui doivent régir la formation d'une image numérique. On notera néanmoins que la transformée de Fourier - même d'une fonction discrète - est toujours une fonction continue. Dans la partie suivante, nous présentons la *transformée de Fourier discrète*, qui peut être représentée sur ordinateur, et donc être utilisée effectivement pour le traitement des images numériques.





(1)



(2)



(3)

FIGURE 1.8 – (1) Une image numérique de taille  $512 \times 512$ , (2) La même image sous-échantillonnée, en prenant 1 pixel sur 2 dans chaque dimension, (3) La même image filtrée, pour éliminer les hautes fréquences, puis sous-échantillonnée en prenant 1 pixel sur 2 dans chaque dimension.

### 1.3 Transformée de Fourier

Dans cette partie, on notera les arguments entre crochets pour souligner leur nature discrète. On notera  $[x, y]$  les coordonnées du domaine spatial,  $[u, v]$  les coordonnées du domaine fréquentiel.

La transformée de Fourier discrète 2d d'une image numérique  $f[x, y]$ , de largeur  $w$  et de hauteur  $h$  est donnée par :

$$F[u, v] = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} f[x, y] e^{-i2\pi(ux/w + vy/h)} \quad (1.16)$$

Cette transformée fournit les coefficients  $F[u, v]$  de la décomposition de l'image  $f$  en combinaison linéaire de sinusoïdes complexes, correspondant à la transformée de Fourier discrète 2d inverse :

$$f[x, y] = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} F[u, v] e^{i2\pi(ux/w + vy/h)} \quad (1.17)$$

$F[u, v]$  et  $f[x, y]$  sont des nombres complexes en général, mais dans notre cas  $f[x, y]$  qui représente l'image numérique dans le domaine spatial sera toujours réel, et donc  $F[u, v]$  sera toujours une fonction complexe paire :

$$F[u, v] = \overline{F[-u, -v]} \quad (1.18)$$

Où  $\bar{z}$  représente le conjugué de  $z$ . En termes d'interprétation, on préfère la représentation (module, argument) des nombres complexes  $F[u, v]$  à la représentation (partie réelle, partie imaginaire). On représentera donc les fréquences par :

$$F[u, v] = ||F[u, v]|| e^{i\varphi[u, v]} \quad (1.19)$$

Avec  $||F[u, v]||$  le module, et  $\varphi[u, v]$  l'argument de  $F(u, v)$ . Le module correspond à l'*amplitude* des sinusoïdes, et l'argument à leur *phase*. Comme  $F$  est paire dans notre cas, on aura donc :  $||F[u, v]|| = ||F[-u, -v]||$  et  $\varphi[u, v] = -\varphi[-u, -v]$ .

$F$  et  $f$  sont des fonctions 2d-périodiques de périodes  $(w, h)$  :

$$F[u, v] = F[u + w, v + h] \quad (1.20)$$

$$f[x, y] = f[x + w, y + h] \quad (1.21)$$

Ceci est un point important dans la modélisation fréquentielle : les images numériques sont supposées infinies périodiques. On doit donc avoir une hypothèse de prolongement des images à l'extérieur de leur support. Généralement, on admet que l'image numérique observée correspond à la plus grande période d'une image infinie périodique : voir Figure 1.9.

Mais cette hypothèse a un effet important sur l'interprétation de la transformée de Fourier, car la périodisation crée en général des contours verticaux et horizontaux artificiels qui se traduisent par des amplitudes importantes pour les fréquences horizontales et verticales, soit des valeurs élevées au niveau des axes des fréquences.

Pour des raisons de symétries, on représente habituellement les valeurs de  $||F[u, v]||$  (resp. de  $\varphi[u, v]$ ) dans une image appelée *spectre d'amplitude* (resp.

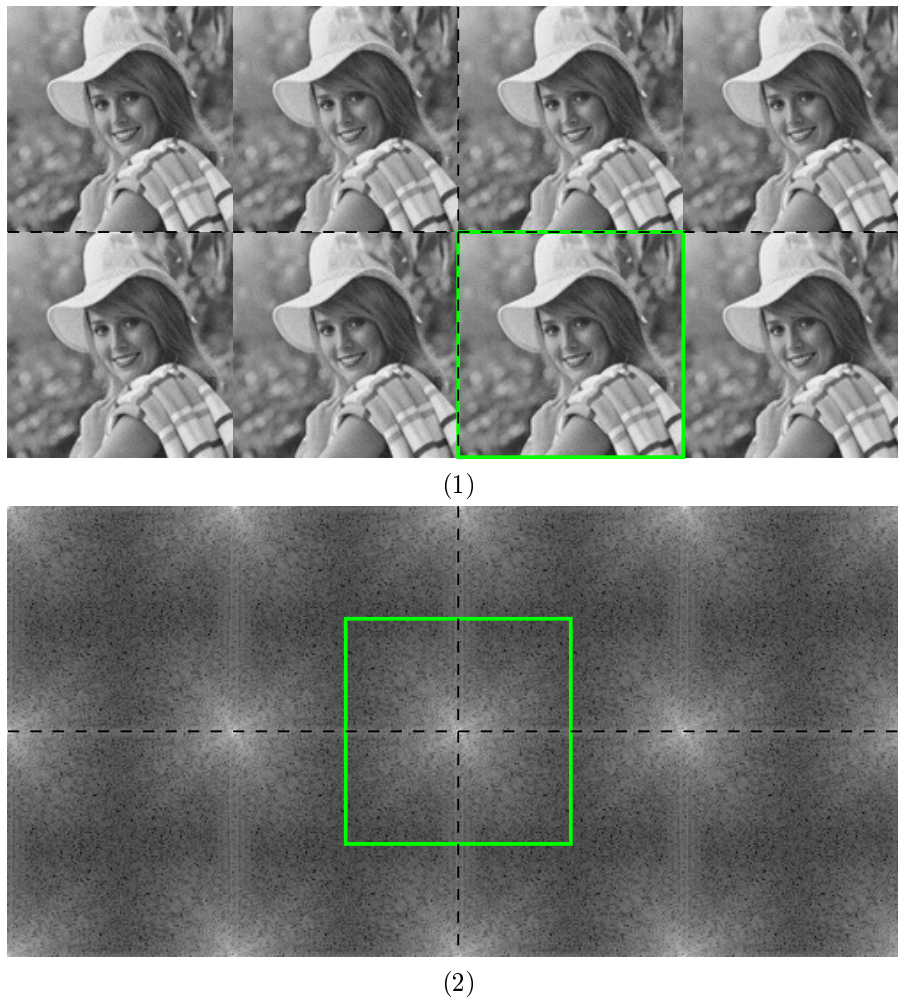


FIGURE 1.9 – Périodicité de l'image (1) et de son spectre d'amplitude (2) dans la modélisation fréquentielle. Le rectangle montre les périodes habituellement représentée.

*spectre de phase*), montrant (en niveaux de gris) les valeurs sur la période  $] -w/2, w/2[\times] -h/2, h/2[$ . La figure 1.10 montre les spectres d'amplitudes et de phase associés à une image exemple.

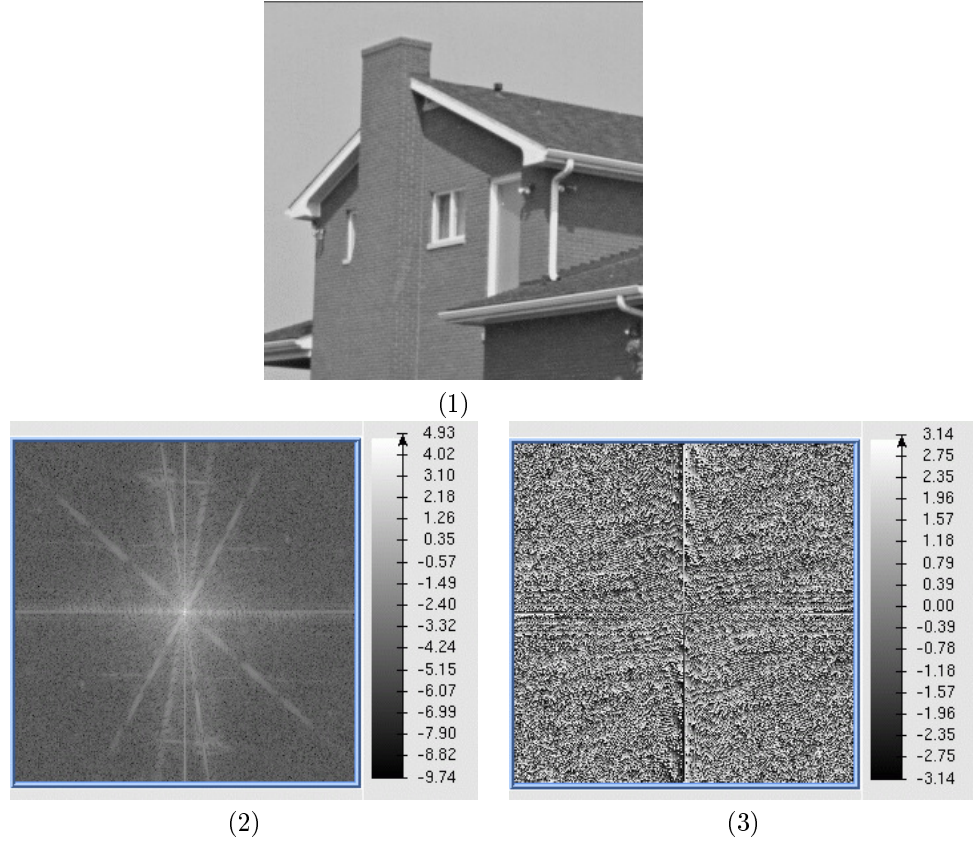


FIGURE 1.10 – Image «House» (1), son spectre d'amplitude (2), représentant (sur 255 niveaux de gris) le logarithme de  $||F[u, v]||$ , et son spectre de phase (3), représentant (sur 255 niveaux de gris)  $\varphi[u, v]$

Les valeurs de l'amplitude présentant en général une très grande dynamique (les basses fréquences plusieurs ordres de grandeur fois plus importantes que les hautes fréquences), le spectre d'amplitude représente le logarithme de  $||F[u, v]||$ .

Outre les effets de la réplication périodique de l'image générant de fortes valeurs au niveaux des axes, le spectre d'amplitude fournit des informations sur les structures présentes dans l'image : fréquences et directions principales. On peut remarquer en particulier que des contours rectilignes très marqués (type en escalier) génèrent des concentrations plus ou moins diffuses de fortes valeurs, le long des directions perpendiculaires à ces contours. Cela s'explique par le fait que les fonctions en escalier ont une représentation fréquentielle riche, correspondant à un ensemble de sinusoides de même orientation (i.e. de fréquences multiples).

Il va de soi qu'en revanche l'interprétation visuelle du spectre de phase n'apporte rien. Mais on ne doit pas en conclure que la phase n'apporte pas d'information sur les images. Au contraire, la phase contient l'information sur la

position relatives des différentes composantes sinusoïdales les unes par rapport aux autres, ce qui joue un rôle déterminant dans la forme des objets. La figure 1.11 montre un exemple de ce qu'on appelle les «fantômes» d'amplitude et de phase, qui consistent à ne conserver que l'un ou l'autre du module ou de la phase, et d'appliquer la transformée de Fourier discrète inverse. On voit que le fantôme d'amplitude se traduit par une «texture» dans lequel on retrouve statistiquement les caractéristiques fréquentielles de l'image, mais que la forme des objets est perdue. En revanche, cette forme est toujours visible sur le fantôme de phase.

## 1.4 Quantification et histogrammes

La quantification, comme on l'a vu dans la première partie de ce chapitre désigne la résolution tonale, c'est-à-dire le nombre de valeurs différentes (niveaux de gris) qu'on peut associer à un pixel.

Idéalement, le nombre de valeurs différentes devrait dépendre de l'amplitude des grandeurs observées (réflectance de lumière visible, luminance infrarouge, ...) dans la scène. Mais en pratique, le nombre de valeurs utilisées pour coder une image lors de son acquisition dépend de la capacité effective du capteur à observer des signaux de grandeurs différentes, qui s'assimile à un *rapport signal sur bruit*. On peut définir cette quantité comme le rapport entre l'amplitude des niveaux de gris mesurables par le capteur ( $n_{max} - n_{min}$ ), et le niveau du bruit, correspondant à l'écart-type  $s_b$  de la perturbation aléatoire qui affecte les niveaux de gris. En prenant le logarithme, on obtient le nombre de bits utiles au capteur pour coder les images.

Outre les capacités du capteur, le nombre de bits réellement nécessaire pour coder une image peut varier beaucoup d'une image à l'autre, en fonction de la richesse informationnelle du contenu. Il est clair qu'une image ne contenant que quelques valeurs différentes (image uniforme par exemple) peut être codée en utilisant peu de bits, mais c'est aussi le cas pour une image où toutes les valeurs seraient représentées, mais dans des proportions très différentes : en effet, il est alors possible d'utiliser un codage à longueur variable, où les valeurs les plus nombreuses seraient représentées sur un codage court, et les valeurs moins nombreuses sur un codage plus long.

La quantité d'information ainsi représentée par l'ensemble des valeurs contenue dans une image se mesure par une grandeur appelée *entropie* [32]. L'entropie se calcule à partir de la distribution des niveaux de gris, c'est-à-dire l'*histogramme*.

Soit  $I$  une image de largeur  $w$  et de hauteur  $h$ , pouvant prendre  $N$  niveaux de gris différents. L'histogramme  $H_I$  est une fonction de  $\{0, N - 1\}$  dans  $\mathbb{N}$  définie par :

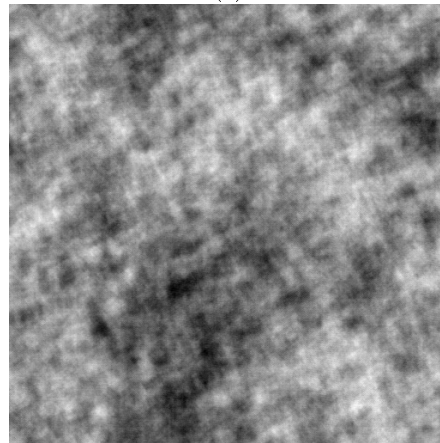
$$H_I(n) = \text{Card}(\{(x, y) \in \{0, w - 1\} \times \{0, h - 1\}; I(x, y) = n\}) \quad (1.22)$$

$H_I(n)$  est donc le nombre de pixels de  $I$  qui ont la valeur  $n$ . On utilise aussi souvent l'histogramme normalisé  $P_I$ , qui est une fonction de  $\{0, N - 1\}$  dans  $[0, 1]$  définie par :

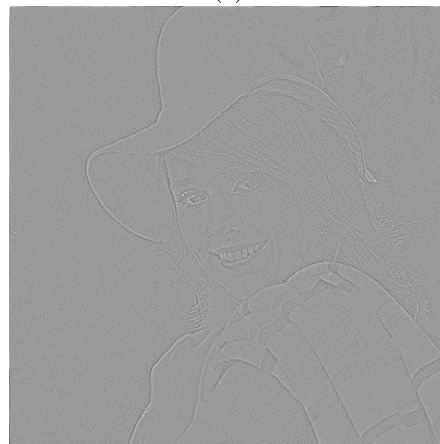
$$P_I(n) = \frac{H_I(n)}{wh} \quad (1.23)$$



(1)



(2)



(3)

FIGURE 1.11 – Image «Elaine» (1), son fantôme d'amplitude (2) consistant à conserver  $\|F[u, v]\|$  et prendre  $\varphi[u, v]$  aléatoire, et fantôme de phase (3) consistant à conserver  $\varphi[u, v]$  et prendre  $\|F[u, v]\|$  constant.

$P_I(n)$  est donc la proportion de pixels de  $I$  qui ont la valeur  $n$ . Enfin, on utilise l'histogramme cumulé normalisé  $R_I$ , qui est une fonction de  $\{0, N-1\}$  dans  $[0, 1]$  définie par :

$$R_I(n) = \sum_{k=0}^n P_I(k) = \frac{1}{wh} \sum_{k=0}^n H_I(k) \quad (1.24)$$

$R_I(n)$  est donc la proportion de pixels de  $I$  dont la valeur est inférieure ou égale à  $n$ .

L'histogramme est une observation statistique, mais il est parfois utilisé comme modèle probabiliste empirique; dans ce cas l'histogramme normalisé correspond à une densité de probabilité, et l'histogramme cumulé normalisé à une fonction de répartition.

L'entropie  $E$  se calcule à partir de l'histogramme normalisé de la façon suivante :

$$E = \sum_{k=0}^{N-1} -P_I(k) \log_2(P_I(k)) \quad (1.25)$$

$E$  représente le nombre moyen de bits par pixel nécessaire pour coder toute l'information présente dans l'image  $I$ . Cette mesure est utilisée dans les techniques de compression sans perte pour adapter le volume de données des images à leur contenu informationnel.

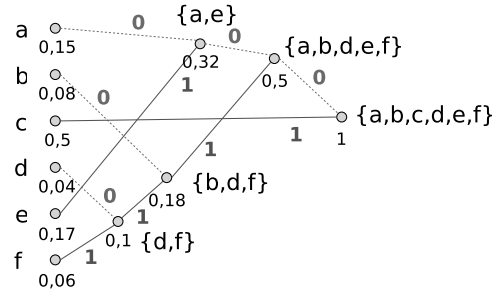
La technique de codage de Huffman [14], par exemple, permet d'associer à chaque valeur un code binaire de longueur inversement proportionnelle à sa fréquence dans l'image  $I$ . Soit  $C_I = \{c_1, \dots, c_K\}$  l'ensemble des valeurs ayant au moins une occurrence dans  $I$ . L'algorithme de Huffman est le suivant :  $\mathcal{C} = \{s_1, \dots, s_L\}$  est un ensemble de symboles, chaque symbole étant associé à une partie de  $C_I$  :  $s_i \subset C_I$ , et  $p(s_i) = \sum_{c_k \in s_i} P_I(c_k)$ .

Initialement,  $\mathcal{C}^0 = \{\{c_1\}, \dots, \{c_K\}\}$ ; c'est l'ensemble des singletons de  $C_I$ . On recherche  $s_1^0$  et  $s_2^0$  les 2 symboles de  $\mathcal{C}^0$  de probabilité la plus faible. On donne à l'un l'étiquette «0» et à l'autre l'étiquette «1». Puis on recommence avec  $\mathcal{C}^1 = \mathcal{C}^0 \setminus \{\{s_1^0\}, \{s_2^0\}\} \cup \{s_1^0, s_2^0\}$ , et ainsi de suite jusqu'à ce que  $\mathcal{C}^n = \{C_I\}$ . Cela revient à construire un arbre binaire où chaque nœud est une partie de  $C_I$ , les feuilles sont les caractères et la racine l'ensemble  $C_I$  tout entier. Le code associé à un caractère  $c$  est alors obtenu en concaténant les étiquettes binaires associées aux symboles (parties de  $C_I$ ) contenant  $c$ , depuis la racine jusqu'à la feuille  $\{c\}$  (voir exemple sur la figure 1.12, avec 6 valeurs différentes symbolisées par les caractères de  $a$  à  $f$ ).

L'histogramme constitue donc un outil de base pour l'étude d'un capteur ou de la dynamique d'une scène. Mais il faut préciser qu'il ne caractérise en rien la géométrie des objets, et on peut transformer radicalement l'histogramme d'une image sans changer significativement son contenu.

Toutefois il existe un certain nombre d'opérateurs de traitements d'images qui modifient les valeurs de niveaux de gris uniquement à partir de l'histogramme. Ces opérateurs sont souvent calculés au niveau même du capteur, et, en général, leur pertinence est très intimement liée aux conditions d'acquisition.

Ainsi la *normalisation d'histogramme*, ou expansion de dynamique, est une transformation affine du niveau de gris des pixels de telle sorte que l'image uti-



Symbole	Probab.	Code
a	0,15	000
b	0,08	010
c	0,5	1
d	0,04	0110
e	0,17	001
f	0,06	0111

FIGURE 1.12 – Algorithme de codage de Huffman. A gauche : arbre binaire construit par l'algorithme d'encodage. A droite : tableau montrant pour chaque symbole (valeur)  $s$ , sa probabilité empirique dans  $I$  ( $P_I(s)$ ), et le code de Huffman résultant.

lise toute la dynamique de représentation. Si l'on note  $d$  la dynamique (nombre de bits utilisés dans la représentation),  $M$  (resp.  $m$ ) la plus grande (resp. petite) valeur dans l'image  $I$ , l'image normalisée  $I_{norm}$  est obtenue par la formule suivante :

$$I_{norm}(x, y) = (I(x, y) - m) \frac{2^d - 1}{M - m} \quad (1.26)$$

Dans l'expression ci-dessus, il suffit qu'un seul pixel ait la valeur 0 et un seul la valeur  $2^d - 1$  pour rendre la normalisation inopérante. On utilise donc plutôt un paramètre  $\alpha$ ,  $0 \leq \alpha \leq 1/2$ , et on prend :

$$m_\alpha \in R_I(\alpha)^{-1} \quad (1.27)$$

$$M_\alpha \in R_I(1 - \alpha)^{-1} \quad (1.28)$$

Remarquons que prendre  $\alpha = 0$  revient à la normalisation absolue, définie par l'égalité 1.26, tandis que prendre  $\alpha = 1/2$  équivaut à réaliser un seuillage à la valeur médiane de l'image.

L'égalisation d'histogramme est une transformation des niveaux de gris dont le principe est d'équilibrer le mieux possible la distribution des pixels dans la dynamique (Idéalement, on cherche à obtenir un histogramme plat).

La technique classique consiste à ajuster les niveaux de gris en fonction de l'histogramme cumulé normalisé, de telle sorte que le nouvel histogramme cumulé normalisé soit «le plus linéaire possible».

$$I_{egal}(x, y) = (2^d - 1)R_I(I(x, y)) \quad (1.29)$$

La Figure 1.13 montre un exemple de ces traitements et de la manière dont ils agissent sur les histogrammes. On obtient une accentuation globale du contraste sur les images. Attention, les opérations de normalisation et d'égalisation telles que décrites par les égalités 1.26 et 1.29 n'augmentent pas l'entropie, car ce sont des transformations bijectives entre deux ensemble de niveaux de gris. On peut même noter que l'opération de normalisation paramétrée fait baisser l'entropie, puisqu'elle diminue le nombre de niveaux de gris en confondant les valeurs extrêmes.



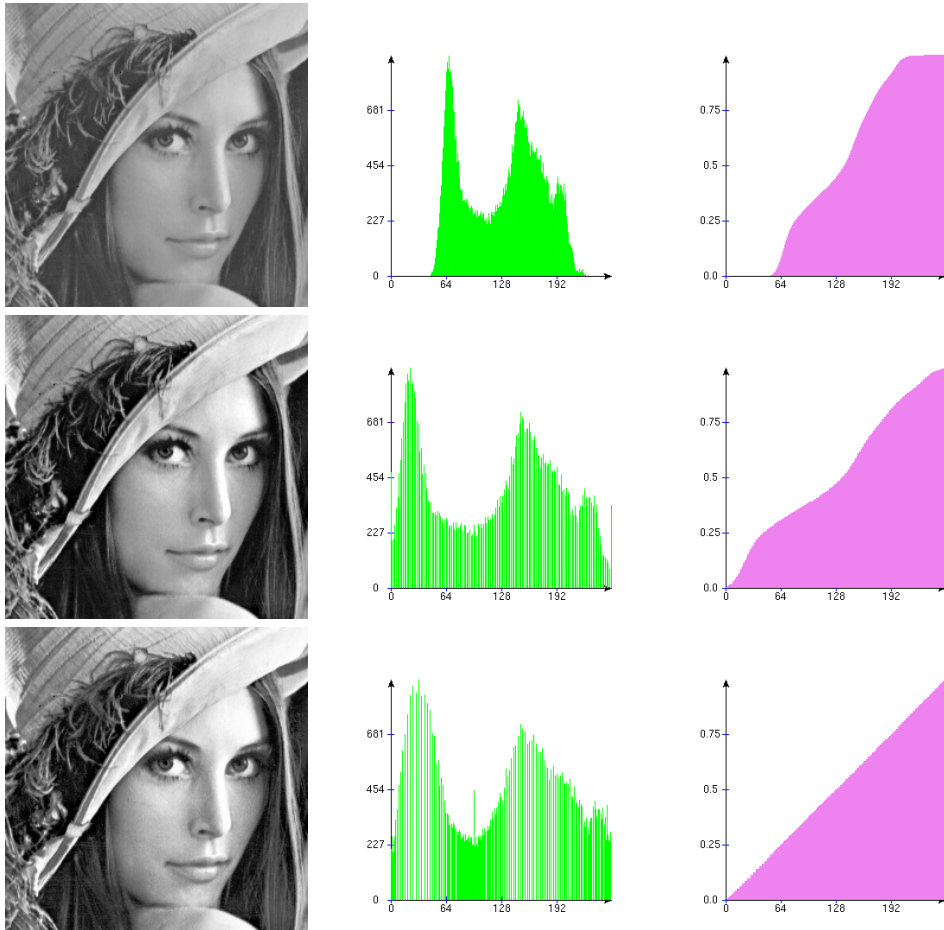


FIGURE 1.13 – 1ère colonne : Image «Lena» (détail), 2ème colonne : Histogramme, 3ème colonne : Histogramme cumulé normalisé. 1ère ligne : Original, 2ème ligne : Normalisation d'histogramme avec le paramètre  $\alpha = 0,005$ , 3ème ligne : Egalisation d'histogramme.

