

Chapitre III : Traitement d'image : Pré-traitements

Le traitement d'image consiste à améliorer le contenu d'une image pour en extraire de l'information : par exemple identifier une séquence de texte [OCR - Optical Character Recognition] ou un chromosome, éviter un obstacle (robotique), détecter des zones soumises à l'érosion (télédétection). En bref traiter une image, c'est lui ajouter du sens par un procédé automatique.

Pour cela il peut être nécessaire de supprimer les facteurs de bruit, de rehausser les signaux utiles contribuant à une meilleure lisibilité. Nous examinerons donc différentes transformations. Certaines affectent tout (ou partie) des pixels de l'image sans qu'intervienne la notion de position spatiale des pixels, ce sont des transformations globales. Les opérations entre images permettent de traiter des séries d'images relatives à la même scène. Enfin des opérations plus élaborées, dites de convolution, portent sur plusieurs pixels voisins et ont donc un caractère local.

Les prétraitements désignent l'ensemble des opérations qui ont pour rôle de donner à l'utilisateur une image libérée des défauts propres au(x) capteur(s). Cette étape peut en fait être très difficilement déconnectée des étapes ultérieures du traitement.

III.1 – Les corrections photométriques

Elles corrigent essentiellement deux types de défauts : d'une part les défauts de linéarité des capteurs, d'autre part les défauts d'homogénéité.

III.1.a : Les défauts de linéarité

Ils sont assez facilement corrigés par des tables de transcodage qui sont rapides et peu chères, les valeurs corrigées du signal n'étant pas calculées mais directement recherchées en adressant une mémoire précalculée. Les défauts de linéarité nécessitent un calibrage des organes d'analyse qui se fait soit sur une rampe de gris continue, soit sur quelques points fixes choisis.

Le meilleur exemple de ce type de correction est le recalibrage (ou correction) Gamma d'un écran cathodique. En effet, un écran CRT (Cathode Ray Tube) n'est pas un système électronique linéaire, c'est à dire que l'intensité lumineuse reproduite à l'écran n'est pas une fonction linéaire de la tension d'entrée. D'un point de vue strictement physique on pourrait s'attendre à ce qu'une correction de gamma soit le processus qui suffise à compenser la non linéarité entre la tension d'entrée et la lumière affichée à l'écran pour obtenir une correcte reproduction de l'intensité lumineuse en un point de l'écran (Figure 3.1).

Cependant la perception humaine de la luminosité n'est pas elle même uniforme mais se comporte plutôt comme une fonction puissance de l'intensité. Finalement pour obtenir une bonne correction de gamma, il faut connaître les caractéristiques physiques de l'écran, le comportement de la vision humaine (ceci engage que l'on définira une correction de gamma pour un observateur de référence avec des conditions d'observation standard) et les conditions de prise de vue des images.

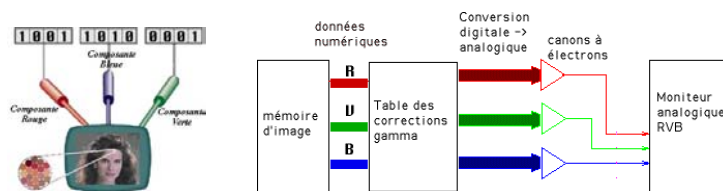


Figure 3.1 : Le principe de la correction Gamma dans la production d'un image sur moniteur CRT

Ce qui intéressant dans le cas de ce facteur Gamma est son influence sur le rendu des couleurs (Figure 3.2).



Figure 3.2 : Influence de la correction Gamma sur le rendu des couleurs

III.1.b : Les défauts d'homogénéité

Ils sont dus aux inhomogénéités spatiales des capteurs ou éventuellement aux défauts des optiques ou des éclairages. Leur correction rigoureuse est très coûteuse et donc rarement effectuée. On lui préfère des corrections approximatives, soit par segmentation de l'image en quelques zones quasi-homogènes, soit en appliquant des corrections locales (soustraction d'une moyenne flottante, division par un contraste flottant, etc...), soit en appliquant des corrections globales approchées (approximation par des quadriques par exemple).

L'assurance d'avoir des documents numériques présentant une très bonne homogénéité de la photométrie permet de gagner en temps de calcul dans les étapes ultérieures.

III.2 – Les manipulations d'histogramme et la fausse couleur

Il ne s'agit pas véritablement de prétraitements, mais ils relèvent d'un type d'opération très voisin des précédents. Ils ont pour objectif de donner à l'utilisateur un document possédant une meilleure lisibilité. Ils utilisent principalement l'information contenue dans l'histogramme et appliquent aux niveaux de gris de l'image une modification indépendante de leurs positions dans cette image.

III.2.a : L'étalement d'histogramme

Egalement appelé **expansion dynamique**. Son but est simplement de donner à une image peu contrastée la dynamique complète qui est autorisée (par exemple 256 niveaux sur 8 bits) par une dilatation artificielle mais **linéaire** de l'échelle de gris utilisée par l'image. On en retire usuellement un **meilleur contraste**.

L'algorithme est le suivant :

Soient V_{\min} et V_{\max} les valeurs extrêmes de la dynamique maximale (en général, $V_{\min} = 0$ et $V_{\max} = 255$).

Notons V_0 et V_1 les limites de la zone non nulle de l'histogramme de l'image I.

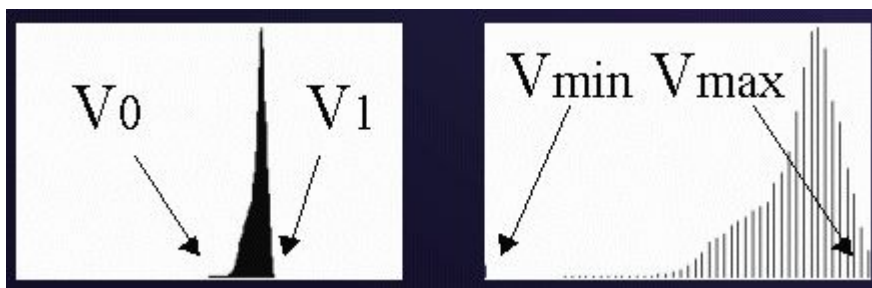
On calcule les valeurs d et g données par :

$$d = \frac{V_{\min} \cdot V_1 - V_{\max} \cdot V_0}{V_1 - V_0} \quad \text{et} \quad g = \frac{V_{\max} - V_{\min}}{V_1 - V_0}$$

La transformation linéaire T suivante transforme le valeur v d'un pixel en valeur v' :

$$\forall v \in [V_0, V_1], v \xrightarrow{T} v' \in [V_{\min}, V_{\max}] \quad \text{avec} \quad v' = g \cdot v + d$$

La figure suivante illustre ce processus :



III.2.b : L'égalisation d'histogramme

De façon générale, une image monochrome possède un histogramme de niveaux de gris dont les composantes non nulles sont groupées dans la partie basse, médiane ou haute de cet

histogramme. Une transformation permettant un étalement (?) de ces niveaux selon des proportions uniformes le long de la gamme est appelée égalisation. C'est une opération qui est appliquée dans deux buts :

- Etendre le contraste, comme précédemment.
- Ramener le document dans une situation où il est aisé de la comparer à un autre.

Le principe est le suivant : on transforme l'histogramme de l'image en un histogramme plat par une table de transcodage que l'on veut strictement monotone afin de préserver l'ordre des contrastes dans l'image.

Dans cette opération, si on reprend l'exemple d'un codage sur 8 bits, la dynamique originale [min,max] est étalée à [0,255]. On cherche à affecter le même nombre de pixels à chaque niveau de gris ; c'est pourquoi on appelle également cette opération **transcodage par équipopulation**. L'effet obtenu permet de mieux séparer les valeurs les plus représentées dans l'image et de rapprocher les valeurs marginales.

Mise en place d'un algorithme d'égalisation d'histogramme

- Calcul de l'histogramme

Soit une image monochrome Im codée sur 8 bits, de taille M par N (N et M entiers) et stockée dans un tableau. On note H le tableau de 256 valeurs entières représentant l'histogramme et Hr l'équivalent en réels, ramené à des proportions (la somme des Hr[i] est égale à 1). Avec le formalisme C, on obtient :

```
char Im[Max_M][Max_N] ;
int H[256] ;
float Hr[256] ;
int i,j ;

/* Initialisation de l'histogramme */

for (i=0; i<256; i++) H[i] = 0 ;

/* Calcul de l'histogramme */

for (i=0; i<M; i++) {
    for (j=0; j<N; j++) {
        H[Im[i][j]] ++
    }
}

/* Calcul de l'histogramme proportionné */

for (i=0; i<256; i++) Hr[i] = H[i] / (M*N) ;
```

- L'histogramme cumulé

Considérons la distribution de niveaux de gris de illustrée sur la figure 1a. L'aire située sous la courbe est égale à 1, représentant la "proportion globale" des niveaux de gris, représentés ici par la fonction f . Il s'agit de l'équivalent d'une fonction de densité de probabilité (fdp) $h_F(f)$ de la variable aléatoire F des niveaux de gris f , que nous considérerons comme une variable continue à valeurs dans $[0,1]$. On a donc :

$$\int_0^1 h_F(f) df = 1$$

La probabilité cumulée à un niveau de gris f est : $\int_0^f h_F(r) dr = H_F(f) = g$ et est appelée fonction de distribution cumulative (fdc). Il s'agit bien évidemment d'une fonction strictement monotone à valeurs dans $[0,1]$ que nous allons utiliser comme transformation non-linéaire des niveaux de gris. L'objectif d'une telle transformation est d'obtenir des niveaux de gris g uniformément distribués, c'est à dire de disposer d'une fonction de densité de probabilité $h_G(g)$ constante pour tout g .

Voici le formalisme C de cet algorithme :

```
Somme = 0 ;

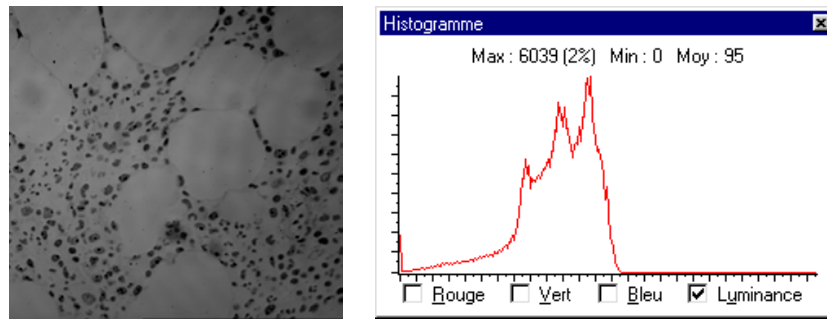
/* Calcul de l'histogramme cumule */

for (k=0; k<255; k++) {
    Somme += Hr[k] ;
    HC[k] = Somme ;
}

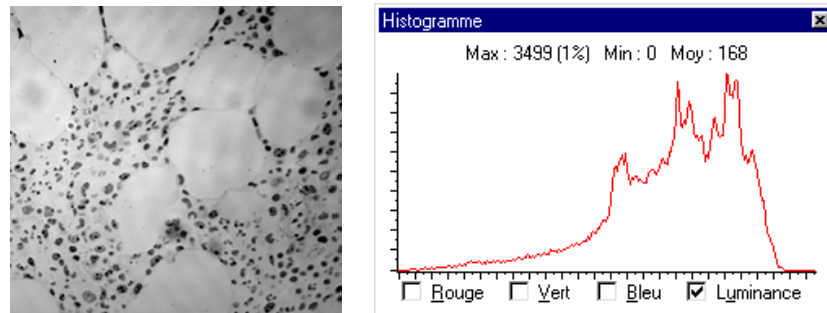
/* Egalisation */

for (i=0; i<M; i++) {
    for (j=0; j<N; j++) {
        Niveau_Initial = Im[i][j] ;
        ImEq[i][j] = 255 * HC[NiveauInitial] ;
    }
}
```

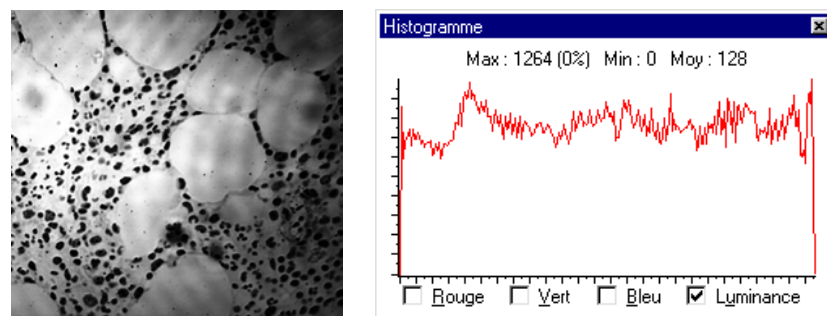
L'aspect discret de notre problème ne permet bien évidemment pas une égalisation rigoureuse, mais les nouveaux niveaux de gris ainsi déterminés seront distribués de façon approximativement uniforme le long de l'échelle 8 bits. La figure 3.3 illustre les deux processus présentés sur un exemple d'image microscopique monochrome.



a : Image et histogramme originaux



b : Etallement d'histogramme



c : Egalisation d'histogramme

Figure 3.3 : Etallement et égalisation d'histogramme

D'autres techniques proposent dans un même esprit des hyperbolisations d'histogramme qui renforcent encore les contrastes apparents

III.2.c : Les manipulations d'histogramme

Si l'on ne s'impose plus d'avoir une table de transcodage monotone, on peut obtenir des effets de renversement local d'histogramme ou exploiter des dynamiques localement très dilatées, qui peuvent être une aide à la détection de structures photométriques subtiles. Néanmoins, ces opérations ne peuvent guère être mises en œuvre que de façon interactive par des moyens conversationnels. Leur généralité demeure très réduite.

III.2.d : La pseudo-couleur

C'est l'étape ultime de la manipulation des niveaux de gris, où l'on s'accorde d'associer à un signal d'entrée noir et blanc trois tables de fausse couleur, chacune fournissant la valeur R, V ou B d'un signal vidéo envoyé sur un moniteur couleur.

III.3 – Les corrections géométriques

Au contraire des corrections de linéarité, les corrections géométriques sont évidemment spatialement variantes, et nous verrons comment ce fait limite la généralité des corrections.

Si nous appelons (x,y) les coordonnées d'un point dans une image de référence (non échantillonnée) et (u,v) les coordonnées de ce même point après déformation géométrique dans l'image distordue, la correction géométrique se caractérise par la transformation G qui associe le couple (u,v) au couple (x,y) :

$$(x,y) = G(u,v) \quad (1)$$

On appelle correction géométrique inverse (il s'agit donc de la déformation) la transformation G_{inv} qui associe (x,y) à (u,v) :

$$(u,v) = G_{inv}(x,y) \quad (2)$$

Dans le cas d'une image continue, la valeur $f(x,y)$ de l'image de référence au point (x,y) est égale après distorsion à la valeur $f'(u,v)$ de l'image distordue au point (u,v) . En traitement numérique, les images sont généralement échantillonnées sur un maillage rectangulaire. Nous ne connaissons donc le signal $f'(u,v)$ de l'image distordue que sur un réseau discret $L_{u,v}$ et nous nous intéressons au signal $f(x,y)$ de l'image de référence que sur un réseau discret $L_{x,y}$.

Les pixels appartenant au réseau $L_{x,y}$ ne correspondent pas aux pixels appartenant au réseau $L_{u,v}$ (voir figure 3.4). Nous devons donc déterminer les bonnes valeurs par interpolation des pixels environnants. Ceci peut être fait selon deux méthodes différentes :

- Dans la première, partant d'un pixel (x_0,y_0) de $L_{x,y}$, on calcule (u_0,v_0) grâce à l'équation (2), puis on détermine la valeur de $f(x_0,y_0)$ en interpolant la valeur équivalente de $f'(u_0,v_0)$ à l'aide des valeurs connues de l'image distordue $f'(u,v)$ sur le réseau $L_{u,v}$ (figure 3.4a).
- Dans la deuxième, on transforme chaque pixel (u,v) de $L_{u,v}$ par l'équation (1). On obtient des pixels corrigés (x,y) extérieurs à $L_{x,y}$, de valeur $f(x,y) = f'(u,v)$ connues. Puis, on calcule la valeur $f(x_I,y_I)$ d'un point du réseau $L_{x,y}$ en interpolant les valeurs $f(x,y)$ des points corrigés environnants (Figure 3.4b).

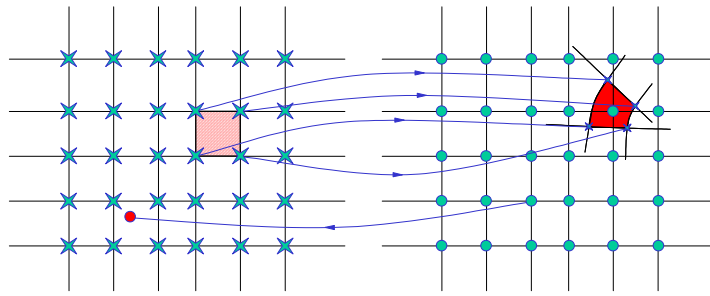


Figure 3.4 : (a) Image numérisée distordue – réseau $L_{u,v}$ (b) Image de référence – réseau $L_{x,y}$

En ce qui concerne la complexité du calcul, ainsi que des besoins en mémoire, les deux méthodes sont équivalentes, mais la première est préférable pour sa meilleure précision de calcul.

L'interpolation à deux dimensions peut se faire par diverses méthodes (par ordre de précision) :

- Identification au plus proche voisin
- Interpolation bi-linéaire (combinaison barycentrique avec 4 voisins)
- Interpolation polynomiale à 12 voisins (ou plus)
- Interpolation par fonctions spline (avec des fonctions B-splines cubiques)

La figure 3.5 illustre ces différentes approches.

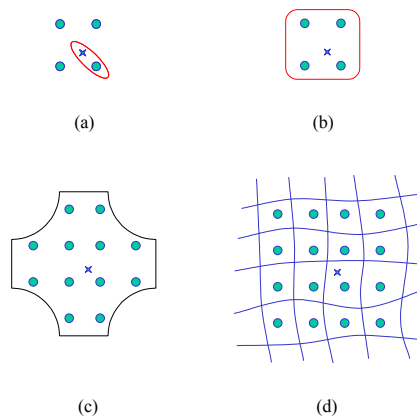


Figure 3.5 : Illustration de 4 techniques d'interpolation bi-dimensionnelles

Les corrections géométriques ont une grande importance pour les photos aériennes ainsi que pour les microphotographies où les aberrations sont notables. L'emploi d'analyseurs à grande ouverture, les perspectives entachées de parallaxe, l'influence des perturbations extérieures pendant la durée de l'exposition constituent quelques unes de ces causes de distorsion géométrique. Nous exposerons brièvement dans les lignes suivantes trois corrections présentant une complexité croissante.

1. Les distorsions géométriques des tubes cathodiques (sans compensation électronique) sont bien connues mathématiquement; qui plus est, on peut les mesurer avec des grilles de calibration. Les lignes horizontales et verticales se transforment en hyperboles. Les corrections géométriques correspondantes s'appliquent facilement.

2. Sur les photos aériennes provenant de spectromètres à balayage, l'attitude de l'avion (roulis, tangage,...) a une importance considérable car elle modifie chacune des lignes. Pour corriger un tel défaut, il faut connaître tous les décalages. On peut y parvenir par une corrélation mono-dimensionnelle de deux lignes successives, en recherchant la valeur la plus élevée de la corrélation, puis en décalant une ligne par rapport à l'autre.

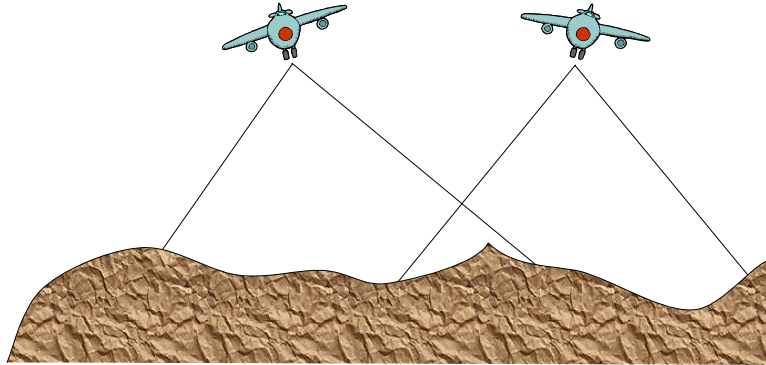


Figure 3.6 : Le problème de la mise en correspondance de deux images

3. L'ultime étape de la correction géométrique est ce qu'on appelle la **registration** des images. Ce mot est d'ailleurs parfois précisé par des mots tels que congruence ou rectification. C'est le cas par exemple lorsqu'on désire mettre en correspondance deux photos aériennes d'une même région, prise de deux points de vue différents (figure 3.6). On doit effectuer des corrections de translation, d'orientation (rotation), d'échelle (homothétie due à la différence d'altitude), mais aussi de différence de perspective; cette différence est très difficile à déterminer et à corriger.

Les différences de translation, d'orientation et d'échelle sont en général compensées à l'aide des paramètres de bord et d'un modèle mathématique des déformations. Les corrections de perspectives sont plus difficiles à mettre en œuvre sur des bases théoriques, car elles demandent une bonne connaissance de la surface du terrain (un modèle numérique du terrain est un échantillonnage de l'altitude en chaque point de ce terrain). On les réalise souvent par morceaux, en extrayant des deux images des points caractéristiques (carrefours, confluent de rivières, orées des bois, fronts de mer, lacs, etc) puis en les mettant en correspondance deux à deux par interconnexion locale, enfin en corrigeant l'une des images par une triangulation basée sur ces points caractéristiques (figure 3.7).

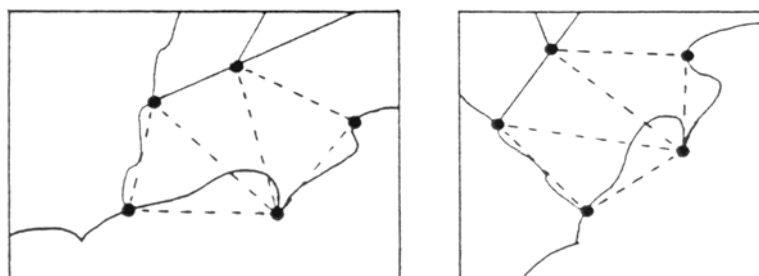


Figure 3.7 : Mise en correspondance par triangulation

III.4 – La suppression du bruit

III.4.a : Filtrage

La notion de filtre est empruntée à la physique et aux techniques de traitement du signal. Si un signal (électrique, radio, image, etc.) présente des composantes de fréquences très différentes, il peut être avantageux d'en éliminer certaines ; on parle de filtrage. Par exemple un filtre passe-bas enlève toutes les composantes du spectre de valeurs supérieures à une valeur de fréquence donnée.

Filtrer une image c'est lui appliquer une transformation mathématique qui modifie la valeurs de gris de tout ou partie des pixels. Si cette fonction est linéaire au sens algébrique du terme, on parle de filtrage linéaire. Si la fonction prend en compte les valeurs de gris au voisinage de chaque pixel transformé, on parle de fonction de convolution. Le terme le plus général est celui d'anamorphose, qui signifie "image grotesque donnée par un miroir courbe".

Contrairement aux manipulations d'histogrammes, qui sont des opérations ponctuelles sur tous les pixels de l'image, les filtrages sont des produits de convolution qui mettent en jeu l'environnement (voisinage) de chaque pixel.

Un filtre est défini par:

- Une fenêtre carrée de dimension impaire (3 ou 5) qui est déplacée sur l'image
- Une matrice de coefficients de même dimension (3 ou 5)

$$\begin{array}{ccc} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{array}$$

- Une transformation combinant les valeurs recouvertes par la fenêtre et les valeurs de la matrice. Le filtre est linéaire si l'opération est une combinaison linéaire. Si $Ne(x,y)$ est le niveau d'entrée du pixel de coordonnées x et y , le niveau de sortie de ce pixel sera donné par la combinaison linéaire :

$$\begin{array}{lcl} Ns(x,y) = & a_{1,1}. Ne(x-1,y-1) + & a_{1,2}. Ne(x-1,y) + & a_{1,3}. Ne(x-1,y+1) \\ + & a_{2,1}. Ne(x,y-1) + & a_{2,2}. Ne(x,y) + & a_{2,3}. Ne(x,y+1) \\ + & a_{3,1}. Ne(x+1,y-1) + & a_{3,2}. Ne(x+1,y) + & a_{3,3}. Ne(x+1,y+1) \end{array}$$

L'application du filtre à l'image consiste à déplacer la fenêtre et à remplacer la valeur de chaque pixel par le résultat de l'opération portant sur ses voisins. Une nouvelle image est ainsi générée (un facteur de normalisation permet de conserver un intervalle de valeurs identique en entrée et en sortie).

III.4.b : Introduction et hypothèses

Parfois certains défauts (poussières, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs, etc.) se traduisent par des taches de faible dimension dont la distribution sur l'image est aléatoire. C'est une approche introductive de la notion de bruit.

Il existe de nombreuses méthodes théoriques destinées à extraire des images un bruit particulier. Nous ne parlerons dans ce chapitre que des méthodes qui reposent sur une connaissance minimale du signal et du bruit. Elles se caractérisent aussi par des algorithmes très simples et une implantation évidente.

Leur principe de base peut s'exprimer de deux façons différentes :

- On utilise la redondance élevée du signal d'image pour éliminer le bruit
- On atténue les hautes fréquences qui contiennent essentiellement du bruit.

Les hypothèses générales sur lesquelles reposent les procédés qui suivent sont essentiellement :

- Les bruit est additif
- Le bruit est indépendant du signal
- Le bruit est fortement décorrélé (bruit à large bande)

Le modèle le plus simple d'un tel bruit est le bruit blanc, c'est à dire un bruit à spectre de densité de puissance uniforme ou à autocorrélation microscopique.

III.4.c : Les filtrages anti-bruit

- Le filtre de Moyenne

Le filtrage des bruits blancs conduit à des filtres de moyenne (ou de lissage) qui remplacent le pixel courant par la valeur moyenne des pixels sur une fenêtre 3x3 ou 5x5.

Ce filtre est bien compatible avec notre observation de départ selon laquelle le bruit est caractérisé par d'importantes hautes fréquences. Mais certaines fréquences (assimilées à de fortes variations dans les niveaux de gris) sont définitivement détruites, le lissage agissant comme un filtre passe-bas. L'image est non seulement lissée mais elle est aussi fortement altérée, les contours étant particulièrement dégradés par les filtres de moyenne.

Le filtre de moyenne est une version d'une gamme plus vaste de filtres de lissages dits linéaires. Ici, le masque représentant ce filtre est le suivant :

1	1	1
1	1	1
1	1	1

Ce filtre, normalisé par un facteur 1/9 pour le ramener dans les valeurs normales, génère un lissage fort. Un lissage moyen peut être réalisé par le biais du filtre suivant :

1	2	1
2	4	2
1	2	1

Le facteur de normalisation est sans surprise $1/16$ et le coefficient plus élevé du pixel central permet d'atténuer l'effet de lissage.

- Le filtre Médian

Les inconvénient du filtre de Moyenne induisent une orientation vers le filtre dit Médian, qui préserve les contours. Ce filtre vise à remplacer la valeur du pixel central par la valeur médiane de la répartition (on trie les luminances dans l'ordre croissant) des niveaux de gris des pixels situés à l'intérieur de cette fenêtre (figure 3.8)

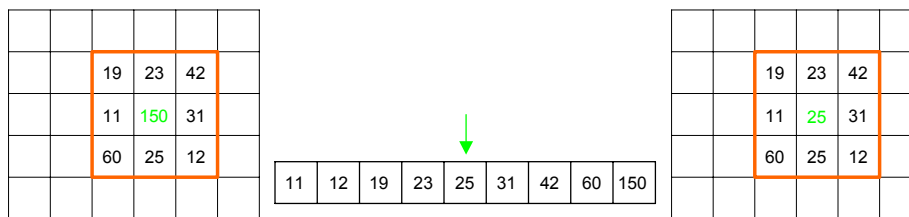


Figure 3.8 : Mise en œuvre du filtre Médian

Ce filtre - non linéaire - permet d'éliminer des pixels isolés ou les traînées fines noyées dans des zones assez homogènes.

- Le filtre de Nagao

Il s'agit d'un filtre utilisé fréquemment dans les images comportant de très fortes structures artificielles (images aériennes de zones urbaines, vision industrielle, etc). Il découpe un fenêtré 5x5 centrée sur le pixel en 9 fenêtrés de 3x3 chacune et mesure sur chacune de ces fenêtrés une valeur de l'*homogénéité* (variance par exemple). Le pixel central est alors remplacé par la valeur de la zone la plus homogène (Figure 3.9).

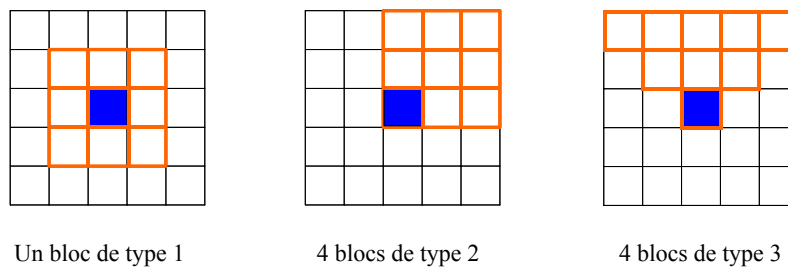


Figure 3.9 : Mise en œuvre du filtre de Nagao

Des indicateurs autres que la variance peuvent être utilisés. A titre d'exemple, on peut citer l'écart entre la plus petite et la plus grande valeur des pixels sur cette image, afin de quantifier une homogénéité plus stricte, plus sensible au bruit.