

Compétences travaillées : 4.1, représentation de l'information – format et compression

## La compression des images

On comprend facilement que le fichier .ppm précédent occupe trop de place. Pour qu'il en occupe moins, il existe de nombreux moyens d'éviter de recopier 13 fois la même ligne. C'est ce que l'on appelle **compresser** les fichiers.

La taille de l'image est un paramètre important à considérer pour des raisons de stockage sur le disque dur (emplacement disponible) mais aussi lors de la construction d'un site Web. Lorsqu'on affiche des images sur une **page Web**, il faut se poser la question de la **durée du chargement des fichiers**.

### 1) En préalable, le calcul de la définition d'une image bitmap

Pour les images bitmap, il y a 3 paramètres :

- Le nombre de colonnes
- Le nombre de lignes
- Le nombre de couleurs par pixel

Les deux premiers paramètres donnent la **définition de l'image**, c'est à dire le **nombre total de pixels** (nombre de lignes \* nombre de colonnes)

Pour calculer **la taille en octets d'une image** Bitmap :

$$\underbrace{\text{Nb\_de\_colonnes} * \text{Nb\_de\_lignes}}_{= \text{Nb de pixels}} * \text{Nb\_de\_bits\_par\_pixel} / 8$$

### 2) Résumons

1. Pour quelques fichiers utilisés dans la séance précédente, compléter le tableau ci-dessous :

Fichier	Nombre de pixels	Nombre maximum de couleurs possibles	Nombre de bits par pixels	Définition de l'image	Taille du fichier en octets calculée	Taille du fichier en octets Indiquée

### 3) Principe général de la compression

- La plupart sont basées sur des formules mathématiques qui permettent de **traiter les redondances** (c'est à dire les éléments qui apparaissent plusieurs fois à l'identique) en les remplaçant par des codes plus courts. Ceci permet après une étape de décompression, de récupérer le fichier tel qu'il était initialement (compression sans perte, par exemple la compression ZIP utilisable sur tous types de données).
- Il existe également des méthodes de **compression avec perte**, appliquées à des types de données bien particuliers (sons, **images JPEG** ou vidéos). Le principe est d'éliminer les informations les moins perceptibles de manière à obtenir un résultat qui paraisse identique à l'original.

#### 4) Un exemple : La compression RLE sans perte (Run Length Encoding)

Cette compression est utilisée par les images bitmap (format .bmp notamment)

Elle est basée sur la répétition d'éléments consécutifs. Par exemple, la chaîne « AAAAHHHHHHHHHH » sera compressée en 4A10H soit un gain de compression de  $(14-5)/14 = 64,3\%$ . En revanche la chaîne REELLEMENT qui deviendrait 1R2E2L1E1M1E1N1T (soit -60% de gain de compression) ne sera pas compressée. En pratique seules les séquences comportant plus de 3 répétitions d'un même élément sont compressées, les autres sont conservées telles quelles, séparées des séquences compressées par un caractère spécial.

**La compression RLE n'a donc d'intérêt que pour des images possédant de nombreux éléments consécutifs redondants, notamment les images avec de larges parties uniformes. C'est ce qui explique que les images BMP atteignent fréquemment des tailles importantes.**

#### 4) Formats PNG, GIF et JPEG

2. **Rechercher les particularités des fichiers GIF. Quelle information particulièrement importante pour l'affichage de l'image le fichier doit-il contenir et qui n'était pas présente dans les formats vus précédemment ?**

3. **Rechercher les particularités des fichiers JPEG.**

4. **Ouvrez une image au format PNG.**
5. **Enregistrez-la sous le même nom au format JPEG en qualité de 40% (avec exporter sous)**

6. Ouvrir les deux images et les comparer à l'œil nu. Ne pas oublier de zoomer.
7. Faire de même avec une image JPEG avec une qualité de 20% puis de 90%.
8. Complétez dans le tableau ci-dessous vos observations :

Fichier	Taille en octets	Qualité perçue de l'image

9. Quel fichier envoyer avec un message électronique ?

10. Créer deux fichiers .pgm contenant 2 images de même taille, l'une unie et l'autre aléatoire. Compresser ces deux fichiers avec le programme gzip et comparer la taille des fichiers.

11. Les photographes utilisent le format RAW ? Pourquoi ?

12. Ce format existe avec plusieurs noms d'extension de fichiers. Pourquoi ?

## Les unités de grandeur

KiloOctets, MégaOctets ...

Longtemps l'informatique s'est singularisée par l'utilisation de différentes valeurs pour les unités du système international. Ainsi beaucoup d'informaticiens ont appris que 1 kilooctet valait 1024 octets. Or, depuis décembre 1998, l'organisme international *IEC* a statué sur la question (<http://physics.nist.gov/cuu/Units/binary.html>). Voici donc les unités standardisées :

- **Un kilooctet (ko ou kB) = 1000 octets**
- **Un Mégaoctet (Mo ou MB) = 1000 ko = 1 000 000 octets**
- **Un Gigaoctet (Go ou GB) = 1000 Mo = 1 000 000 000 octets**
- **Un Téraoctet (To) = 1000 Go = 1 000 000 000 000 octets**

Attention ! De nombreux logiciels (parfois même certains systèmes d'exploitation) utilisent toujours la notation antérieure à 1998 pour laquelle :

- Un kilooctet (ko) =  $2^{10}$  octets = 1024 octets
- Un Mégaoctet (Mo) =  $2^{20}$  octets = 1024 ko = 1 048 576 octets
- Un Gigaoctet (Go) =  $2^{30}$  octets = 1024 Mo = 1 073 741 824 octets
- Un Téraoctet (To) =  $2^{40}$  octets = 1024 Go = 1 099 511 627 776 octets

→ L'IEC a également défini le kilo binaire (kibi), le méga binaire (Mébi), le giga binaire (Gibi), le tera binaire (Tebi).

Voici leurs définitions :

- **Un kibioctet (kio ou kiB) vaut  $2^{10} = 1024$  octets**
- **Un Mébioctet (Mio ou MiB) vaut  $2^{20} = 1 048 576$  octets**
- **Un Gibioctet (Gio ou GiB) vaut  $2^{30} = 1 073 741 824$  octets**
- **Un Tébioctet (Tio ou TiB) vaut  $2^{40} = 1 099 511 627 776$  octets**

Il est également utile de noter que la communauté internationale dans son ensemble utilise préférentiellement le nom de « byte » plutôt que le terme « octet » purement francophone. Cela donne les notations suivantes pour kilobyte, mégabyte, gigabyte et terabyte : kB, MB, GB, TB

**13. J'ai acheté un disque dur de 500 Go mais sa capacité réelle est de 465 Go. Y a-t-il tromperie sur la marchandise ?**