

Chapitre 1

Traitement d'image

Préambule

Ce premier chapitre traite de l'image numérique individuelle : on y aborde les notions essentielles de quantité d'information, de contraste, de spectre, de contour, de forme et de couleur ; et l'on présente les outils mathématiques associés : codage statistique, transformations intégrales, filtrages, morphologie mathématique, segmentation et classification. On en profite pour exposer aussi les principales propriétés du système visuel humain reposant sur la trichromie.

Le second chapitre portera sur le traitement de séquences d'images numériques (vidéos) introduisant la notion de mouvement. Enfin le troisième chapitre présente les formats et les standards de codage des images et vidéos, développés notamment pour la compression.

1.1 Introduction

La vision est la **perception** du monde extérieur [20, 57]. L'image est la **représentation** bidimensionnelle d'une scène 3D (à trois dimensions). C'est l'information issue d'un capteur de vision (caméra, œil). Le traitement d'image [81, 104, 117] comporte donc différentes tâches : capture, analyse et traitement (bas niveau), interprétation et décision (haut niveau). On distingue différents traitements :

- amélioration, restauration et correction d'image : augmentation du contraste, correction des distorsions optiques, filtrage du bruit,
- analyse [40, 26] : détection et localisation d'objets, segmentation, reconnaissance de formes, estimation et mesures,
- codage pour la compression, l'encryptage, l'archivage et la transmission numérique [65].

Le traitement d'image couvre plusieurs domaines [12, 21, 5] :

- l'**électronique** (capteur, optique, acquisition),
- le **traitement de signal** donc la mathématique (détection, segmentation, estimation, décision) [16],
- l'**informatique** [14, 15, 28] et l'intelligence artificielle [44] (reconnaissance, interprétation, guidage robotique).

On ne s'intéresse ici qu'aux **images numériques**, c'est-à-dire échantillonnées et quantifiées. Une image numérique est définie par un tableau de pixels de taille $L \times C$ (typ. 256×256). Chaque pixel, portant l'information d'intensité lumineuse perçue en ce point, est codé sur n bits. Typ. $n = 8$ pour une image à 256 Niveaux de Gris (NdG) ou $n = 24$ pour une image couleur Rouge Vert Bleu (RVB), avec 8 bits par composante couleur.

On définit dans l'image un repère Oxy et une notion de distance et de voisinage (Fig. 1.1).

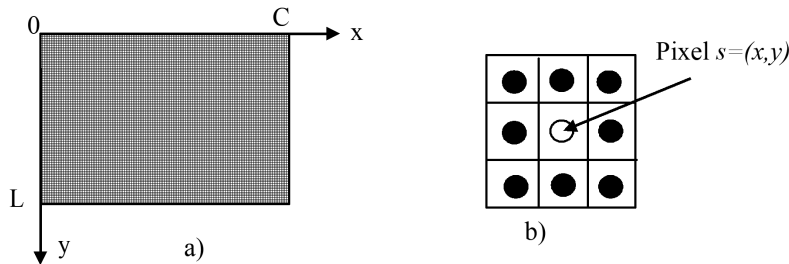


FIGURE 1.1 – a) Repère relatif à l'image ; b) 8-voisinage d'un pixel.

1.2 Rappels de théorie de l'information

1.2.1 Entropie d'une source discrète

L'entropie H (exprimée en bits) d'une source discrète d'observations à valeurs dans $\{0 \dots M - 1\}$ (typ. $M = 256$) est définie par [118] :

$$H_{bit} = - \sum_{i=0}^{M-1} p_i \log_2 p_i \quad (1.1)$$

où p_i représente la probabilité pour que l'observation en le site $s = (x, y) \in S$ vaille : $o_s = i$ (S étant le support des observations, en l'occurrence ici une image de taille $L \times C$).

Rappelons que le choix de la base du logarithme est arbitraire et correspond au choix de l'unité de mesure (le bit si l'on choisit le logarithme binaire, le nat

si l'on prend le logarithme népérien) ; et l'on a :

$$H_{nat} = H_{bit} \cdot \log_e 2 \approx 0.7H_{bit} \quad (1.2)$$

En effet, le passage d'une base a à une base b requiert simplement une multiplication par la constante $K = \log_b a$.

Comme l'information élémentaire apportée par un événement de probabilité p_i vaut par définition¹ : $I_i = \log \frac{1}{p_i}$, on voit que l'entropie est une mesure de l'**information moyenne** débitée par une source : $H = \sum_i p_i I_i$.

Shannon a démontré que l'entropie (exprimée en nats) d'une source gaussienne centrée d'écart-type σ vaut $H_{nat} = \log(\sigma \cdot \sqrt{2\pi e})$ et qu'à σ fixé, la source donnant l'entropie maximale est la distribution gaussienne.

Pour une source quelconque d'entropie H donnée, Shannon définit la notion de « puissance entropique » P_e :

$$P_e = \frac{1}{2\pi e} \exp(2H) \quad (1.3)$$

C'est la puissance de la distribution gaussienne équivalente à la distribution d'entropie H . Pour un bruit gaussien centré d'écart-type σ , on retrouve bien sûr : $P_e = \sigma^2$. Comme le bruit blanc gaussien possède l'entropie maximale à puissance fixée, la puissance entropique d'un bruit quelconque est toujours inférieure ou égale à sa puissance effective. Toujours selon Shannon, un bruit blanc gaussien a la propriété d'absorber tout autre signal qui lui est ajouté. La puissance entropique résultante est à peu près égale à la somme de la puissance du bruit blanc et de la puissance du signal (supposé centré), à condition que la puissance du signal soit faible, « dans un certain sens », comparée au bruit [118].

En supposant le bruit additif gaussien et majoritaire sur le support, on peut donc estimer un écart-type entropique σ_e équivalent en calculant l'entropie H des observations [97] :

$$\sigma_e = \sqrt{P_e} = \frac{\exp(H)}{\sqrt{2\pi e}} = \frac{2^{H_{bit}}}{\sqrt{2\pi e}} \quad (1.4)$$

En prenant une mesure de seuil θ estimée par :

$$\theta = 4\sigma_e \approx 2^{H_{bit}} \quad (1.5)$$

on obtient une technique automatique de seuillage adaptatif, applicable entre autres à la détection du mouvement dans les séquences d'images [96, 98]

§4.62

1. Que l'information varie à l'inverse de la probabilité relève du sens commun : un message très probable, tel que « Demain, le soleil se lèvera », apporte très peu d'information. Au contraire, un message très peu probable, comme « Le soleil ne se lèvera pas demain » apporte beaucoup d'information (s'il est vrai!). C'est la notion même des « nouvelles », au sens journalistique (cf. le *scoop* du journaliste).

1.2.2 Redondance d'une source d'information

L'entropie est une notion fondamentale car elle renseigne sur la redondance spatiale présente dans une image et permet de mettre en œuvre les outils de codage définis par Shannon pour réaliser une compression sans perte de l'information contenue dans l'image (économie de débit pour la transmission : codage de Huffman, de Shannon-Fano). La redondance d'une source X est définie par :

$$R = 1 - \frac{H(X)}{H_{max}} \quad (1.6)$$

Pour une source discrète à M événements, l'entropie maximale est atteinte quand les événements sont équiprobables et elle vaut :

$$H_{max} = \log_2 M. \quad (1.7)$$

1.2.3 Codage entropique

Les méthodes de codage entropique visent à coder les messages les plus probables avec les mots-codes les plus courts : on obtient un code à longueur variable, en anglais Variable Length Coding (VLC). Elles permettent d'assurer la propriété du préfixe (obtention d'un code « irréductible »), et de faire en sorte que $p(0) \approx p(1)$, pour réduire la redondance et atteindre un taux de compression optimal sans perte.

1.2.3.1 Méthode de Shannon-Fano

La procédure est la suivante §4.53 :

1. Classer dans un tableau les messages (ou événements) par ordre de probabilité non croissante.
2. Découper l'ensemble X des messages en 2 sous-ensembles X_1 et X_2 ayant des probabilités voisines (aussi proches que possible de l'équiprobabilité).
3. Attribuer à l'un des sous-ensembles le préfixe 0 et à l'autre 1 (comme première lettre du mot-code).
4. Répéter la séquence d'opérations 2 & 3 sur X_1 puis sur X_2 séparément, etc... jusqu'à avoir isolé chaque message individuellement.
5. Le code obtenu se lit directement dans le tableau, simplement de gauche à droite.

1.2.3.2 Méthode de Huffman

La procédure est la suivante §4.46 :

1. On classe les événements dans l'ordre des probabilités décroissantes (idem Shannon-Fano).
2. A chaque étape, on attribue aux 2 derniers messages m_{M-1} et m_M (donc les 2 moins probables) un symbole 0 et un symbole 1 respectivement.
3. On construit à partir d'eux un super-message (texte $m_{M-1}m_M$) de probabilité somme $p_{M-1} + p_M$ que l'on reclasse dans un nouveau tableau en l'intercalant pour respecter toujours l'ordre de probabilités décroissantes.
4. On continue jusqu'à ce qu'il ne reste que 2 messages isolés.
5. On revient en arrière dans le tableau pour le codage (lecture du code de droite à gauche).

1.3 Histogramme

L'histogramme est la courbe $n(i)$ où n est le nombre de pixels d'intensité i . Il peut comporter plusieurs pics ou modes (Fig. 1.2), ou bien être uniforme.

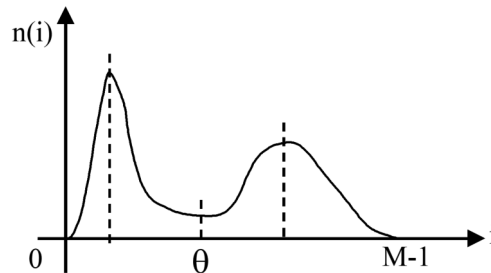


FIGURE 1.2 – Exemple d'histogramme comportant deux modes, avec un seuil θ positionné en un minimum local.

Si l'on note $N = L \times C$ le nombre total de pixels de l'image, et M le nombre de niveaux de gris possibles ($i = 0 \dots M - 1$), la probabilité d'un niveau de gris i vaut $p_i = \frac{n_i}{N}$, et la fonction de densité de probabilité cumulée (CPDF) s'exprime :

$$F(i) = \sum_{k=0}^i p_k \quad \text{avec } F(0) = p_0 \quad \text{et } F(M - 1) = 1. \quad (1.8)$$

Les deux principales corrections d'histogramme sont l'égalisation et l'étalement.

L'**égalisation** consiste à redistribuer les niveaux de gris en remplaçant le niveau i par le niveau l_i tel que **§4.3** :

$$l_i = (M - 1)F(i) = \frac{M - 1}{L \times C} \sum_{k=0}^i n_k \quad (1.9)$$

La Fig. 1.3 illustre l'effet typique d'une égalisation d'histogramme, réalisée grâce au logiciel libre éditeur d'images GIMP (GNU Image Manipulation Programme) sur une photo prise en sous-exposition (personnages dans l'ombre).



FIGURE 1.3 – Image originale **Vernet**, et image égalisée.

L'**étalement** consiste à utiliser toute la dynamique des NdG **§4.4** :

$$l_i = (M - 1) \frac{i - i_{min}}{i_{max} - i_{min}} \quad (1.10)$$

où i_{min} et i_{max} sont les plus petit et plus grand NdG présents dans l'image. L'intérêt de ces deux techniques est d'être complètement automatiques, donc facilement programmables.

Dans le cas général, une correction d'histogramme consiste à appliquer une transformation non linéaire T sur les niveaux de gris **§4.2** :

$$l_i = T(i) \quad (1.11)$$

T doit être une fonction monotone croissante et bornée dans l'intervalle $[0, M - 1]$ si l'on veut conserver la gradation des nuances du sombre au clair. Cela permet amélioration ou trucage, en appliquant un facteur d'amplification non linéaire sur l'échelle de gris (Fig. 1.4).

La Fig. 1.5 présente en haut à gauche une image de visage avec l'histogramme correspondant, et illustre les variations réalisables par modification de cet histogramme avec diverses transformations T linéaires par morceaux.

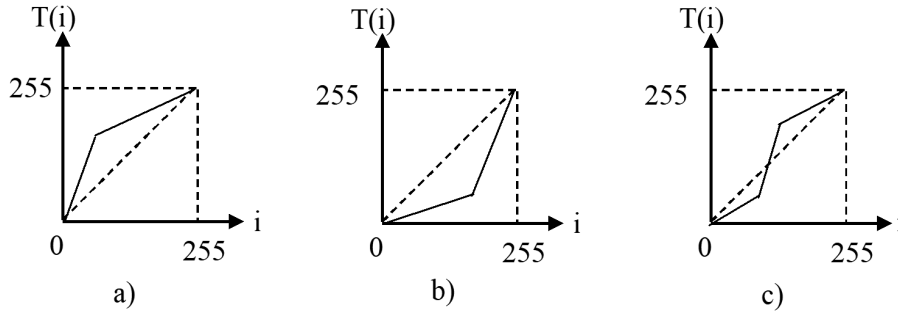


FIGURE 1.4 – Exemple de correction d’histogramme par rehaussement : a) du sombre ; b) du clair ; c) central.



FIGURE 1.5 – Image de visage, et variations obtenues par transformations de son histogramme.

1.4 Seuillage

Le seuillage (sur l’histogramme des NdG, sur leurs dérivées, sur l’entropie, sur les transformées ...) est une opération très fréquente (pour ne pas dire systématique) en traitement d’image (nombreuses applications en détection de

contours, détection de mouvement ...) Il permet de passer d'une image en NdG à une image Noir et Blanc (N&B) (étiquetage binaire $e(s)$) :

$$\forall s = (x, y) \in S, \quad e(s) = \begin{cases} \text{"1"} & \text{si } I(s) > \theta \\ \text{"0"} & \text{sinon} \end{cases} \quad (1.12)$$

La difficulté est évidemment le choix du seuil adéquat θ . Sur un histogramme, on positionne souvent le seuil en un minimum local entre deux maximums (Fig. 1.2). Il existe de nombreuses méthodes et techniques adaptatives, semi-automatiques, ... de sélection du seuil [110, 97]. L'intérêt est de pouvoir ensuite appliquer des outils spécifiques aux images binaires (morphologie mathématique binaire, suivi de contours binaires) **§4.60**.

1.5 Contraste

Il existe différentes définitions du contraste dans une image, mais c'est toujours de la forme $\Delta L/L$, pour exprimer un rapport relatif contextuel (écart local) de luminance. On note ici L la luminance d'un pixel (*i.e.* son NdG), et L_{moy} , L_{min} , L_{Max} les valeurs moyenne, minimale et maximale de la luminance sur une zone donnée **§4.35**.

1.5.1 Contraste de Michelson

- c'est à l'origine une mesure de visibilité de franges d'interférence,
- Michelson considère le cas d'une luminance L sinusoïdale,
- l'estimateur de contraste vaut : $C_M = \frac{L_{Max} - L_{min}}{L_{Max} + L_{min}}$
- il est utilisé avec succès en psychophysique,
- il existe des variantes : $C_M = \frac{L_{Max} - L}{L_{Max} + L}$ et $C_M = \frac{L - L_{min}}{L + L_{min}}$

1.5.2 Contraste de Weber

- Weber considère le cas d'un fond uniforme,
- l'estimateur est : $C_W = \frac{L - L_{moy}}{L_{moy}}$
- il est utilisé en psychophysique (CIE, météorologie).

N.B. : La courbe expérimentale de seuil de visibilité de Weber est largement utilisée.

1.5.3 Contraste de Gordon

- Gordon [63] considère le cas d'un stimulus complexe (image naturelle de mammographie pour diagnostic de cancer du sein),