

Mohammed Merzougui et M'barek Nasri

Contrôle qualité par segmentation évolutionniste d'images



Notations

<i>AE</i>	Algorithme Évolutionniste
<i>AG</i>	Algorithme Génétique
<i>CR</i>	Croissance de Régions
<i>CRE</i>	Croissance de Régions Évolutionniste
<i>ISODATA</i>	Algorithme ISODATA
<i>KM</i>	Kmeans
<i>EKM</i>	Kmeans Évolutionniste
<i>SE</i>	Stratégie d'Évolution
<i>SE_XB</i>	Critère de Xie et Beni évolutionniste
V_{MENP}	Principe du maximum d'entropie
V_{XB}	Critère de Xie et Beni

Avant-propos

Les journalistes étaient les premiers utilisateurs d'images numériques. A l'époque, il n'était même pas question d'ordinateurs. Il s'agissait de trouver des méthodes pour transmettre le plus rapidement possible des photos d'un continent à un autre. Le système Bartlane mis au point par Harry G. Batholomew et Maynard D. McFarlane et développé en Grande Bretagne dans les années 1920 fut un des premiers procédés de numérisation et transmission d'images. L'information contenue dans l'image était au départ transcrite sur une carte perforée composée de 5 types de trous différents, d'où des images codées en 5 niveaux de gris. Ces images étaient ensuite transmises télégraphiquement par câble sous-marin puis retraduites et imprimées à l'autre bout de la ligne. Cette technique permettait de transmettre des photos en moins de 3 heures alors qu'une semaine était auparavant nécessaire.

En 1925, de nouvelles améliorations rendaient possible la transmission d'images en 15 niveaux de gris. Bien que ces images soient le résultat d'une sorte de processus de numérisation, elles ne rentrent pas dans le cadre de ce qu'on appelle classiquement les images numériques car leur fabrication n'a pas été réalisée à l'aide d'un ordinateur. L'histoire des images numériques proprement dites est donc intimement liée à l'histoire et l'évolution des ordinateurs. Si les principes d'automatisation de calculs remontent à l'Antiquité (avec notamment l'Abacus, boulier utilisé à Babylone), et si de nombreuses machines à calculer mécaniques ont été inventées au cours des siècles, la notion moderne d'ordinateur remonte aux années 1940. Elle repose sur deux concepts clés amenés par Von Neumann : une mémoire pour stocker les programmes et les données et la notion de branchements

conditionnels. Cependant, il faudra attendre les années 1960 pour disposer de machines suffisamment puissantes permettant de traiter informatiquement des images. Le programme spatial de la NASA à l'époque a beaucoup contribué à la création et à l'évolution du traitement d'images par l'informatique. Il s'agissait d'être capable d'améliorer et d'interpréter les images ramenées par les premières sondes spatiales (Ranger 7 et suivantes). La fin des années 1960 et le début des années 1970 voient ensuite la naissance de l'imagerie médicale avec l'invention de la tomographie.

Depuis, l'utilisation d'images numériques s'est généralisée à de nombreux domaines et les progrès réalisés en informatique permettent de manipuler des images de plus en plus complexes.

A notre époque, et surtout la dernière décennie, un grand nombre d'articles est publié chaque année dans le domaine de la recherche de traitement d'images. En effet, la montée en puissance et en capacité de stockage des ordinateurs ainsi que la mise à portée des systèmes d'acquisitions numériques de haute précision ont contribué au développement des systèmes d'aide à la décision dans le domaine de traitement d'images. L'objectif final consiste à obtenir une interprétation de l'image au niveau sémantique répondant aux différentes attentes des utilisateurs humains. Pour cela, de nombreux travaux ont été consacrés au développement de chaînes de traitement d'images basées essentiellement sur la segmentation. Un des plus grands domaines d'application de la segmentation est la VIAO (Vision Industrielle Assistée par Ordinateur). Bien évidemment, il en existe d'autres parmi lesquels nous pouvons citer la télésurveillance et gestion des alertes, l'imagerie médicale et la télédétection qui est utilisée essentiellement dans les domaines de la météo, la cartographie, l'analyse des ressources terrestres.

Nous nous sommes intéressés dans cet ouvrage à un domaine très important qui est le contrôle qualité où l'image numérique s'est imposée comme un support et une source d'information privilégiée. En production industrielle, le contrôle qualité peut nécessiter la mise en œuvre d'un système de vision artificielle pour la détection des défauts. Ce système de vision offre plus de fiabilité et de rapidité à la décision qu'un opérateur humain.

Vers la fin des années 70, David Marr [MAR 82] a proposé un modèle calculatoire pour le traitement et la représentation de l'information visuelle. En pratique, le paradigme de David Marr se traduit par trois étapes de traitement : segmentation, reconstruction et reconnaissance. Notre attention est focalisée particulièrement sur la première étape.

La segmentation d'images est une étape primordiale en traitement d'images puisqu'elle conditionne la qualité de l'interprétation. Dans le domaine de contrôle qualité, son but est de détecter les défauts de l'objet à contrôler d'une façon automatique. La segmentation automatique d'images est cependant une tâche difficile, du fait de la variabilité des images. Beaucoup de méthodes existantes dans la littérature donnent de bons résultats mais elles nécessitent des connaissances a priori qui sont généralement difficilement accessibles. En outre, les algorithmes de segmentation proposés nécessitent un choix délicat d'opérateurs à utiliser et un calcul compliqué de paramètres.

Afin de rendre la méthode la plus générale possible et qui s'adapte à différentes applications, une chaîne de traitement doit être mise en place. Cette dernière est évidemment composée d'un certain nombre d'opérateurs de traitement d'images choisis soigneusement en fonction de la qualité de l'image et de l'information recherchée dans celle-ci. En effet, la difficulté de choix des opérateurs demeure dans l'adaptation du comportement théorique de ces derniers et leur comportement sur des cas réels afin de les rendre exploitables en routine de contrôle.

Dans cet ouvrage, nous essayons de porter des améliorations à des algorithmes classiques et largement utilisés en classification des données et en segmentation d'images. Notre intérêt est essentiellement réservé aux trois algorithmes bien connus dans la littérature et qui sont : K-means, ISODATA et Croissance de Régions. Le premier souffre de la phase d'initialisation, des optimums locaux et du nombre de classes qui doit être donné a priori. ISODATA est un algorithme pour lequel le réglage de ses paramètres est très difficile surtout les deux seuils pour fusionner deux classes ou pour diviser une seule classe. Quant à l'algorithme de Croissance de Régions, le nombre de germes doit être connu auparavant. Notre contribution est de trouver des solutions aux différents inconvénients mentionnés ci-dessus.

Cet ouvrage est divisé en six chapitres :

Le chapitre 1 concerne un état d'art portant sur l'image numérique et la segmentation d'images et ses différentes méthodes et applications. Les avantages et les inconvénients de ces méthodes sont présentés afin de situer la solution proposée pour chacune d'elles.

Le chapitre 2 est consacré à une présentation assez détaillée des algorithmes évolutionnistes. Nous nous intéressons plus précisément aux deux méthodes, à savoir les algorithmes génétiques et les stratégies d'évolution.

Le chapitre 3 traite le contrôle qualité par vision artificielle, ce qui explique le contexte dans lequel cet ouvrage est inscrit ainsi que les objectifs derrière.

Le chapitre 4 décrit les algorithmes de classification K-means et *ISODATA* ainsi que leur optimisation avec les stratégies d'évolution. Nous détaillons pour chacun les étapes le constituant, sa réalisation et sa validation à travers des simulations.

Le chapitre 5 concerne l'application des algorithmes développés dans le chapitre 4 pour segmenter quelques images de synthèse et réelles. Nous choisissons surtout des images à des fins de contrôle qualité.

Dans le chapitre 6, nous proposons d'utiliser l'algorithme Croissance de Régions *CR*. Nous l'optimisons avec les stratégies d'évolution et nous estimons le nombre de germes au début par le critère de Xie et Beni. Nous évaluons l'algorithme ainsi proposé sur des exemples d'images de synthèse et réelles dans un contexte de contrôle qualité.

Chapitre 1

Concepts et problématiques de la segmentation d'images

1. Introduction

Pour un ordinateur, une image est un ensemble de pixels (points image) caractérisés par des couleurs ou par des niveaux de gris. Elle est vue donc comme un « tout » et non pas comme une composition d'éléments. Par exemple, une image représentant un livre posé sur une table est perçue par un ordinateur comme étant une matrice de pixels. Si on veut mener une analyse sur la partie de l'image associée au livre, on doit pouvoir séparer les pixels qui représentent le livre de ceux représentant la table. C'est un problème de traitement d'image connu sous le nom de segmentation d'images. Cette segmentation est basée sur la détection d'indices visuels primaires. Ce sont des informations sur l'image obtenues directement à partir de calculs sur les pixels et formant les premières structures de données importantes de niveau juste supérieur à celui des pixels [SAL 94]. Les indices visuels primaires les plus utilisés sont les régions (points ayant une luminosité proche) et les contours (points présentant des changements de luminosité remarquables). Ainsi, on distinguera les méthodes de segmentation par détection de régions et / ou de contours.

Le traitement numérique des images trouve des applications dans un nombre croissant de domaines grâce à deux facteurs :

- La grande montée en puissance des ordinateurs, l'amélioration et la

miniaturisation des systèmes d'acquisition d'images et la chute du coût des matériels électroniques ont récemment permis un formidable essor de l'utilisation de l'image numérique ;

- Développement de différentes techniques telles que le traitement du signal, l'intelligence artificielle, la morphologie mathématique, les approches évolutionnistes, la logique floue, la théorie des graphes, etc. Les domaines d'application du traitement d'images sont très vastes, parmi lesquels :

- Le milieu industriel (contrôle qualité et suivi des processus de fabrication) ;
- La télésurveillance et gestion des alertes (l'analyse des séquences d'images) ;
- L'imagerie médicale (diagnostic, localisation et suivis des maladies) ;
- L'imagerie satellitaire (la météorologie, le suivi des récoltes, etc.)

La vision artificielle ou vision assistée par ordinateur (VAO) symbolise toute la chaîne visuelle partant de l'image brute jusqu'à l'interprétation de son contenu. Il existe donc un aspect décisionnel très important lors de la phase de la compréhension, d'où l'interaction avec l'intelligence artificielle.

Le traitement d'images se définit comme un ensemble de tâches destinées à extraire de l'image des informations qualitatives et quantitatives. Ces dernières seront transmises au module d'un niveau supérieur. Le traitement d'image est né de l'idée et de la nécessité de remplacer l'observateur humain par la machine [COC 95]. Les étapes d'une chaîne de vision par ordinateur sont, figure 1.1.



Figure 1.1 Etapes d'un système de vision par ordinateur.

- Étape d'acquisition d'image : c'est la transformation de la réalité (scène ou document) en fichier informatique ;
- Étape de prétraitement : consiste à améliorer la qualité de l'image

- acquise. Elle permet l'élimination de bruit ou de rehausser le contraste ;
- Étape de segmentation : a pour objectif de diviser l'image en zones homogènes appelées régions selon certains critères (intensité, couleur, texture...) [EHR 03]. Cette étape peut être suivie par une étape facultative de post traitement qui consiste à raffiner la segmentation afin d'améliorer les résultats obtenus ;
 - Étape d'interprétation : consiste à extraire l'information sémantique de l'image en identifiant les différents objets qui la constituent.

2. Image numérique

2.1. Qu'est-ce qu'une image numérique ?

Une image est la représentation planaire d'une scène acquise à l'aide de systèmes de production d'images (appareils photographique, caméra, radiographies, scanner, sonar,...). Sa forme peut être analogique (ex : négatif, photographie, vidéo.) ou numérique (images numérisées suivant divers formats (images compressées ou non...) ou obtenues par des capteurs fournissant des images numérisées) et dans ce cas un traitement par ordinateur est possible. Son élaboration résulte de la volonté de proposer une entité observable par l'œil humain. Ceci explique d'une part son aspect planaire et d'autre part le fait que l'information élémentaire associée à chaque point de l'image (pixel) soit transcrite en niveaux de gris ou en couleur [COC 95] [TOD 10].

Une image réelle est obtenue à partir d'un signal continu bidimensionnel pris par un outil d'acquisition. Sous cette forme, l'image est inexploitable par l'ordinateur qui ne connaît que le langage binaire, ce qui nécessite sa numérisation.

La numérisation d'une image analogique est la conversion de celle-ci de son état analogique (distribution continue d'intensités lumineuses), en une image numérique représentée par une matrice bidimensionnelle de valeurs numériques $I(i,j)$ où :

- i,j : coordonnées cartésiennes d'un point de l'image ;
- $I(i,j)$: niveau de gris en ce point.

L'opération de numérisation comprend trois phases :

– **Échantillonnage** : l'échantillonnage commence par découper l'image en surfaces carrées élémentaires d'une matrice carrée ou rectangulaire ; chacun des carrés éléments d'image est appelé *pixel* (Picture Element) et repéré par ses coordonnées x et y .

– **Quantification** : les lignes sont étudiées les unes après les autres et sur chacune, la valeur de chaque pixel est mesurée ; parfois ce pixel a une structure hétérogène ; la valeur retenue est alors une moyenne, des détails seront donc perdus.

– **Codage** : représentation informatique (binaire) des valeurs représentant les pixels.

Ainsi, l'image se présente sous la forme d'une matrice I de M lignes et P colonnes, Chaque élément $I(i, j)$ représente un pixel de l'image. La valeur de ce dernier est associée à l'intensité du point (niveau de gris) : 0 pour le noir et N pour le blanc (souvent $N = 255$) [RUS 07] [RUS 11], figure 1.2.

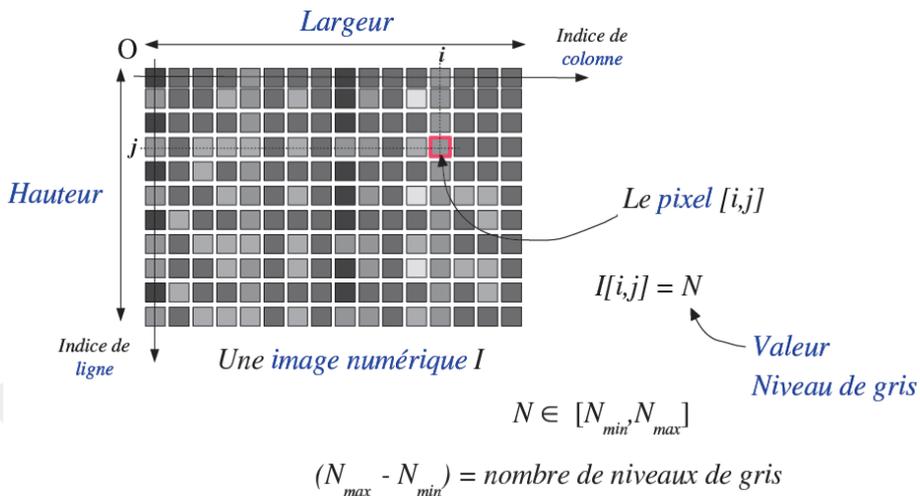


Figure 1.2 Image numérique (ensemble de pixels).

L'acquisition d'image est la transformation de la réalité (scène ou document) en fichier informatique. Elle est faite par des capteurs que l'on appelle plus spécifiquement Capteur à Charges Défiante (Charge Couple Device) CCD. La lumière est projetée sur le CCD qui convertit ces informations en données numériques. Cette étape consiste à prendre des images issues de différentes sources et de les transférer sur ordinateur.

Les scanners et les appareils photos numériques ou caméras vidéo utilisent généralement le principe des capteurs CCD. La qualité de numérisation dépend entre autre de la qualité de résolution de ces appareils. Un capteur de rayons X mesure la quantité de rayonnement transmise par la pièce (contrôle qualité) ou l'organe (imagerie médicale) [KUN 93].

On peut conclure qu'une image est définie par :

- Le nombre de pixels qui la composent en largeur et en hauteur ;
- L'étendue des teintes de gris ou des couleurs que peut prendre chaque pixel (on parle de dynamique de l'image).

A partir d'une image numérique, il convient d'extraire les informations les plus pertinentes en regard de l'application concernée, de les traiter puis de les interpréter. Le terme générique d'analyse d'images désigne l'ensemble de ces opérations. Un codage sur k bits, par exemple, permet de définir 2^k niveaux de gris. Avec $k = 6$, on aboutit à 64 niveaux de gris, ce qui est approximativement le pouvoir de discrimination de l'œil humain. Nous énumérons dans ce qui suit les caractéristiques les plus importants.

2.2. Caractéristiques d'une image numérique

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants :

2.2.1. Définition et résolution

On appelle définition le nombre de points (pixel) constituant l'image, c'est-à-dire sa « dimension informatique » (le nombre de colonnes de l'image que multiplie son nombre de lignes). Une image possédant 640 pixels en largeur et 480 en hauteur aura une définition de 640 pixels par 480, notée 640×480 .

La résolution, terme souvent confondu avec la « définition », détermine par contre le nombre de points par unité de surface, exprimé en *points par pouce* (PPP, en anglais DPI pour *Dots Per Inch*) ; un pouce représentant 2.54 cm . La résolution permet ainsi d'établir le rapport entre le nombre de pixels d'une image et la taille réelle de sa représentation sur un support physique. Une résolution de 300 dpi signifie donc 300 colonnes et 300

rangées de pixels sur un pouce carré ce qui donne donc 90000 pixels sur un pouce carré. La résolution de référence de 72 dpi nous donne un pixel de $1''/72$ (un pouce divisé par 72) soit 0.353mm, correspondant à un *point pica* (unité typographique anglo saxonne).

2.2.2. Niveau de gris

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la « couleur » de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux.

Exemple :

- Si on code l'image avec un bit par pixel on obtient 2 nuances de gris 0 pour le noir et 1 pour le blanc. C'est une image, dite binaire, pour laquelle chaque pixel ne peut avoir pour valeur que 0 ou 1. C'est typiquement le type d'image que l'on utilise pour scanner un texte lorsque celui ci est composé d'une seule couleur [BOV 09] ;
- Si on code l'image avec un codage sur 8 bits, on obtient 256 niveaux de gris. Un pixel de l'image en niveau de gris (ou image d'intensité) est un entier compris entre 0 et 255. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 représente le blanc (intensité lumineuse maximale) [BOV 09].

2.2.3. Image en couleurs

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images en couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (modèle RGB).

L'image couleur est donc la superposition de trois couleurs de base. Il existe plusieurs modes de représentation de la couleur. Le plus utilisé pour le traitement des images numériques est l'espace couleur (*RGB*). On définit, alors un plan de niveau de gris pour chacune des trois couleurs.

2.2.3.1. Représentation en couleurs réelles

Dans les images en niveaux de gris, la valeur d'un pixel est codée sur un seul canal et correspond à son intensité lumineuse. Une image numérique 2D est une approximation de l'image réelle formée par la projection de la scène 3D sur le plan image (plan correspondant à la surface photosensible du capteur ayant acquis l'image). La qualité d'une image numérique dépend de sa résolution (c'est-à-dire du nombre de pixels dans l'image) et de la quantification de son intensité lumineuse, à savoir le nombre de bits utilisé. Pour des images en niveaux de gris, l'intensité d'un pixel est codée sur un octet (8 bits) on a alors $2^8 = 256$ nuances de gris possibles. Pour des images *RVB*, l'intensité lumineuse est codée sur trois octets (24 bits) [VEL 08].

Elle consiste donc à utiliser 24 bits pour chaque point de l'image. Huit bits sont employés pour décrire la composante rouge (R), huit pour le vert (G) et huit pour le bleu (B). Il est ainsi possible de représenter environ 16,7 millions de couleurs différentes simultanément. Cela est cependant théorique, car aucun écran n'est capable d'afficher 16 millions de points. Dans la plus haute résolution (1600 x 1200), l'écran n'affiche que 1 920 000 points. Par ailleurs, l'œil humain n'est pas capable de distinguer autant de couleurs.

2.2.3.2. Représentation en couleurs indexées

Afin de diminuer la charge de travail nécessaire pour manipuler des images en 24 bits, on peut utiliser le mode de représentation en couleurs indexée. Le principe consiste à déterminer le nombre de couleurs différentes utilisées dans l'image, puis à créer une table de ces couleurs en attribuant à chacune une valeur numérique correspondant à sa position dans la table. La table, appelée palette, comporte également la description de chacune des couleurs, sur 24 bits.

2.2.3.3. Autres modèles de représentation

Le modèle *RVB* représentant toutes les couleurs par l'addition de trois composantes fondamentales, n'est pas le seul possible. Il en existe de nombreux autres. L'un d'eux est particulièrement important. Il consiste à séparer les informations de couleurs (chrominance) et les informations d'intensité lumineuse (luminance). Il s'agit du principe employé pour les enregistrements vidéo. La chrominance est représentée par deux valeurs (selon des modèles divers) et la luminance par une valeur.

2.2.4. Histogramme

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image, c'est-à-dire le nombre de pixels pour chaque intensité lumineuse. Par convention un histogramme représente le niveau d'intensité en abscisse en allant du plus foncé (à gauche) au plus clair (à droite). Ainsi, l'histogramme d'une image en 256 niveaux de gris sera représenté par un graphique possédant 256 valeurs en abscisses, et le nombre de pixels de l'image en ordonnées. Prenons par exemple l'image suivante composée de niveaux de gris, figure 1.3 :

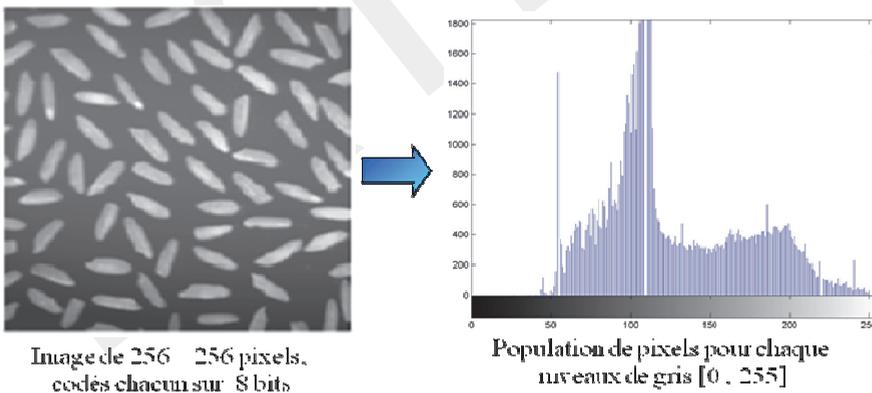


Figure 1.3 Exemple d'image en niveaux de gris.

L'histogramme permet de donner un grand nombre d'informations sur la distribution des niveaux de gris de l'image, ce qui est communément appelé la dynamique de l'image [RUS 11].

L'égalisation d'histogramme est une méthode de transformation non

linéaire de l'image. Elle consiste, d'une part, à améliorer l'aspect visuel de l'image, et d'autre part de réduire le nombre de niveaux de gris de cette image. En pratique, ce nombre est réduit généralement à 64,32, 16 ou 8 valeurs [RUS 11].

La figure 1.4 présente un exemple d'égalisation d'histogramme d'une image I quantifiée sur 256 niveaux de gris, puis réduit à 64 valeurs.

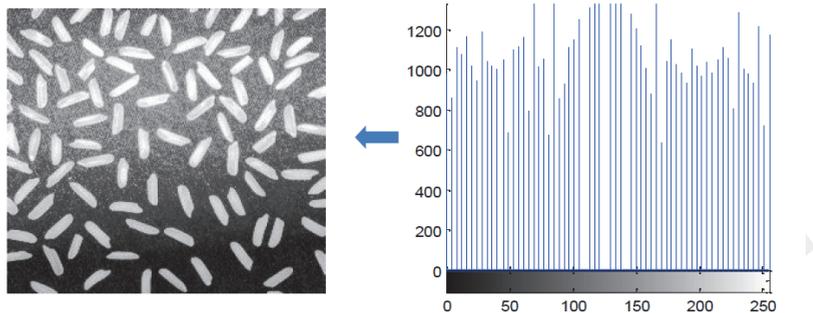


Figure 1.4 Exemple d'égalisation d'histogramme.

Pour les images en couleur plusieurs histogrammes sont nécessaires. Par exemple pour une image codée en RGB :

- Un histogramme représentant la distribution de la luminance ;
- Trois histogrammes représentant respectivement la distribution des valeurs respectives des composantes rouges, bleues et vertes.

2.2.5. Contours

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative. L'extraction de contour consiste à identifier dans l'image les points qui séparent deux régions différentes suivant un critère. Nous donnons plus de détails sur ce concept dans la suite de ce chapitre.

2.2.6. Luminance

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet.

La moyenne ou luminance (brillance) d'une image numérique en niveau de gris est définie comme la moyenne des pixels de l'image :

$$Lum(I) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \quad (1.1)$$

Avec

M : Nombre de colonnes,

N : Nombre de lignes,

$f(x,y)$: La valeur de niveau de gris dans le point (x,y)

2.2.7. Contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images. Si L_1 et L_2 sont les degrés de luminosité respectivement de deux zones voisines A1 et A2 d'une image, le contraste C est défini par le rapport :

$$C = \frac{L_1 - L_2}{L_1 + L_2} \quad (1.2)$$

2.2.8. Bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur.

C'est aussi une altération (déformation) de l'image pouvant être causée par les processus d'acquisition, de transmission ou de stockage. Généralement, le bruit est caractérisé par une moyenne et par une variance [AUB 06]. Les sources de ce dernier sont les étapes de formation des images numériques (l'acquisition, la numérisation et la transmission). Ces sources induisent des fluctuations aléatoires de la valeur des pixels, d'où l'apparition des bruits gaussiens et des bruits impulsionnels [RUS 11]. La figure 1.5 présente une image bruitée par un bruit gaussien.