

Chapitre V : **Traitement d'images**

Introduction

Une image est une source d'informations. L'analyse et le traitement d'images visent donc l'extraction d'informations utiles, à partir d'une scène, pour une application/un objectif bien défini(e).

Le traitement d'images trouve son application dans différents domaines tels que :

- **Contrôle de présence / absence.** Sur des chaînes de production, on vérifie en bout de chaîne avec une caméra vidéo la présence d'une pièce dans un ensemble plus complexe. Pour cela bien souvent il suffit de faire un simple seuillage dans une région spécifique.
- **Contrôle du niveau de maturation des fruits sur une chaîne de conditionnement.** Il s'agit de reconnaître à la couleur et à la texture du fruit son degré de maturité et donc la catégorie sous laquelle il sera emballé puis vendu.
- **Construction et correction de cartes géographiques d'après des images satellites ou des images aériennes.** On recalcule d'après des informations topographiques les images reçues, puis on les met sur la carte en correspondance avec les informations trouvées dans l'image : voies de communication, voies et plans d'eau, parcelles agricoles...
- **Surveillance et évaluation de la production agricole.** Il est possible de déterminer le degré de maturation des cultures, la quantité d'eau nécessaire pour l'irrigation, le rendement moyen... On peut ainsi établir des prévisions à large échelle de la récolte à venir.
- **Reconnaissance de l'écriture.** La reconnaissance de l'écriture manuscrite progresse de jour en jour. Elle est suffisamment opérationnelle pour que la majorité des adresses, même manuscrites, soient reconnues automatiquement sur le courrier postal.
- **Recherche d'image par le contenu.** L'objectif de cette technique est de rechercher, parmi une base de données d'images, les images similaires à une image exemple, ou ayant certaines caractéristiques, par exemple rechercher toutes les images comportant un vélo.
- **Analyse de la vidéo.** L'objectif de cette technique devenue une discipline depuis les années 2000 (lorsque la puissance des processeurs peu onéreux et en particulier des PC a permis des traitements puissants en temps réel) est d'interpréter les faits observés à l'image afin de signaler ou d'enregistrer des faits marquants. Le plus souvent, la caméra est fixe et observe les mouvements d'une scène. Les applications sont nombreuses : Protection des biens (détection d'intrusion, détection d'objet abandonné ou déposé...), Identification (biométrie faciale), Sécurité des personnes (détection de chutes de personnes, franchissement de rambardes, ...), Animations (planchers animés selon les mouvements des danseurs en boîte de nuit), Détection de feux (industriel, forêts, tunnels, ...), Surveillance de tunnels (comptage, mesure de vitesse, détection de fuites/anomalies dans les plafonds), surveillance de tuyaux et autres process industriels.

I. Historique du traitement d'images

Le traitement d'images commence à être étudié dans les années 1920 pour la transmission d'images par le câble sous-marin allant de New York à Londres. Harry G. Bartholomew et Maynard D. McFarlane effectuent la première numérisation d'image avec compression de données pour envoyer des fax de Londres à New York. Le temps de transfert passe ainsi de plus d'une semaine à moins de trois heures. Il n'y a pas vraiment d'évolution par la suite jusqu'à la période d'après-guerre.

Le traitement du signal prend de l'importance vers la fin de la Seconde Guerre mondiale avec l'arrivée du radar. La prospection pétrolière participe aussi beaucoup au développement des techniques de traitement du signal.

Le véritable essor du traitement d'images n'a lieu que dans les années 1960 quand les ordinateurs commencent à être suffisamment puissants pour travailler sur des images. Peu après, la redécouverte de la transformée de Fourier rapide (FFT) révolutionne le domaine, en rendant possible les manipulations du contenu fréquentiel des signaux sur ordinateur.

Cependant, l'essentiel des recherches porte encore, à cette époque, sur l'amélioration des images et leur compression.

En 1980, David Marr formalise la détection de contours de manière précise (D. Marr et E. Hildreth : *Theory of Edge Detection*, Proc. R. Soc. London, B 207, 187-217, 1980).

Au cours des années 1980, un véritable engouement se fait pour le traitement de l'image et surtout pour la compréhension de l'image par des systèmes experts. Les ambitions étaient beaucoup trop grandes, l'échec fut d'autant plus cuisant.

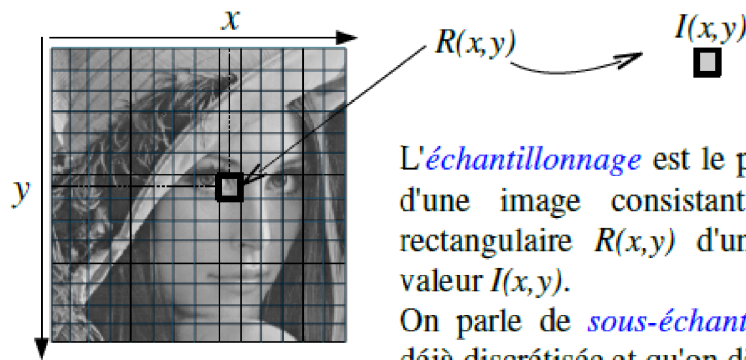
Les années 1990 sont témoin de l'amélioration constante des opérateurs. La recherche médicale devient un très gros demandeur en traitement d'images pour améliorer les diagnostics faits à partir des nombreuses techniques d'imagerie médicale, la technique reine étant l'IRM. Les publicitaires, puis le grand public se familiarisent avec la retouche d'image grâce au logiciel Photoshop, et le traitement d'images dans un objectif esthétique se répand avec l'apparition d'autres logiciels dédiés (The Gimp, Paint Shop Pro). Enfin, la décennie s'achève sur l'engouement pour les ondelettes et les images multimodales.

II. Représentation numérique d'une image

Tout d'abord il faut distinguer les images analogiques des images numériques.

Une image numérique est représentée par un ensemble de valeurs. Si on possède l'ensemble de ces valeurs, il est possible de la reproduire à l'identique. En ce qui concerne les images analogiques, elles existent grâce à un support matériel (exemple une toile pour une peinture) et ne sont pas reproductibles à l'identique. Pour les images numériques, il s'agit de valeurs discrètes contrairement aux images analogiques dont les valeurs sont continues. Nous allons nous intéresser uniquement aux images numériques.

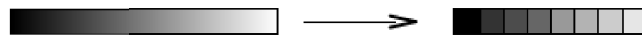
La numérisation d'une image analogique permet la génération de son équivalent numérique. Ceci nécessite l'application de deux opérations fondamentales : l'échantillonnage et la quantification.



L'*échantillonnage* est 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 $I(x,y)$.

On parle de *sous-échantillonnage* lorsque l'image est déjà discrétisée et qu'on diminue le nombre d'échantillons.

La *quantification* désigne la limitation du nombre de valeurs différentes que peut prendre $I(x,y)$.



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

Les résultats de l'*échantillonnage* sont sensibles à la **résolution spatiale** comme le montre la figure ci-dessous.

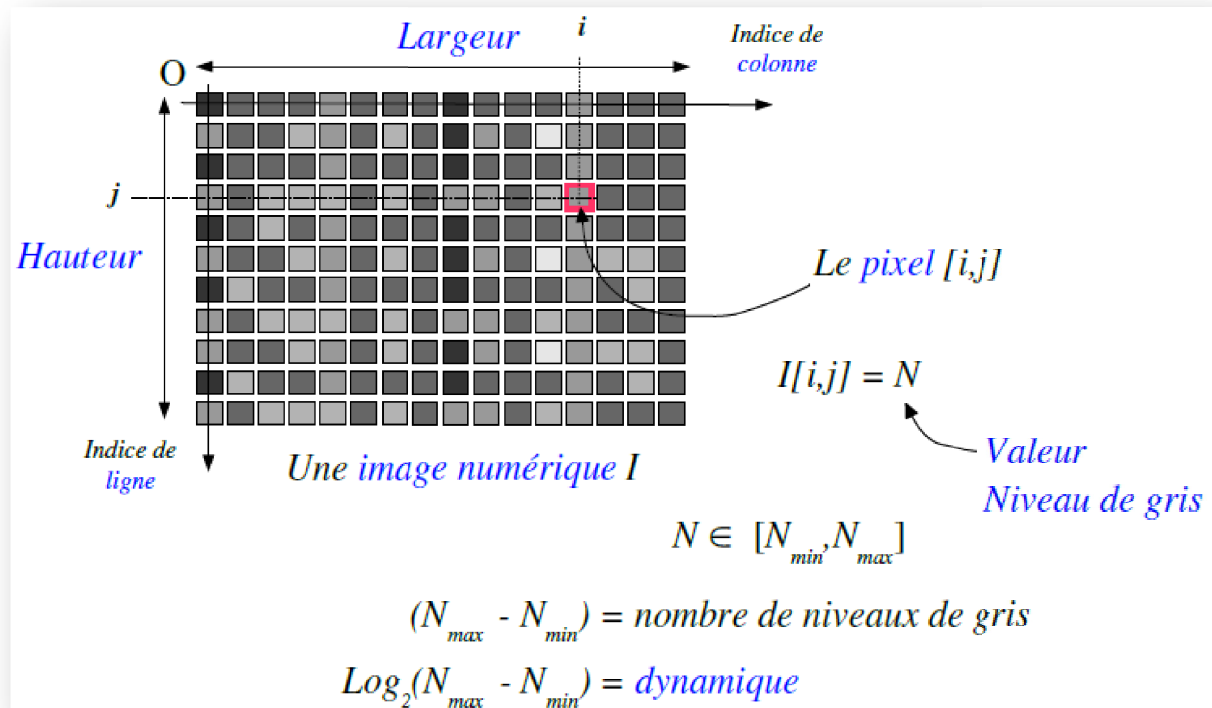


Les résultats de la *quantification* sont sensibles à la **résolution tonale** comme le montre la figure ci-dessous.



Les images sont composées d'un nombre n_i de pixels auxquels est associé une couleur ou une nuance de gris. La résolution d'une image, quant à elle, est liée au nombre de pixels qui la constituent. Plus il y aura de pixels, plus la résolution sera

bonne. Plus il y aura pixels plus on pourra zoomer également (dans la limite de la résolution de l'écran sur lequel nous affichons l'image).



Ce que nous appelons couramment une image en noir et blanc est en fait une image en nuances /niveaux de gris (256 niveaux avec 0 pour le noir et 255 pour le blanc).

Une image en noir et blanc est une sous-catégorie des images en nuances de gris avec seulement 2 nuances différentes (0 pour le blanc et 1 pour le noir). Autrement dit, pour les images en noir et blanc les pixels ne peuvent prendre que deux valeurs, c'est une image binaire. De plus, plus il y a de nuances de gris, plus la qualité de l'image est bonne. On peut le constater avec l'exemple ci-dessous. L'image a été traitée par des fonctions que nous présenterons durant les séances de travaux pratiques.

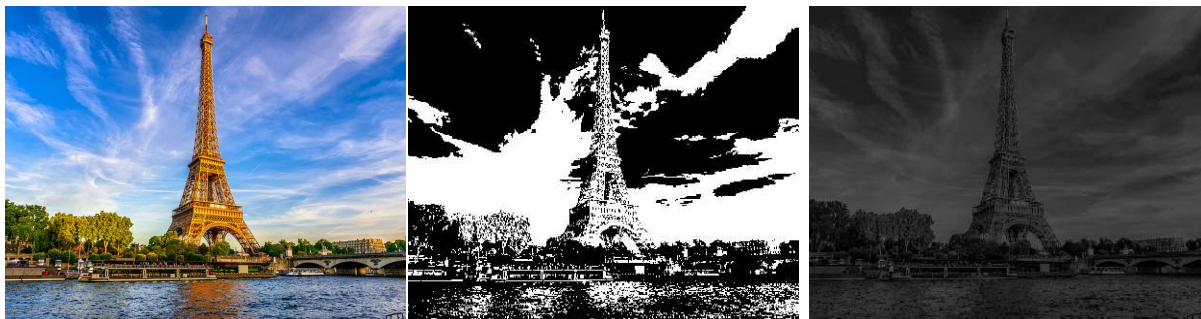


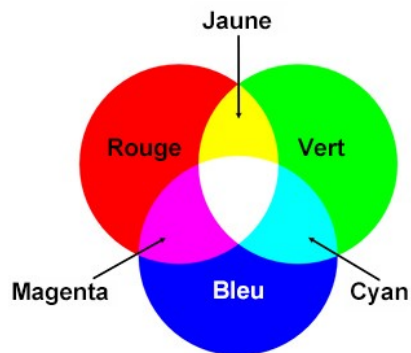
Image couleur / noir et blanc / nuances de gris

Pour les images en couleur il y a en fait 3 couleurs pour chaque pixel. Il existe 256 nuances pour chacune de ces 3 couleurs. Cela fait plus de 16 millions de couleurs

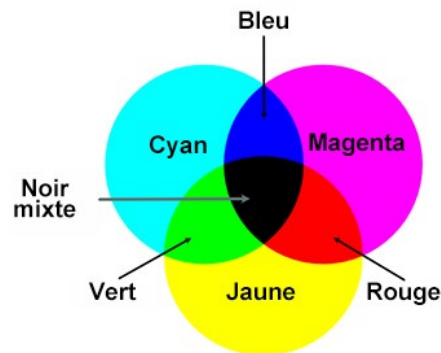
différentes. Une image numérique est représentée par un tableau ou une matrice contenant les différentes valeurs de chaque pixel. Pour les images en couleur il s'agit d'une matrice 3D de dimensions $(m,n,3)$, il y a donc une matrice pour chacune des 3 couleurs citées précédemment.

Exemples de bases couleurs :

- RGB (Red Green Blue): Rouge, Vert et Bleu
- CMY (Cyan Magenta Yellow) : Cyan, magenta et jaune



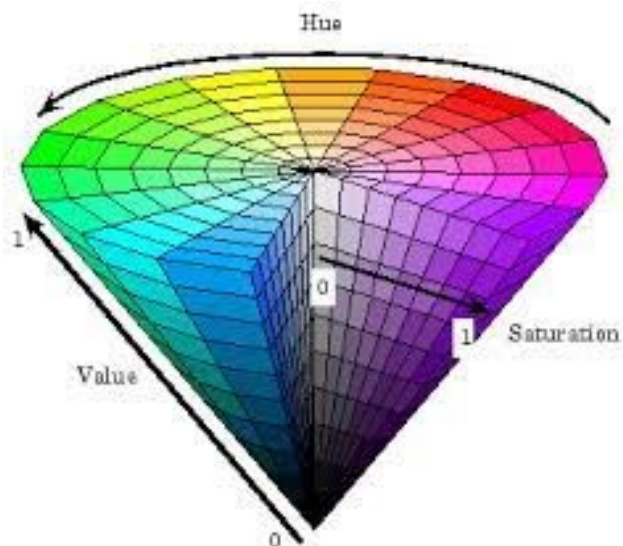
Couleurs additives



Couleurs soustractives

- HSV (Hue Saturation Value) : Teinte, saturation et valeur

La teinte est mesurée par un angle et est associée à la longueur d'ondes, la saturation est mesurée par un pourcentage indiquant la saturation ou le degré de blanc de la couleur et la valeur par un pourcentage indiquant la luminance (clair ou sombre).



- NTSC (National Television Standards Committee USA) : Le NTSC est le premier codage couleur électronique de télévision développé dans le monde. Chaque couleur est codée sur trois composantes traduisant les doses du rouge, du vert et du bleu.

III. Formats d'images

Une image est soit matricielle (Bitmap), soit vectorielle.

1. Images Bitmap (image matricielle, constituée d'une matrice de points colorés)

a. TIFF (Tagged Image File Format)

Ce format est utilisé pour l'impression et a l'avantage de ne pas être compressé et d'accepter les zones de transparence. Cela permet de n'avoir aucune perte de qualité, ce qui lui donne un avantage pour la réalisation d'un imprimé professionnel. Par contre, les fichiers réalisés dans ce format sont plus lourds, donc ne permettent pas une utilisation dans un site web.

b. JPEG (Joint Photographic Experts Group)

Lors de l'enregistrement de l'image dans ce format, une compression plus ou moins forte est effectuée afin de limiter le poids du fichier final. De ce fait, bien qu'il soit possible d'utiliser un mode colorimétrique adapté, son utilisation est très déconseillée pour la réalisation d'imprimés. Le fait de pouvoir définir le taux de compression lors de l'enregistrement a l'avantage de pouvoir créer des fichiers très légers, particulièrement utiles pour la réalisation de sites internet. Le format JPEG ne supportant pas la transparence, il faudra être vigilant lors de son utilisation.

c. PNG (Portable Network Graphics)

PNG est un format d'image destiné uniquement à l'affichage sur écrans, car il ne supporte que le mode colorimétrique RVB. La très bonne gestion de la transparence par ce format en faisait un standard pour la réalisation de sites internet avant que ceux-ci ne prennent en compte les formats vectoriels. Il existe 2 variantes, le PNG-8 et le PNG-24, le dernier étant plus utilisé car sa profondeur de 24 bits permet l'affichage de 16777216 couleurs, contre 256 pour le premier, ainsi qu'une résolution supérieure de la transparence. Il utilise également un haut taux de compression, sans perte de qualité.

d. GIF (Graphics Interchange Format)

Le format GIF est très proche du format PNG-8, dans le sens où il permet l'affichage de 256 couleurs ainsi qu'une transparence de base. Il utilise également le mode colorimétrique RVB, ce qui limite son utilisation à l'affichage sur écran. Son atout majeur est qu'il supporte les animations images par images, par exemple pour animer certaines parties d'un site web. C'est également ce qui en fait un allié incontournable des créateurs de MEMEs.

2. Images Vectorielles (Constituées de vecteurs, où les formes sont calculées à partir de points d'ancrages)

a. EPS (Encapsulated PostScript)

Ce format est beaucoup utilisé pour l'impression et étant un format vectoriel permet un agrandissement de l'image sans limite. Ce format est très fréquent lors de la création de logos ou d'images illustrées. Dans le web, on lui préférera un autre format vectoriel, le SVG.

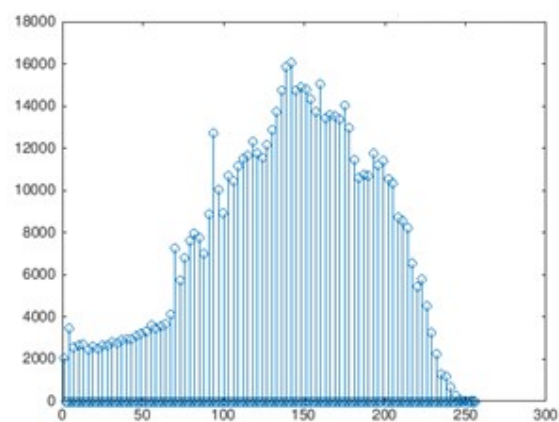
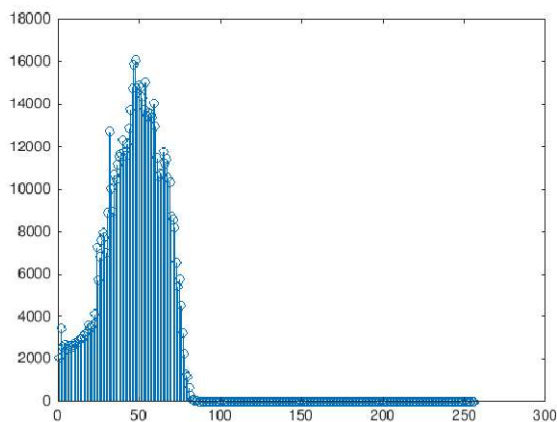
b. SVG (Scalable Vector Graphics)

Ce format est couramment utilisé lors de la création de sites internet car il permet l'ajout d'éléments (logo, illustration, graphique) dans les pages web en tant qu'image (fixe) ou directement en tant que code intégré à la structure HTML (permettant ainsi des petites animations). Ce format est particulièrement apprécié car les fichiers sont légers et permettent un agrandissement sans perte.

Il existe bien sûr d'autres formats d'images, moins répandus car utilisés plutôt par les professionnels, ou n'étant pas supportés par tous les systèmes d'exploitation. On entendra parler de PSD, BMP, DXF, Targa, Scitex CT, Dicom, etc.

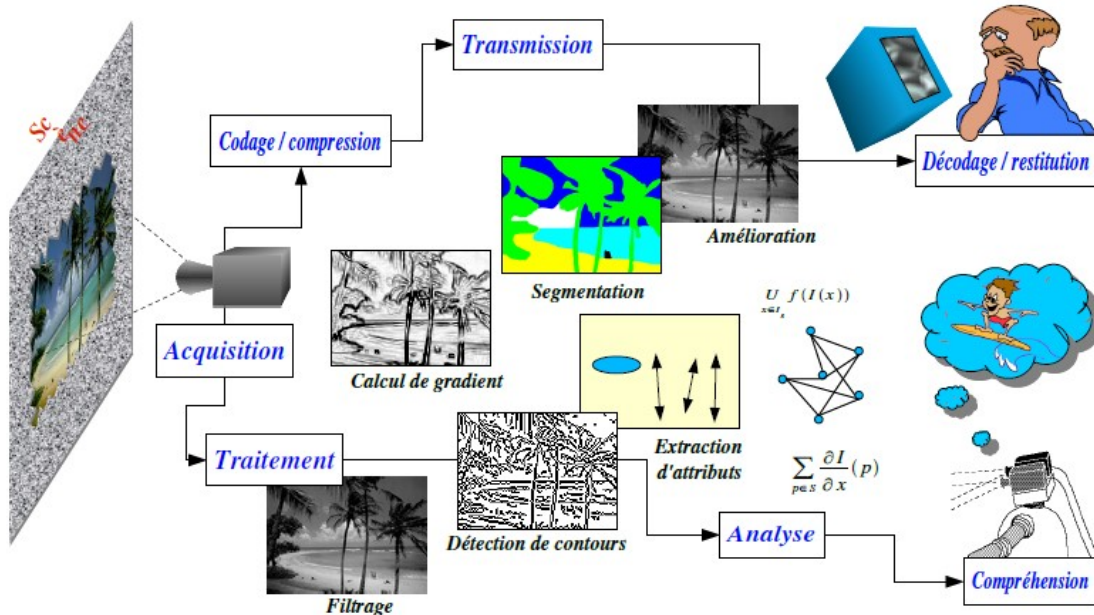
IV. Histogramme

L'histogramme d'une image à niveaux de gris est un outil de description permettant d'afficher, pour l'image considérée, d'un diagramme dont l'abscisse illustre tous les niveaux de gris L_i contenus dans l'image et en ordonnée le nombre de pixels ayant le niveau L_i .



V. Outils de traitement et d'analyse d'images

L'acquisition d'images à partir d'une scène réelle permet de générer un ensemble d'images numériques sur lesquelles plusieurs opérations peuvent être exécutées afin d'en interpréter le contenu.



Parmi les principales opérations souvent appliquées aux images numériques, on retrouve la convolution et la transformée de Fourier discrète.

1. Convolution

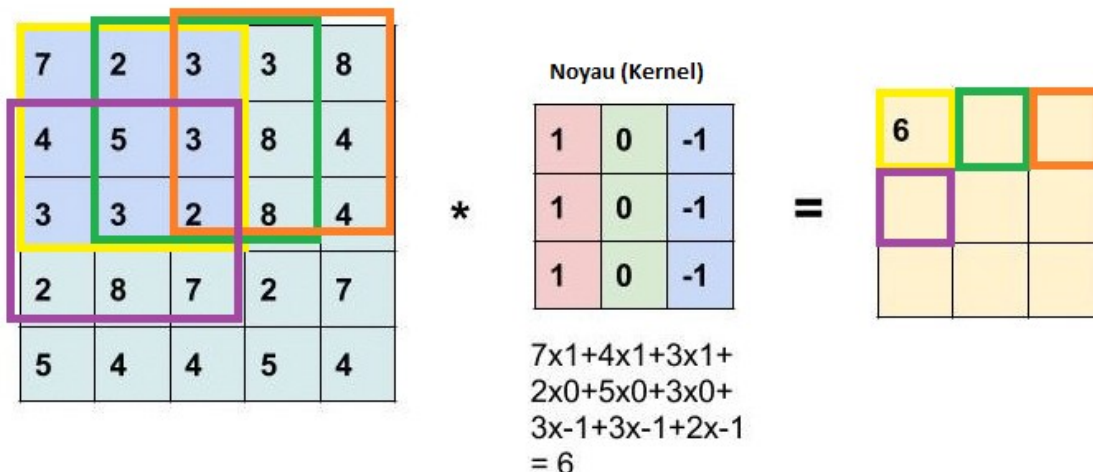
Soit I une image numérique.

Soit h une fonction de $[x_1, x_2] \times [y_1, y_2]$ à valeurs réelles

La convolution de I par h est définie par :

$$(I * h)[x, y] = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} h[i, j] \cdot I[x - i, y - j]$$

La fonction h est dite le noyau de la convolution.

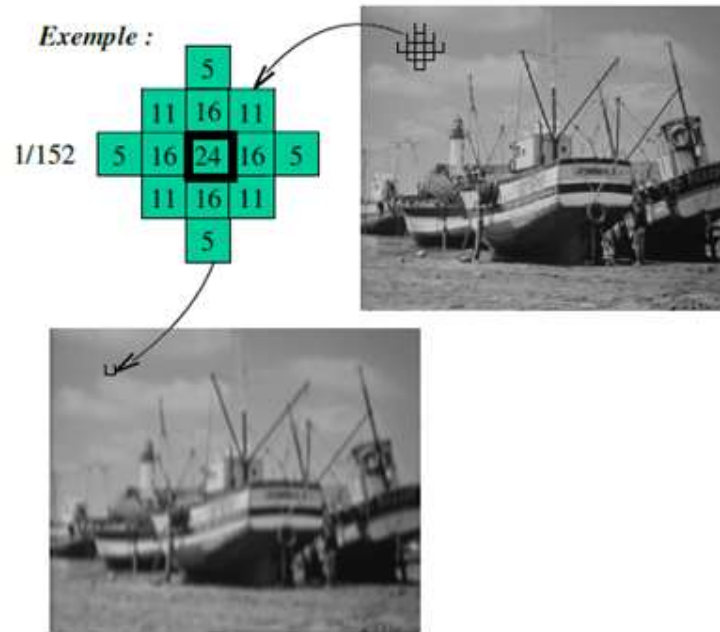


Propriétés

Commutativité: $h * g = g * h$

Associativité : $(h * g) * k = h * (g * k) = h * g * k$

Distributivité : $h * (g + k) = (h * g) + (h * k)$



2. Transformée de Fourier

La transformée de Fourier permet la décomposition d'un signal f en combinaison linéaire de sinusoides complexes, dont les coefficients $F[u,v]$ dit coefficients de Fourier, fournissent des informations sur les fréquences (u,v) et permettent des manipulations dans le domaine fréquentiel.

Transformée de Fourier discrète bidimensionnelle :

Si (x,y) sont les coordonnées du domaine spatial et (u,v) sont celles du domaine fréquentiel, la transformée de Fourier directe est définie par :

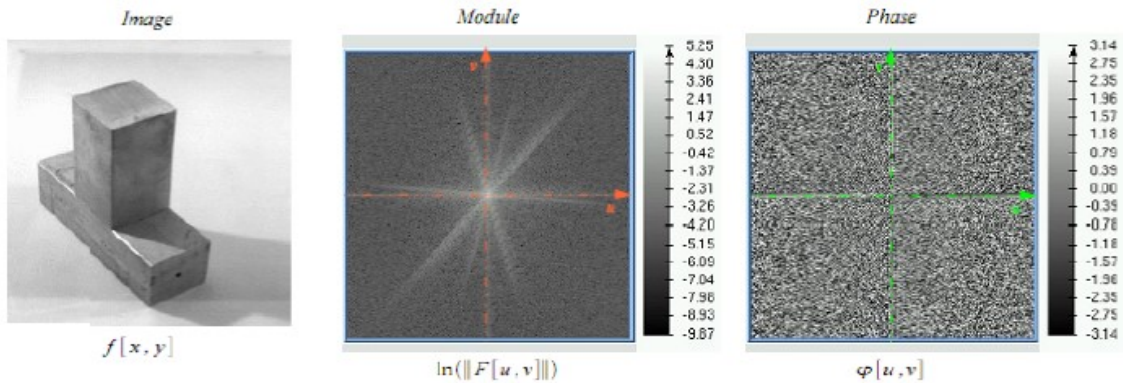
$$F[u,v] = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} f[x,y] \cdot e^{-2i\pi(ux+vy)/wh}$$

La transformée inverse est définie par :

$$f[x,y] = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} F[u,v] \cdot e^{2i\pi(ux+vy)/wh}$$



la transformée de Fourier d'une image donnée est souvent caractérisée par son module et sa phase.



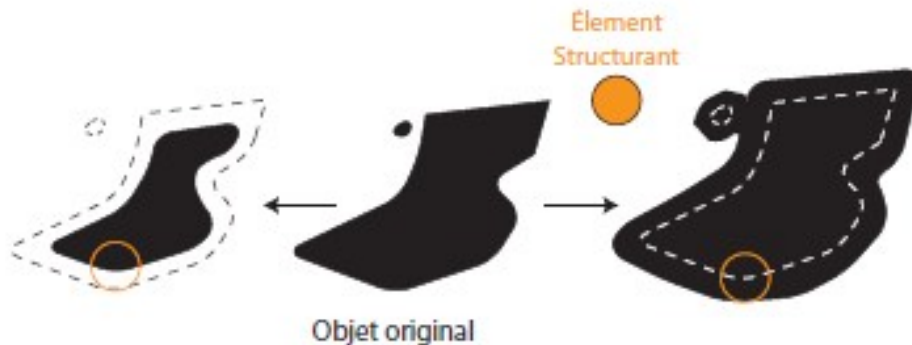
3. Opérateurs morphologiques

Le domaine de la morphologie mathématique est issu de la théorie des sets visant à filtrer une image donnée.

3.1 Erosion et Dilatation

L'érosion (en MATLAB, imerode) et la dilatation (imdilate) sont deux techniques de base de la morphologie mathématique.

Ce groupe d'opérateurs se base sur le filtrage des hautes et basses valeurs de l'image I en appliquant un élément structurant B (filtre). L'image érodée/ dilatée par l'élément structurant B est l'ensemble des pixels $x \in I$ (souvent noté I_B) tel que B soit entièrement contenu (érosion) ou tel que le centre de l'élément soit contenu (dilatation) dans l'objet considéré. Le pixel filtré x prend la valeur booléenne 1 si l'élément est entièrement contenu (ou partiellement pour la dilatation) ou 0 sinon, pour le cas de base où on considère des images en 1 bit (binaires).



Erosion (gauche) et dilatation (droite) de l'objet binaire en utilisant un élément structurant B circulaire

Plus précisément érosion $\epsilon_B(I)$ et dilatation $\delta_B(I)$ sont définis (pour une image binaire en noir et blanc) comme:

$$\begin{aligned}\epsilon_B(I) &= I \ominus B, \\ \delta_B(I) &= I \oplus B.\end{aligned}$$



Erosion (gauche) et dilatation (droite) des objets originaux (centre) en utilisant un élément structurant B circulaire de rayon 2 (lx) et 5 (dx) pixels.

L'extension de ces techniques aux cas de l'image à échelle de gris, nécessite une adaptation par rapport à l'assignation des nouvelles valeurs de l'image filtrée. Si avant il s'agissait d'opérations booléennes, ces opérations deviennent arithmétiques assignant la valeur maximale (dilatation) ou minimale (érosion) contenue dans la surface de l'élément structurant au pixel filtré x sur lequel B est centré. On définit l'érosion $\epsilon_B(I)$ et dilatation $\delta_B(I)$, pour une image à échelle de gris, comme suit:

$$\begin{aligned}\epsilon_B(I) &= \inf(I_B), \\ \delta_B(I) &= \sup(I_B).\end{aligned}$$



Erosions (vers gauche) et dilations (vers droite) de l'image originale (centre) en utilisant un élément structurant B circulaire de rayon [2, 4, 6, 8] pixels. (Image exemple de MATLAB, 'cameraman.tif')

D'un point de vue purement pratique, l'érosion (en prenant le minimum de valeurs de gris contenu dans l'élément B) efface les éléments clairs contenus dans la taille de l'élément. Au contraire, la dilatation efface les éléments sombres de l'image.

(Rappel: plus une valeur est grande, plus elle tend vers le blanc!)

Erosion et Dilatation sont les deux opérations de base dans le domaine de la morphologie mathématique. Par la suite, on verra comment la combinaison ou l'itération de ces éléments nous amènent à obtenir d'autres filtres et opérateurs très utilisés lors de l'extraction d'information structurale d'une image ou bien lors d'un lissage, ou encore pour éliminer des objets dans l'image.

Dans MATLAB ces deux opérations sont contenues dans la toolbox de traitement d'images :

IM2 = imdilate(I, SE); % I l'originale; SE l'élément structurant B (SE = strel (forme,size))

IM2 = imdilate(I, V); % I l'image originale; V matrice binaire spécifiant le voisinage de B

IM2 = imerode(I, SE || V); % fonctionne de même manière.

3.2 Ouverture et Fermeture

Les opérateurs ouverture et fermeture sont une combinaison des opérateurs élémentaires qu'on a vus dans la section précédente appliqués en série, en respectant un certain ordre. La définition de ouverture ($\gamma_B(I)$) et fermeture ($\phi_B(I)$), par un élément structurant B, est donnée par la combinaison de dilatation et érosion, comme suit:

$$\begin{aligned}\gamma_B(I) &= \delta_B(\epsilon_B(I)) = (I \ominus B) \oplus B = I \odot B, \\ \phi_B(I) &= \epsilon_B(\delta_B(I)) = (I \oplus B) \ominus B = I \oslash B.\end{aligned}$$

Pratiquement, ces opérateurs appliquent un filtrage externe ou interne à l'objet. Plus l'élément structurant B est grand, plus la géométrie de l'objet (en termes d'angles et formes) sera uniforme.

Un objet qui présente une géométrie régulière sera peu filtré par rapport aux autres. Plus B est grand, plus les objets clairs ou sombres seront filtrés sur la base de leur tonalité et dimension. Avec ces types d'opérateurs, la forme de B (le type de voisinage à considérer) est aussi très influente.

Ces deux opérateurs exploitent la combinaison de l'érosion et de la dilatation (dans un ordre défini), pour filtrer (et non éliminer comme on a vu auparavant) des parties sombres ou claires de l'image.



Ouvertures (vers gauche) et fermetures (vers droite) de l'image originale (centre) en utilisant un élément structurant B circulaire de rayon [2, 4, 6, 8] pixels.

Ces deux opérations sont présentes en tant que fonctions indépendantes en MATLAB, qu'on exploitera lors du TP. A titre informatif, les fonctions sont:

IM2 = imopen(IM,SE) ; % IM image, SE él. struct. B, IM2 image filtrée

IM2 = imopen(IM,NHOOD); % IM image, NHOOD voisinage de B, IM2 image filtrée.

IM2 = imclose(IM, SE || NHOOD); % fonctionne comme la précédente.