Le traitement analogique d'une image est limité actuellement aux fonctions classiques de l'optique: grandissement géométrique, symétrisation par un miroir, déformation géométrique simple (panavision par exemple). Le traitement de la fonction intensité est quasi absent, la majorité des principes physiques connus ne réagissant pas à la fonction intensité autrement que sous forme linéaire (la photographie ne fait que restituer et non traiter). Seule l'optique en lumière cohérente permet d'envisager des traitements avancés (Transformée de Fourier par diffraction, holographie...)

Pour envisager l'analyse du contenu d'une image, on fait appel au traitement numérique donc discret de l'image. Pour obtenir une image "numérique", deux aspects de la notion de discrétisation sont à prendre en compte:

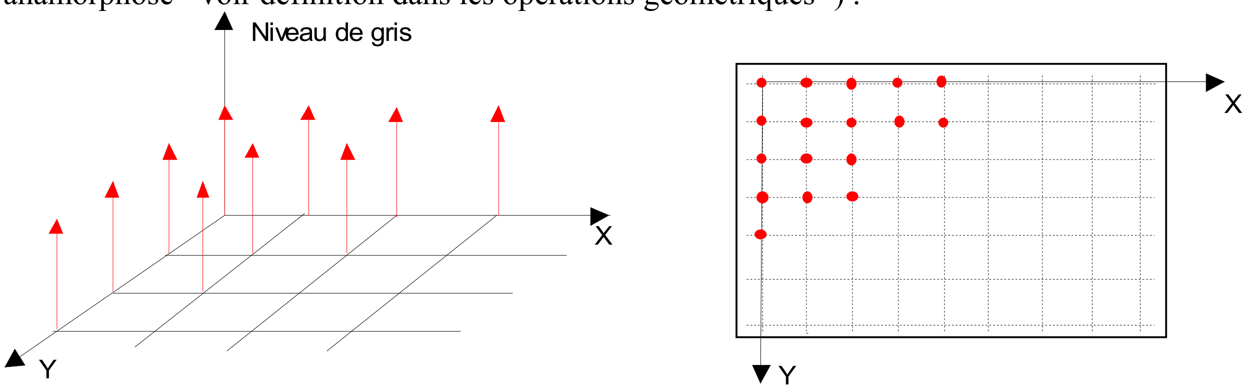
- la discrétisation du domaine spatial de définition de l'image
- la quantification de sa fonction intensité .

DISCRETISATION SPATIALE

La discrétisation spatiale est une opération d'échantillonnage d'un signal au sens classique . Elle fait donc appel à des échantillonneurs parfaits ou imparfaits, régulièrement disposés sur la surface de l'image.

Mailles de discrétisation

Un échantillonneur parfait étant représenté par une impulsion de Dirac, la figure formée par les point d'échantillonnage forme le *maillage* ou *réseau de discrétisation* de la surface image; il est le plus souvent à maille constante sauf dans certains cas spécifiques (images ayant subies une anamorphose - voir définition dans les opérations géométriques -) .



1 - Fig

Fig 1 - Echantillonneurs ponctuels en brosse de Dirac et maillage rectangulaire

Chaque point échantillonné est repéré par des coordonnées discrètes (m,n) et possède pour niveau de gris f(m,n).

En considérant que chaque élément de la brosse de Dirac de pas (T_x, T_y) a pour coordonnées (mT_x, nT_y) et pour expression $\delta(x-mT_x, y-nT_y)$, le niveau de gris du point échantillonné est:

$$f(m,n) = \iint f(x,y)\delta(x-mT_x, y-nT_y)dxdy$$

L'image échantillonnée est une brosse de Dirac, dont chacun des éléments s'appuie sur la surface représentative de l'image (représentation 3D surfacique).

Autres mailles

D'autres éléments *maillants* sont proposés par les chercheurs car certaines propriétés des maillages sont déterminantes dans les algorithmes de traitement.

La plus critique est l'*isotropie* du maillage: quelles sont les modifications de l'image échantillonnée lorsque le contenu subit une *rotation*.? Le maillage carré est très peu performant (fortes déformations de l'image échantillonnée) car il ne possède 4 directions principales.

Pour améliorer l'isotropie, on propose le maillage "hexagonal". Son intérêt est que chaque pixel est équidistant de ses 6 voisins dans 6 directions principales.

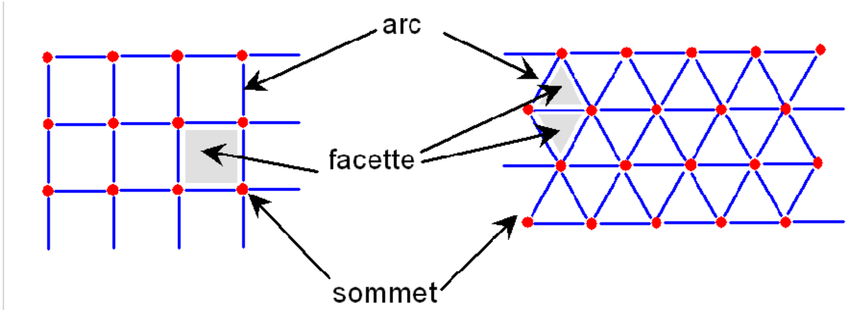


Fig 2 - Maillages carré et "hexagonal"

D'autres notions comme les *relations de connexité* seront beaucoup mieux satisfaites sur le maillage "hexagonal". C'est pourquoi certains logiciels proposent un rééchantillonnage en maille "hexagonales" ainsi que des traitements (traitements dits "morphologiques" par exemple). Jusqu'à maintenant, aucun fabricant ne propose de caméra à maillage hexagonal.

Pavage

A chaque échantillonneur (sommet du maillage), on associe une surface élémentaire dite "pavé" ou "tesselle". Ce pavé peut représenter la surface photosensible du capteur ou bien la surface irradiante d'un afficheur (reproduction de l'image échantillonnée).

Il y a indépendance entre la forme du pavé et le maillage. Parmi les pavages, on met en évidence les notions suivantes:

- pavage régulier ou irrégulier
- pavage alterné associant deux types de pavés
- pavage jointif, non-jointif, recouvrant

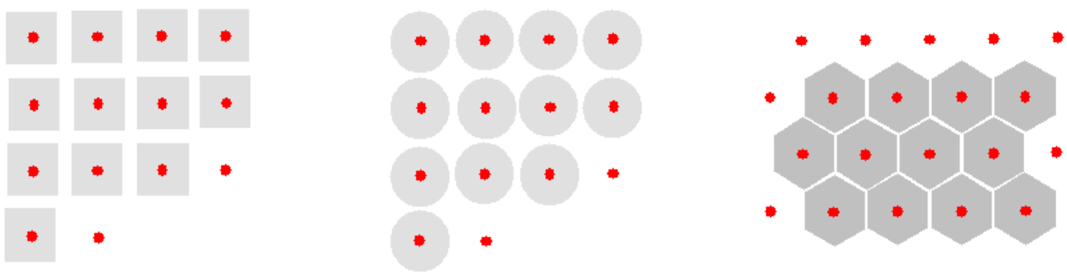


Fig 3 - Exemples de pavages réguliers non-recouvrants

En particulier, on retient souvent le pavage *régulier, jointif de taille maximale*. C'est le pavage le plus efficace en reproduction (pas de zone obscure)

Exemple de pavage spécifique :

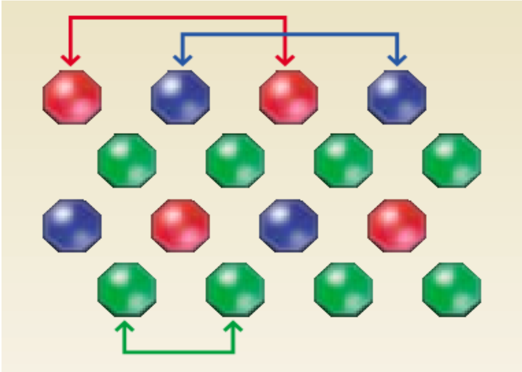


Fig 4 - Pavage du capteur Fuji Super CCD

Physiquement, les capteurs d'image utilisent un pavage pour faire la mesure de lumière.

C'est pourquoi les capteurs fournissent comme valeur de niveau de gris la ***moyenne*** (ou l'intégrale) sur une surface centrée sur le point d'échantillonnage formant le *pavé photosensible*. Cette solution présente cependant l'inconvénient de limiter la bande passante du capteur par rapport à l'échantillonnage ponctuel.

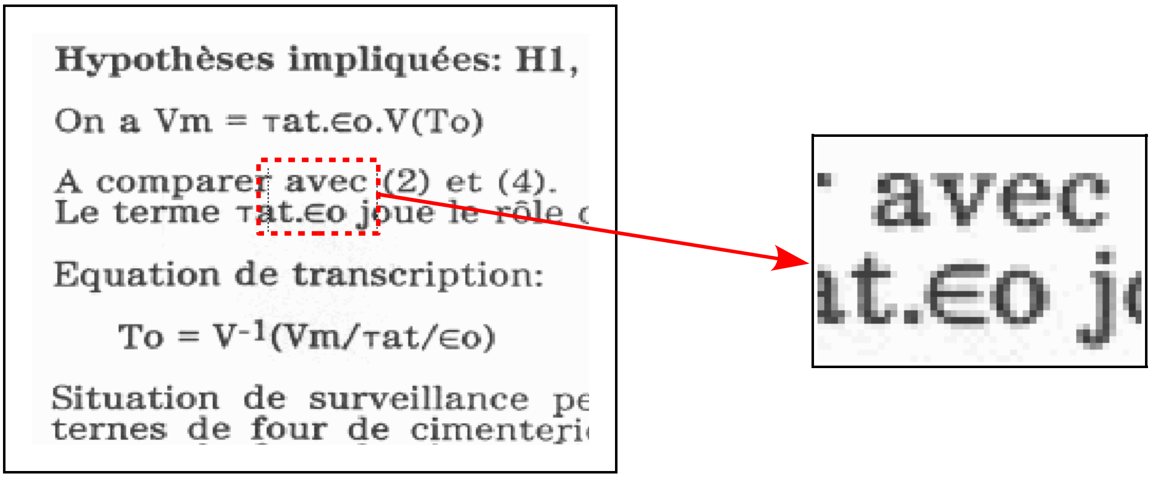


Fig 5 - Effet de moyenne provoqué par le pavage sur le contenu des échantillons

L'échantillonnage imparfait fournit une amélioration *visuelle* de la qualité des images. En effet, la mesure globale d'énergie par sommation pondérée sur les pavés limite le phénomène de *crénage* [parfois appelé à tort *aliasing*] en introduisant des pixels gris sur les bords des objets. De plus, les objets de très petite dimension (de l'ordre du pixels) sont "visibles" sous forme d'un pavé gris.

L'échantillonnage spatial défini ci-dessus (moyenne) est un modèle théorique. Les pavés sont généralement non-jointifs, leurs propriétés photosensible n'est pas constante sur toute leur surface. La mesure photométrique est influencée par ces propriétés, ainsi que les propriétés de l'image échantillonnée.

Soit $B(x,y)$ la loi de sensibilité [pondération] sur le pavé, le signal échantillonné en (mT_x, nT_y) a pour expression le *produit de convolution* (au signe près pour les variables *x* et *y*) entre *f* et *B* :

$$f(m,n) = \iint f(x,y) B(x - mT_x, y - nT_y)\, dxdy$$

On peut montrer que l'échantillonnage imparfait peut conduire à une inversion de constraste locale pour des informations à caractère fréquentiel, c'est à dire que les zones sombrent deviennent claire et réciproquement.

Résolution

La résolution qualifie le nombre de point d'échantillonnage ou le pas relatif. Elle s'exprime par différentes mesures possibles:

- le nombre de point dans chaque dimension (les pixels sont parfois rectangulaires)
- le nombre total de points de mesure (le rapport hauteur/largeur étant supposé connu)
- la taille des pixels et la taille totale du capteur
- la taille du pixel projeté sur la scène et le champ total
- le nombre de points par unité de longueur (ex : 600dpi pour un scanner A4)

La discrétisation spatiale introduit une limite de résolution spatiale, c'est à dire une limite du nombre d'informations pouvant être transcrites discrètement, *chaque détail significatif (blanc ou noir) contenu dans une image devant être traduit par un pixel au moins*.

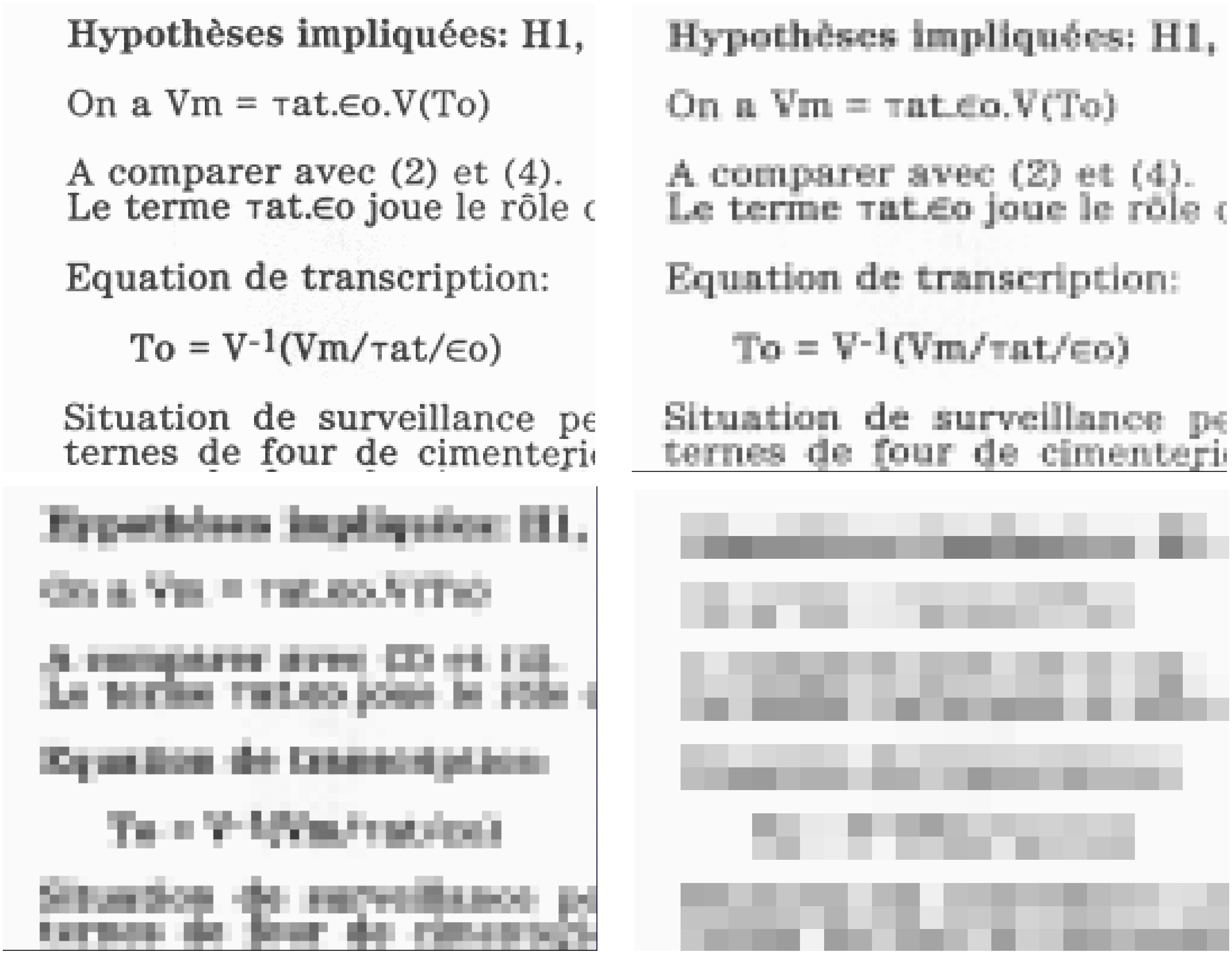


Fig 6 - Influence de la résolution spatiale (250x200, 125x100, 75x50, 25x20

Les images en sous-résolution sont souvent utilisées pour faire une pré-détection des informations de l'image (les lignes de texte dans l'exemple ci-dessous).

Théorème de Shannon

Les choix pratiques de résolution (perception de l'élément le plus petit à traiter) rejoint la règle définie par le théorème de Shannon:

Une information à caractère répétitif peut être assimilée à une fonction fréquentielle de période T_0 (exprimée en m). Pour une période du signal, il correspond deux informations utiles en termes d'image, l'une correspondant à la crête positive (le blanc), l'autre à la crête négative (le noir). Le nombre minimum de pixels pour traduire cette information est de deux. Il en résulte que *la période d'échantillonnage T_x ou T_y doit être au moins deux fois plus petite que celle du signal T_0* (ou que la fréquence d'échantillonnage doit être double de la fréquence maximale contenue dans le signal)

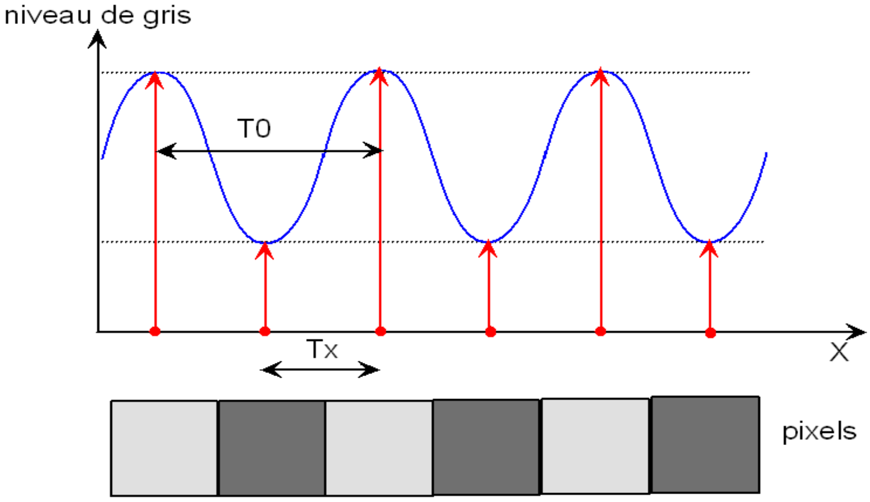


Fig 7 - Échantillonnage à la fréquence limite de Shannon

La présence de détails "sub-pixels", c'est à dire correspondant à des fréquences spatiales élevées, non gênantes dans certains cas (détails isolés se traduisant sous forme d'un gris de faible contraste), peut conduire à une image discrète non représentative de l'image originale. C'est le phénomène d'*aliasing* ou de *repliement de spectre* qui se produit lorsque la fréquence du signal est supérieure au double de la fréquence d'échantillonnage (Théorème de Shannon). Il apparaît des informations fréquentielles superposées au contenu visible de l'image, voire une image fantôme.

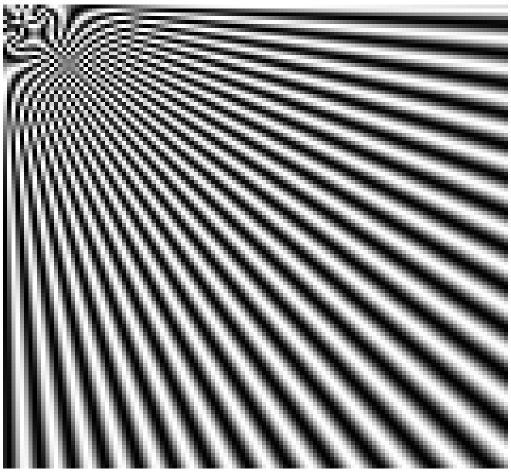


Fig 8 - Phénomène d'aliasing

La seule méthode d'élimination de *d'aliasing* est le filtrage préalable de l'image dans le domaine optique[continu] par un filtre à front raide. C'est une opération très difficile à réaliser [transformateur optique de Fourier, filtrage du spectre, transformateur inverse]. Paradoxalement, un manque de netteté optique introduit typiquement un filtrage passe-bas; le phénomène d'aliasing en est fortement diminué. C'est pourquoi de nombreuses caméras sont fournies avec un objectif de faible qualité.

Influence de l'Echantillonnage imparfait

L'échantillonnage imparfait intègre l'image sur une certaine largeur, ce qui peut provoquer des effets spécifiques :

- l'intégrale du signal périodique sur une surface de pavage peut donner un résultat constant pour chaque point. Ce phénomène apparaît pour un signal de période $3\mathrm{T}_x$ avec T_x pas d'échantillonnage et un pavage de taille T_x (pavés jointifs).

- le contraste est inversé pour un signal de période comprise entre $2\mathrm{T}_x$ et $3T_x$. Le blanc prend la place du noir et réciproquement.

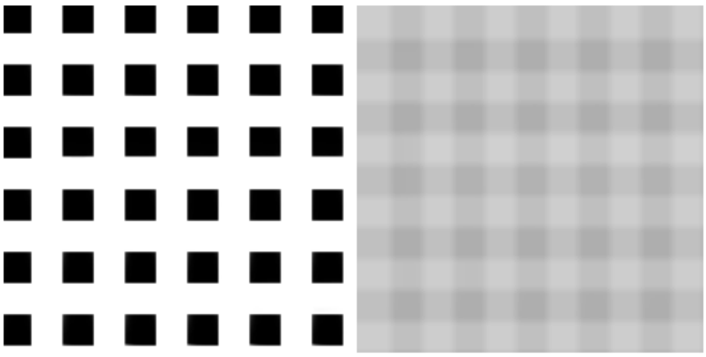


fig 9 - Inversion de contraste par effet de moyenne

Réponse impulsionnelle

Pour tester la performance globale d'un système de prise de vue, on peut utiliser la réponse impulsionnelle, c'est à dire l'image d'un point objet à distance infinie et projeté sur le capteur. L'expérience est facile à mettre en oeuvre et elle permet de prendre en compte l'ensemble des dégradations du système (optique et capteur).

Interpolation SubPixel

La mesure par pavé permet de travailler en mode "sub-pixels", c'est à dire de faciliter l'interpolation d'informations spécifiques comme la position d'un front brusque: le "bord" est sous forme de niveau de gris.

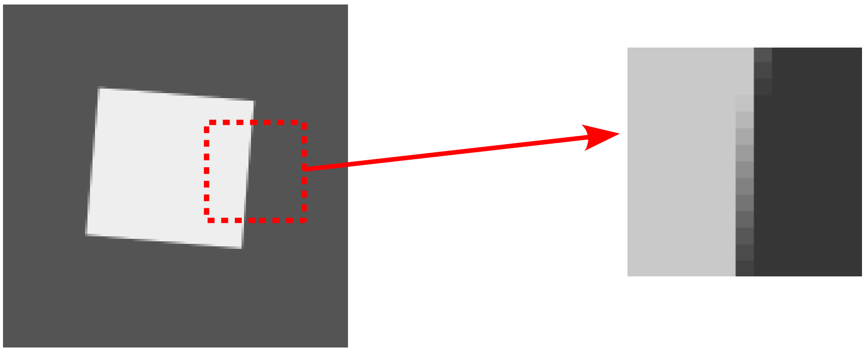


Fig 10 - Image d'un bord parfait

Pour des pavés parfaits, on considère que le niveau de gris représente l'exacte proportion de la transition noir-blanc selon la position du front.

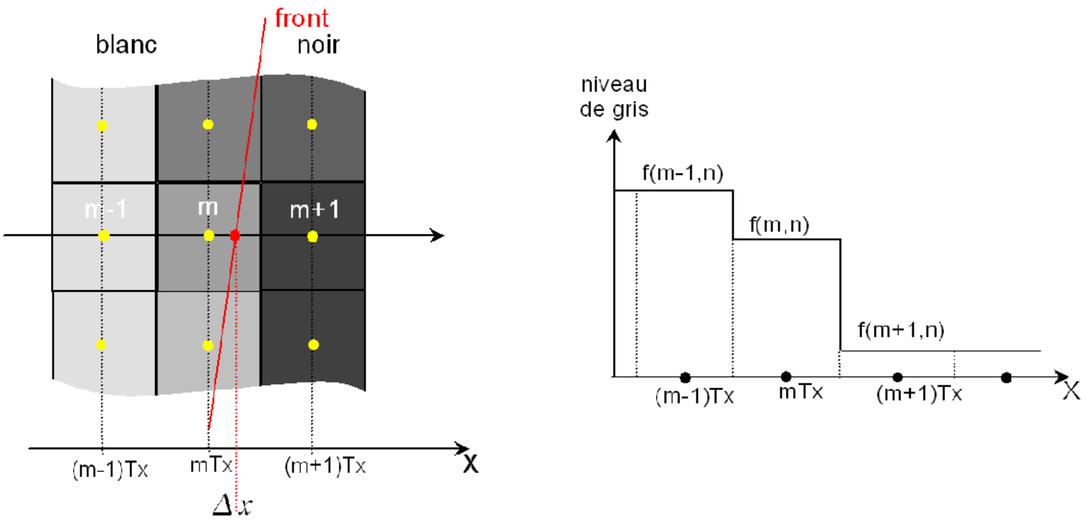


Fig 11 - Interpolation du niveau sur un front :

- soit $f(m-1,n)$, $f(m,n)$ et $f(m+1,n)$ les niveaux de gris des pixels de la transition voisine du point de coordonnées (m,n)

- soit Δx la position relative du front : $-T_x/2<\Delta x<+T_x/2$

L'interpolation linéaire a pour hypothèse que le niveau de gris du pixel $f(m,n)$ est directement proportionnel à la hauteur du front et au décalage relatif du front par rapport au centre du pixel.

$$f(m,n)=f(m-1,n)+\left(\frac{T_x/2-\Delta x}{T_x}\right)[f(m+1,n)-f(m-1,n)]$$

A partir de la mesure de f(m,n), la position interpolée est donnée par

$$\Delta x=\frac{T_x}{2}-\frac{f(m,n)-f(m-1,n)}{f(m+1,n)-f(m-1,n)}T_x$$

QUANTIFICATION DES NIVEAUX DE GRIS

Pour traiter numériquement une image, il est nécessaire de discrétiser la fonction intensité. Deux notions interviennent dans la numérisation :

- la *quantification*, c'est à dire la loi de découpage de la fonction continue (linéaire ou non linéaire).
- le *codage* des différents niveaux (codage en numération classique ou autre).

Quantification linéaire

La quantification linéaire avec codage en binaire naturel est la méthode la plus employée pour passer de la mesure de l'énergie lumineuse continue à la valeur numérique.

Le nombre de niveaux, c'est à dire le nombre de bits du mot dépend de l'utilisation.

- 1 bit : il n'y a que deux niveaux le noir et le blanc. C'est un mode de fonctionnement très simple, couramment utilisé en robotique. On parle alors d'i*mages binaires*. La difficulté est de définir le seuil de décision, elle ne donne de bons résultats que pour des images très contrastées.
- 8 bits pour un usage général: permet de traduire l'ensemble des niveaux de gris que l'oeil humain sait distinguer. L'image digitalisée apparaît donc de qualité identique à l'original. De plus, cette valeur est en correspondance avec les outils informatiques actuels.
- 10 à 14 bits permet de garder toute la dynamique des capteurs modernes. C'est le format des fichiers "RAW" des appareils numériques. Cette précision permet d'appliquer des fonctions numériques de correction (contraste, gain, gamma) sans créer d'artefact de niveau (par exemple, saut important d'une valeur de gris à une autre après correction pour une variation très faible sur l'original)

Un faible nombre de niveaux de quantification présente l'inconvénient de faire apparaitre des *contours fantômes* correspondant à la ligne de séparation de deux niveaux successifs. La distinction entre vrais contours et faux contours devient plus difficile.

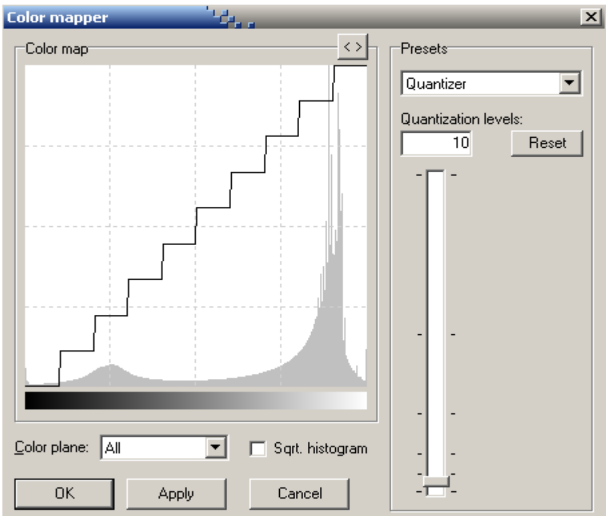


Fig 12 - Quantificateur linéaire (logiciel Image Analyser)

Image Originale 256 niveaux de gris

16 Niveaux de Gris

4 niveaux de gris

2 Niveaux de Gris

Fig 13 - Influence de la quantification du niveau de gris

Sur des images à caractère industriel, les contours fantômes sont très visibles.

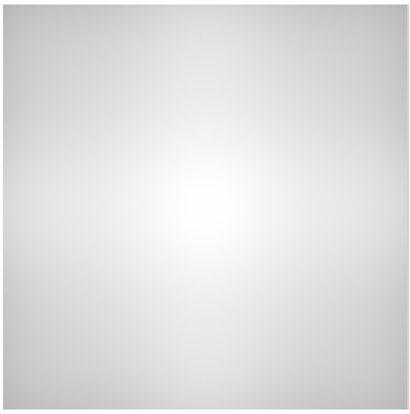
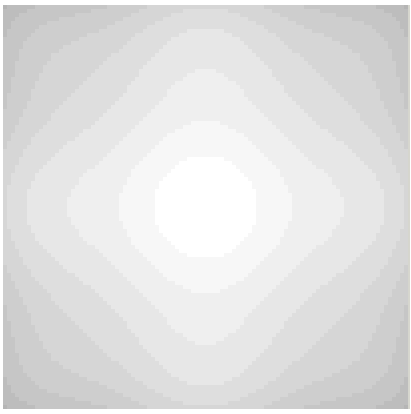


Fig 14 - Image sur 256 niveaux



Contours fantomes sur 32 niveaux-

Bruit de Quantification

La conversion Analogique/Digitale transforme les variations continues du niveau de gris en variations discrètes qui se traduisent par un effet de "grain" sur les surfaces homogènes. Ce phénomène se caractérise quantitativement par un rapport signal/bruit.

Analyse classique

Considérons un convertisseur parfait, c'est à dire caractérisé par une fonction de transfert idéale en escalier. Soit q le pas de quantification correspondant au LSB (Least Significant Bit) du convertisseur. Pour un convertisseur n bits, le pas q vaut $1/2^n$ de la pleine échelle de mesure.

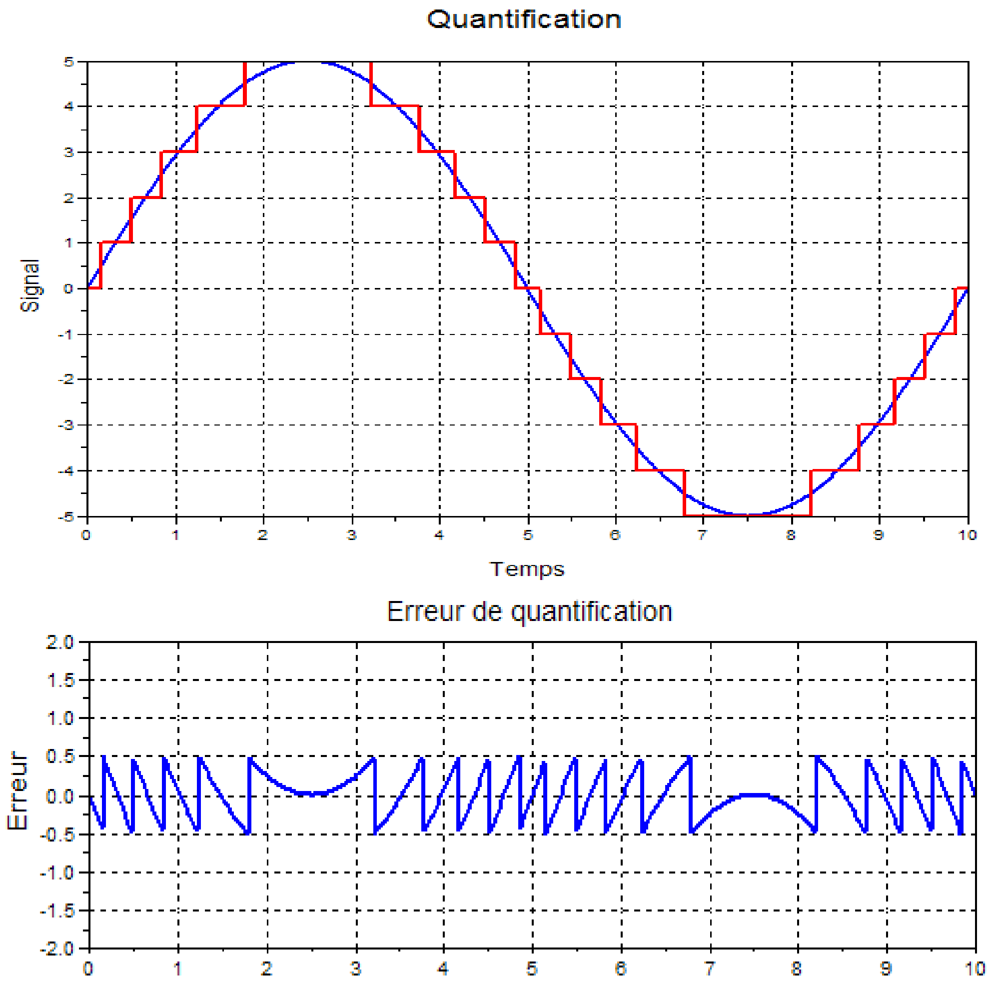


Fig 14 - Erreur de quantification

L'erreur de quantification correspond à l'écart entre le signal analogique et sa représentation digitale. Cette erreur a une valeur comprise entre $-q/2$ et $+q/2$ quel que soit le signal d'entrée considéré.

Pour un signal utile de densité de probabilité uniforme, l'erreur de quantification ε présentera elle-même une densité uniforme dans l'intervalle $[-q/2,+q/2]]$. En normalisant la probabilité de ε, on aura :

$$p(\varepsilon)=1/q \text{ pour } -q/2<e<+q/2$$
$$p(\varepsilon)=0 \quad \text{ailleurs}$$

La moyenne quadratique de ε ou puissance de bruit est $E(\varepsilon^2)=\frac{1}{q}\int_{-q/2}^{+q/2}\varepsilon^2 d\varepsilon=\frac{q^2}{12}$

En considérant que ce bruit se superpose à un signal utile de forme sinusoïdale (notion contestable en Traitement d'Images) , on calcule la puissance ou efficacité du signal:

soit $f(x) = A \sin \omega x$ avec ω pulsation du signal

Sa valeur efficace est $f_{eff} = \frac{1}{T} \int_0^T A^2 \sin^2 \omega x \, dx = \frac{A^2}{2}$

Le rapport Signal/Bruit en puissance est donc :

$$SNR_{dB} = 10 \, \mathrm{Log}_{10} \frac{A^2/2}{q^2/12}$$

En considérant que la dynamique du convertisseur correspond à la valeur crête à crête du signal soit $2A$ et que le convertisseur possède n bits de résolution soit 2^n niveaux de quantification, le pas de quantification est :

$$q = \frac{2\mathrm{A}}{2^n} = \frac{A}{2^{n-1}}$$

Le rapport Signal/Bruit pour un signal sinusoïdal d'amplitude maximale est donc:

$$SNR = 10 \, \mathrm{Log}_{10} \frac{A^2/2}{q^2/12} = 10 \, \mathrm{Log}_{10} \left(\frac{3}{2} 2^{2\mathrm{n}} \right)$$

La valeur pratique est :

$$SNR_{dB} = (1,76 + 6,02\, n)$$

On remarque que le SNR gagne environ 6dB par bit de résolution supplémentaire (soit un facteur 2). Pour une binarisation (n = 1) , le SNR n'est que de 8 dB environ.

Particularités du Traitement d'Images

Les signaux Image sont toujours positifs, ce qui est équivalent à la présence d'un niveau continu sur un signal alternatif. Cette composante continue ne représente aucune information utile et ne participe pas à la définition d'un SNR.

De plus, la notion de puissance d'un signal ou d'efficacité, liée au carré du signal ($P = RI^2$ ou $P = U^2/R$ pour un signal électrique) n'a pas de signification pour les images puisque le signal est déjà représentatif d'une énergie lumineuse. Il convient donc de modifier les relations en prenant la racine carrée des valeurs efficaces de Bruit ou de Signal. Pour la conversion en dB, on utilisera le facteur d'échelle 20. Globalement, on remarque que les résultats pratiques sont inchangés.

Nombre de Bits Equivalent à un SNR

Inversement, il est possible de donner une valeur de n équivalente à un rapport SNR donné en dB. Cette notion permet de choisir un convertisseur pour une application donnée. Le nombre équivalent de bits (Equivalent Number of Bits ou ENOB) est :

$$\mathrm{ENOB} = \frac{SNR_{dB} - 1,76}{6,02}$$

Par exemple, un signal de caméra de SNR = 40 dB correspond à un ENOB de 6,35 bits. Le bruit de quantification venant s'ajouter au bruit propre de la caméra, il faudra choisir un convertisseur d'au moins 7 bits.