

# Images numériques matricielles "Bit Map"

## Codage informatique:

Qu'est-ce que l'image, sur support argentique (photo) ou sur support numérique? De l'information !

Alors qu'en photo la qualité de l'information est liée à la qualité du support (type de film, papier photo), dans le monde numérique, la qualité est indépendante du support physique (disque dur, CD...). Seule compte la quantité d'informations que l'on souhaite voir prises en compte lors de la numérisation (scanner, appareil photo numérique, caméra numérique...)

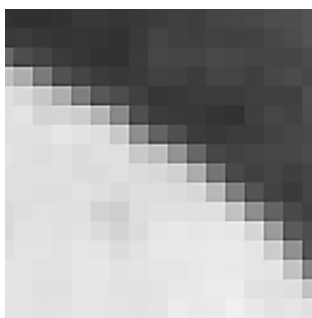
L'information contenue dans une image est toujours très complexe et difficile à représenter. Son codage sous forme numérique utilise obligatoirement une quantité très importante de valeurs, souvent plusieurs millions.

Pour que cette masse de données soit interprétable il faut qu'elle soit structurée de manière très précise, et qu'elle rende compte d'un modèle de description de l'image représentée.

Ces modèles précisent comment les images peuvent être décomposées en éléments simples, qui sont eux-mêmes susceptibles d'être numérisés, c'est-à-dire représentés par une ou plusieurs valeurs numériques.

Pour décomposer une image existante le plus simple est de la découper en une matrice de cellules par simple quadrillage.

Ces cellules sont considérées comme indivisibles et caractérisées par une couleur unique. On les appelle des Pixels, de l'anglais picture element. Plus la taille des Pixels est faible, plus le modèle est proche de l'image originale.



• *Mise en évidence de la Pixelisation lors d'un agrandissement.*

Ce modèle revient à considérer une image comme un ensemble de points.

Dans certain cas il est plus intéressant de décomposer l'image à l'aide d'objets plus complexes, comme des segments, cercles, et autres formes géométriques prédéfinies. L'image est alors décrite par un modèle mathématique, on parle alors d'images vectorielles.

La précision de l'image ou son "piqué": problèmes de Résolution et de Définition.

Un des éléments les plus surprenants, pour l'infographiste débutant est la façon dont on adapte la taille du fichier, "le Poids de l'image", avant la manipulation sur ordinateur, à l'utilisation qui sera faite de l'image terminée.

Parce que le document numérisé est composé de la variation d'une suite de nombres, 0 et 1, la place nécessaire pour stocker l'information de 1 Pixel est importante.

Un élément binaire -1 Bit- est la plus petite quantité d'information existant en informatique. Parce que l'électricité est le véhicule de l'information, celle-ci est binaire: 0=ouvert, 1=fermé. Comme il est difficile de manipuler une suite de données ininterrompue, on la groupe par paquets de 8 Bits, ce que l'on appelle un Octet.

On parle d'image 24 Bits, cela veut dire que chaque Pixel est défini par 24 Bits, soit un Octet par composante Rouge, Vert et Bleu.

La pellicule photo est caractérisée par son format et par le piqué que procure son émulsion. Personne ne penserait à définir une pellicule photo par le nombre d'éléments photosensibles qu'elle contient. L'information contenue dans une image numérique se mesure précisément par le nombre de ses Pixels. Ceux-ci sont disposés régulièrement et servent donc d'unité de mesure: une image peut avoir une largeur de 1000 Pixels et une

longueur de 2000 Pixels.

La définition se mesure en nombre de points par pouce, ppp en français, dpi (dots per inch) en anglais, ou plus rarement en points par centimètre ppc. Entre ici en compte la notion de densité: exemple 300 Pixels peuvent être répartis sur 1 pouce ou sur 10 pouces. Dans le premier cas, l'image aura une définition de 300 ppp la seconde de 30 ppp; la même image peut faire 10 pouces de longueur ou 1 pouce et contenir la même quantité d'informations.

On comprend que visuellement le résultat ne soit pas le même: vu d'une même distance, l'une donnera l'impression d'être "piqué" l'autre pas.

La définition concerne l'image numérisée: sur un pouce le scanner va dégager 300 informations distinctes, converties en autant de Pixels.

La résolution concerne la sortie, que se soit sous forme d'impression quadri ou de film elle est définie par un nombre de lignes par pouce, lpp, ou par millimètre. On parle aussi de linéature ou fréquence de trame. Elle est par exemple de 53 lpp pour une imprimante à jet d'encre standard.

Pour compliquer le sujet, Photoshop utilise les notions de Résolution de Pixel ou quantité d'informations disponibles pour chaque Pixel: on parle d'image 16 Bits ou 24 Bits, cela équivaut aussi à ce que l'on appelle la Profondeur d'image et correspond en fait au nombre de nuances colorées maximum contenu dans l'image.

## **Le codage informatique de la couleur**

On détermine le nombre de couleurs disponibles pour une image par le nombre de Bits servant à coder chaque pixel:

1 Bit = 2 valeurs possibles: noir et blanc

2 Bits =  $2^2 = 4$  couleurs

4 Bits =  $2^4 = 16$  couleurs

8 Bits =  $2^8 = 256$  couleurs

15 Bits =  $2^{15} = 32768$  couleurs

16 Bits =  $2^{16} = 65536$  couleurs

24 Bits =  $2^{24} = 16777216$  couleurs

Une fois l'image décomposée en éléments simples, il faut disposer d'un modèle susceptible de coder la couleur de chacun de ceux-ci, et de la caractériser par un nombre.

Chaque couleur est décomposée en une combinaison de trois couleurs de référence, Rouge, Vert et Bleu, qui sont dosés selon une échelle de nuances, plus ou moins large pour chacune, en fonction du nombre de Bits disponibles pour le codage des Pixels.

## **La compression des données**

En infographie, la place c'est de l'argent: même si le prix des disques durs et autre support de stockage a considérablement baissé, la place occupée par les éléments d'un montage d'images destiné à être publié est énorme.

Plus important que l'argent qu'il faudra pour stocker ces volumes d'informations: le temps nécessaire à leur traitement augmentera en fonction de la taille des fichiers.

Le doublement des dimensions d'une image se traduira par le quadruplement de son poids en mégaoctet (1 million d'octets). On peut imaginer que le temps de calcul sera multiplié par quatre.

Cela pose également et surtout des problèmes de transmission qui sont eux plus difficiles à régler car les canaux de diffusion, la fibre optique mise à part, sont incapables de supporter des débits élevés. Il a donc fallu mettre en œuvre des techniques de compression des données.

Imaginons un reporter devant transmettre à son agence des images d'un poids total de 100Mo, via le réseau

téléphonique qui permet un débit d'environ 10 kilobit/seconde (dans de bonnes conditions), il lui faudra environ 3 heures pour achever sa transmission!

Les signaux numériques sont susceptibles de subir grâce aux ordinateurs des traitements spécifiques, d'origine mathématique, définis par des algorithmes plus ou moins complexes. Les algorithmes sont l'ensemble des règles opérationnelles propres à un calcul; ils se manifestent dans le domaine informatique par des suites d'instructions données à un ordinateur afin qu'il puisse résoudre un problème.

Une panoplie de ces algorithmes, qui sont la clef du développement de la transmission d'information pour les années à venir, permet de réduire fortement la quantité des données numériques sans altérer l'information du message.

Dans le cas des images, il est possible de choisir entre des programmes tels qu'aucune information ne disparaisse (on parle alors de compression transparente ou sans perte); ou accepter de sacrifier une partie de l'information initiale à condition que cela ne soit pratiquement pas perceptible par un spectateur (on parle alors de compression virtuellement transparente); on peut enfin se montrer plus drastique en acceptant, pour s'adapter à des conditions difficiles, de sacrifier la qualité de l'image en ne conservant qu'un faible pourcentage de l'information initiale.

Plusieurs techniques ont été mises en œuvre et testées au cours de ces dernières années; d'autres utilisant par exemple les "fractales" apparaissent. Il est bien évident que des progrès seront encore effectués.

On peut dire de manière cursive que la compression des données numériques, qui permet une importante autant qu'indispensable réduction du débit numérique (BRR: Bit Rate Réduction), est basée sur l'analyse et la mise à profit de l'organisation interne, notamment des redondances, qui existent toujours de manière plus ou moins affirmée dans le contenu des données, qu'il s'agisse de textes, de musique ou d'images.

Une de ces idées consiste à découvrir à l'intérieur de l'information numérique des motifs (au sens d'éléments qui se répètent, comme on le dit d'un motif musical) ou "patterns" dont l'occurrence est élevée, c'est-à-dire que l'on retrouve souvent. Ces motifs peuvent représenter aussi bien des caractères alphanumériques (représentés par un octet dans le code ASCII) que des pixels d'une image. Une fois ces motifs identifiés, plusieurs types d'algorithmes pourront être mis en œuvre.

Une méthode basique relativement évidente consiste à s'occuper de motifs élémentaires (pixels d'une image, lettres d'un texte) et à repérer si plusieurs de ces éléments consécutifs sont identiques, autrement dit, pour une image, détecter s'il existe des plages uniformes (redondance spatiale). On code alors l'information en notant simplement deux informations: le nombre d'éléments successifs identiques et leur référence (octet du code ASCII pour l'alphabet; valeur de la luminance pour un pixel). Ce type de compression simple est appelé codage à longueur variable ou RLC (Run Length Coding).

Le codage télégraphique Morse met en œuvre ce qu'on appelle un codage entropique. L'entropie est une notion issue de la thermodynamique qui permet de caractériser le degré d'organisation d'un système. Le Morse est en fait un langage binaire puisqu'il ne dispose que de deux éléments informatifs, la longue et la brève, correspondant à la durée d'un bip. Il faut donc des combinaisons de ces deux symboles (on dirait aujourd'hui qu'il faut utiliser plusieurs bits) pour coder les 26 lettres de l'alphabet, plus les dix chiffres, plus les signes de ponctuation.

Le code Morse a intelligemment choisi, afin que la transmission soit la plus rapide possible, de représenter les lettres par des messages d'autant plus courts que les lettres sont rencontrées fréquemment. Ainsi le E est représenté par une brève, le T par une longue, le I par deux brèves, le A par une brève et une longue... .

On constitue ainsi une sorte de dictionnaire dit table d'affectation ou LUT (Look Up Table) établissant la correspondance entre les motifs (ici les signes typographiques) et le mot de code qu'on leur attribue en fonction de leur probabilité d'occurrence.

Ces deux types de compressions sont mis en œuvre pour la transmission d'images simples par télécopieur (norme T4 du CCITT).

1977: Le codage LZW

Le codage LZW (Lempel, Ziv et Welch), fort utilisé pour le compactage des fichiers de texte, est apparu dans ses premières versions en 1977. C'est un codage entropique faisant correspondre, grâce à une LUT, un mot de code à un motif.

La table de correspondance est généralement définie à l'aide de 12 bits ce qui correspond à 4 096 "cases" permettant d'indexer autant de motifs. L'originalité du codage LZW réside dans le fait que la table est construite dynamiquement au fur et à mesure que l'on analyse le fichier.

Lorsque l'ordinateur rencontre pour la première fois un motif nouveau, il ne le traite pas en tant que tel; il apprend à le reconnaître (une sorte de vaccination en somme) et le repère par un index simple qui est le numéro d'une case de la table; lorsque ce motif réapparaît ultérieurement, l'ordinateur le remplace automatiquement par son index. La structure de la table est donc implicitement contenue au sein des données compressées; elle pourra se trouver reconstruite, d'une manière analogue à celle utilisée pour sa construction, lors de la décompression du fichier.

Il n'est donc pas nécessaire de la transmettre.

## JPEG et MPEG

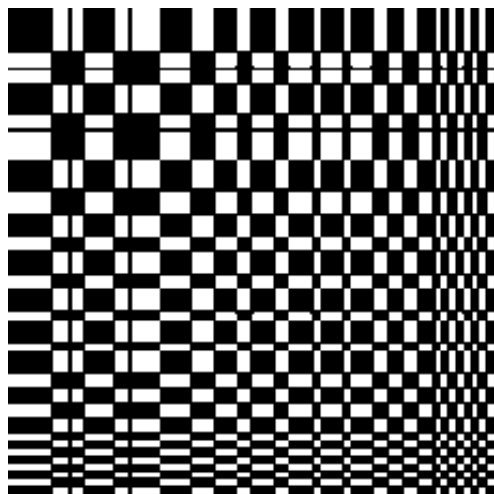
Deux familles d'algorithmes sont aujourd'hui normalisées et couramment utilisées ce sont les algorithmes JPEG (Joint Photographic Experts Group) pour les images fixes et MPEG (Motion Picture Experts Group) pour les images animées.

La première étape a consisté à mettre au point une méthode permettant de décrire l'image suivant un catalogue de motifs prédéfinis que l'on peut repérer à l'aide de méthodes mathématiques assez simples.

Cette méthode repose sur une opération mathématique nommée "Transformation Cosinus Discrète" (TCD) ou "Discret Cosine Transform" (DCT). Le premier terme indique que l'on va modifier la manière de caractériser le signal; le deuxième terme indique que l'on va profiter des "bonnes" fonctions mathématiques que sont les fonctions sinus ou cosinus; le troisième terme réfère au caractère discontinu ou discret de l'information à traiter, il indique que le processus portera sur des signaux numérisés.

On sait qu'on peut définir un son musical, à variation périodique, comme étant constitué par une somme de sons purs, à variations purement sinusoïdales. On trouve une composante à la fréquence du son musical considéré, c'est la fondamentale, ainsi que des composantes d'amplitudes variables dont les fréquences sont des multiples de cette fondamentale: ce sont les harmoniques. On nomme ce processus "Analyse Harmonique" ou en termes mathématiques "Décomposition en série de Fourier". Inversement, les synthétiseurs électroniques reconstituent les timbres des instruments en mélangeant avec des niveaux divers des sons purs provenant d'oscillateurs électroniques. On peut donc utiliser deux manières équivalentes pour décrire un phénomène périodique: soit exprimer sa variation dans le temps (domaine temporel), soit donner son spectre c'est-à-dire la liste des fréquences qui permettent de le recomposer (domaine fréquentiel).

On procède d'une manière analogue pour l'image en l'analysant en termes de composantes de différentes fréquences spatiales (en vertical et horizontal); ces différentes fréquences correspondent à des motifs comportant des bandes (verticales ou horizontales) régulièrement disposées ce qui donne naissance à un catalogue de motifs qui ressemble à un catalogue de tartans écossais.



• Les 64 motifs que l'on peut obtenir pour un bloc de 8x8 pixels.

L'image est habituellement décomposée en blocs de 8 x 8 pixels auxquels on fait correspondre des coefficients obtenus par la DCT. La DCT, qui est assez facilement mise en œuvre sur un ordinateur, permet de remplacer la représentation dans le domaine spatial de type bit-map par une représentation dans le domaine fréquentiel. On substitue au tableau constitué par les 64 nombres affectés aux points du bloc (domaine spatial) un autre tableau comportant les coefficients d'amplitude des composantes fréquentielles permettant de caractériser 64 motifs périodiques simples (domaine fréquentiel) qu'il est possible de fabriquer avec les 64 pixels du bloc. Le tableau est tel que les fréquences horizontales croissent de la gauche vers la droite et les fréquences verticales de haut en bas. Le motif en haut à gauche correspond à la valeur moyenne de luminosité du bloc. Lorsqu'on s'éloignera en diagonale en partant du coin haut à gauche pour aller vers le coin bas à droite, on rencontrera des motifs de plus en plus fins correspondants à des détails de plus en plus ténus.

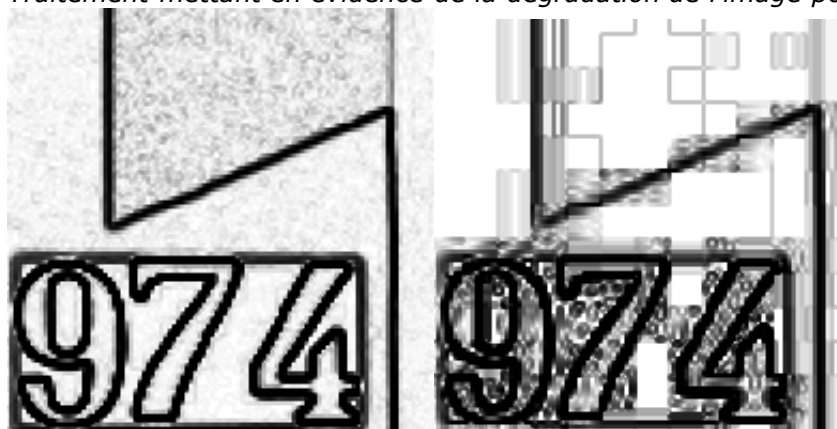
On constate, pour la plupart des images naturelles, que les coefficients de la partie du tableau situé en dessous et à droite de la diagonale ont presque toujours des valeurs soit très faibles soit nulles. On conçoit donc déjà qu'il est plus économique, en termes de flux numérique, de transmettre la fiche d'identité des quelques fréquences présentes au-dessus de la diagonale (les fréquences basses) plutôt que celles des 64 pixels individuels du bloc. On obtient ainsi une transformation totalement transparente, c'est-à-dire que toute l'information présente dans le bloc est conservée.

On peut, au prix de quelques sacrifices, augmenter le taux de compression en réalisant une compression virtuellement transparente (c'est-à-dire passant inaperçu lors de la restitution).

Ceci est obtenu en négligeant tout d'abord les coefficients inférieurs à un certain seuil puis en affectant à chaque composante fréquentielle conservée un poids tenant compte de l'importance que représente pour la psychophysologie de la perception la fréquence correspondante; les coefficients les moins importants (les hautes fréquences c'est-à-dire les détails peu perceptibles) sont codés avec une faible précision sur peu de bits, leur valeur initiale étant arrondie à la valeur la plus proche d'un des larges échelons de quantification ainsi créés.

Un taux de compression important et non pénalisant pour le spectateur peut être obtenu par un choix adapté des valeurs des seuils et des coefficients de pondération permettant de n'éliminer que des informations non pertinentes (irrelevant en anglais) pour l'œil.

- *Traitement mettant en évidence de la dégradation de l'image pour un taux de compression maximum.*



Le tableau est ensuite lu en zigzag en commençant par le haut à gauche (les basses fréquences); ce mode d'exploration conduit à rencontrer de longues séquences de zéro qui seront économiquement codés sous forme condensée (longueur de la séquence) par un codage à longueur variable (RLC). On effectue ensuite une analyse statistique de l'occurrence des différents coefficients pour réaliser un "codage entropique" par l'intermédiaire d'une table d'indexation (LUT).

## Les Fractales

C'est le mathématicien français Benoît Mandelbrot qui inventa en 1975 le nom fractal pour définir des entités mathématiques complexes mais qui pouvait être définis d'une manière récursive grâce à des règles génératrices relativement simples.

En voici deux exemples élémentaires et classiques: la courbe de Koch et le tamis de Sierpinski.

Les objets fractals possèdent une propriété essentielle qui est l'auto-similarité. Les détails les plus fins ont une structure identique à celle de la figure dans son ensemble, le zoom dans un objet fractal fait retrouver indéfiniment des images identiques.

On s'est vite aperçu que certaines des images obtenues pouvaient ressembler à des images naturelles. Mandelbrot dans son livre *The fractal geometry of nature* montra que des objets naturels aussi divers que des galaxies des montagnes, des arbres, la côte de la Bretagne possédaient la propriété d'auto-similarité et qu'ils pouvaient être décrits économiquement en termes de structures fractales.

Le mathématicien américain Michaël F. Barnsley déposa en 1988 un brevet décrivant une méthode de compression des images utilisant une méthode de type fractale.

Il décrit avec humour le principe de la méthode sous l'appellation d'algorithme de la photocopieuse: une photocopieuse bénéficiant d'une fonction de changement d'échelle. Son exemple porte sur la construction d'une feuille de fougère, élément simple et auto similaire s'il en est.

Plus généralement, la méthode de compression fractale consisté à trouver des motifs récurrents à l'intérieur de

l'image, à en dresser un catalogue puis à spécifier les zones de l'image où ils ont été rencontrés. À la différence de la méthode JPEG, ces motifs ne relèvent pas de fonctions mathématiques mais sont des éléments de l'image elle-même.

Il s'agit donc de prélever une petite région d'une image puis de la comparer à l'ensemble des autres parties de cette image en tentant de les faire coïncider. Il serait exceptionnel que l'on obtienne une coïncidence parfaite; on estimera que deux régions sont identiques si la somme des différences de valeurs des pixels mis en coïncidence est inférieure à un seuil prédéfini. Plus ce seuil sera élevé moins la ressemblance des deux éléments sera précise mais plus on aura de chance de trouver des zones susceptibles de s'apparier; le taux de compression pourra donc être élevé mais on dégradera l'image.

Un logiciel simpliste permet de prélever un petit carré de l'image et de le faire glisser pixel par pixel et ligne à ligne pour le comparer à l'ensemble de l'image.

La procédure de recherche de coïncidence mise réellement en œuvre dans les logiciels de compression est en fait un peu plus complexe puisqu'elle utilise l'ensemble des transformations géométriques dites affines en ajoutant aux effets de translation mentionnés plus haut des effets de rotation et de changement d'échelle (ou homothétie) de la zone test. Il est ainsi possible de comparer tout d'abord un bloc par exemple de 64 x 64 pixels à des zones de 128 x 128 (processus rapide) puis de 64 x 64 puis de 32 x 32, puis de 16 x 16 puis enfin de 8 x 8, voire de 4 x 4 ou, pourquoi pas, de 2 x 2.

On conçoit que ces tentatives multiples pour apparier ces zones sont longues à effectuer.

Par contre, lorsque le catalogue des motifs qu'on a trouvés et dupliqués a été établi, et qu'on a simultanément la connaissance de l'endroit où ils ont été reconnus ainsi que les indications d'angle de rotation et de facteur d'échelle qu'il faut leur appliquer, il devient très facile de reconstruire l'image initiale. Compression et décompression, encore une fois ne sont pas symétriques et fort heureusement, le travail le plus simple est celui qui est effectué chez l'utilisateur.

On commence à disposer de logiciels de compression fractale. Compte tenu du fait que les motifs répertoriés - et la manière dont ils ont été choisis - sont parfaitement adaptés au contenu de l'image traitée, la compression fractale permet d'obtenir de manière virtuellement transparente des taux de compression sur images fixes de l'ordre de 100 ce qui est tout à fait intéressant aussi bien pour le stockage des images que pour leur transmission sur des lignes téléphoniques. On pense bien entendu à la presse, mais il ne faut pas oublier d'autres secteurs comme par exemple le médical, où ces besoins sont en rapide croissance.