



## La vidéo

### La vidéo

Informations obtenues sur le site <http://www.commentcamarche.net>

Cette fiche est assez technique mais ce degré de technicité est néanmoins nécessaire pour bien comprendre la vidéo.

### Qu'est-ce qu'une vidéo ?

Une vidéo est une succession d'images à une certaine cadence. L'œil humain a comme caractéristique d'être capable de distinguer environ 20 images par seconde. Ainsi, en affichant plus de 20 images par seconde, il est possible de tromper l'œil et de lui faire croire à une image animée. On caractérise la fluidité d'une vidéo par le nombre d'images par secondes (en anglais frame rate), exprimé en FPS (Frames per second, en français frames par seconde). D'autre part la vidéo au sens multimédia du terme est généralement accompagnée de son, c'est-à-dire de données audio.

### Vidéo numérique et analogique

On distingue généralement plusieurs familles d'«images animées»

- Le cinéma consiste à stocker sur une pellicule la succession d'images en positif. La restitution du film se fait alors grâce à une source lumineuse projetant les images successives sur un écran
- La vidéo analogique représente l'information comme un flux continu de données analogiques, destiné à être affichées sur un écran de télévision (basé sur le principe du balayage. Il existe plusieurs normes pour la vidéo analogique. Les trois principales sont : PAL, NTSC, SECAM
- La vidéo numérique consiste à coder la vidéo en une succession d'images numériques.

### PAL

Le format PAL/SECAM (Phase Alternating Line/Séquentiel Couleur avec Mémoire), utilisé en Europe pour la télévision hertzienne, permet de coder les vidéos sur 625 lignes (576 seulement sont affichées car 8% des lignes servent à la synchronisation) à raison de 25 images par seconde à un format 4:3 (c'est-à-dire que le rapport largeur sur hauteur vaut 4/3).

Or à 25 images par seconde, de nombreuses personnes perçoivent un battement dans l'image. Ainsi, étant donné



*Avec le soutien de la Commission européenne*

qu'il n'était pas possible d'envoyer plus d'informations en raison de la limitation de bande passante, il a été décidé d'entrelacer les images, c'est-à-dire d'envoyer en premier lieu les lignes paires, puis les lignes impaires. Le terme «champ» désigne ainsi la «demi-image» formée soit par les lignes paires, soit par les lignes impaires. L'ensemble constitué par deux champs est appelé trame entrelacé. Lorsqu'il n'y a pas d'entrelacement le terme de trame progressive est utilisé.

Grâce à ce procédé appelé «entrelacement», le téléviseur PAL/SECAM affiche 50 champs par seconde (à une fréquence de 50 Hz), soit 2x25 images en deux secondes.

## NTSC

La norme NTSC (National Television Standards Committee), utilisée aux Etats-Unis et au Japon, utilise un système de 525 lignes entrelacées à 30 images/sec (donc à une fréquence de 60Hz). Comme dans le cas du PAL/SECAM, 8% des lignes servent à synchroniser le récepteur. Ainsi, étant donné que le NTSC affiche un format d'image 4:3, la résolution réellement affichée est de 640x480.

## La vidéo numérique

La vidéo numérique consiste à afficher une succession d'images numériques. Puisqu'il s'agit d'images numériques affichées à une certaine cadence, il est possible de connaître le débit nécessaire pour l'affichage d'une vidéo, c'est-à-dire le nombre d'octets affichés (ou transférés) par unité de temps.

Ainsi le débit nécessaire pour afficher une vidéo (en octets par seconde) est égal à la taille d'une image que multiplie le nombre d'images par seconde.

Soit une image true color (24 bits) ayant une définition de 640 pixels par 480. Pour afficher correctement une vidéo possédant cette définition il est nécessaire d'afficher au moins 30 images par seconde, c'est-à-dire un débit égal à :

$$900 \text{ Ko} * 30 = 27 \text{ Mo/s}$$

## Sous-échantillonnage

Etant donné que l'œil est peu sensible aux variations de chrominance, la technique dite de sous-échantillonnage en chrominance (en anglais chroma subsampling), appelée également décimation, consiste à supprimer des informations de chrominance dans un groupe de 4x4 pixels.

## Notion de codec

Une image d'une vidéo non compressée occupe une taille d'environ 1 Mo. Afin d'obtenir une vidéo paraissant fluide il est nécessaire d'avoir une fréquence d'au moins 25 ou 30 images par seconde, ce qui produit un flux de données d'environ 30 Mo/s, soit plus de 1.5 Go par minute. Il est évident que ce type de débit est peu compatible avec les espaces de stockage des ordinateurs personnels ni même avec les connexions réseau de particuliers ou de petites ou moyennes entreprises.

Ainsi, afin de pallier à cette difficulté, il est possible de recourir à des algorithmes permettant de réduire significativement les flux de données en compressant / décompressant les données vidéos. On appelle ces algorithmes CoDec (pour COmpression / DECompression).

## Le M-JPEG

La première idée qui vient à l'esprit après s'être intéressé à la compression d'images est d'appliquer . Le principe du Motion JPEG (noté MJPEG ou M-JPEG, à ne pas confondre avec le MPEG) consiste à appliquer successivement l'algorithme de compression JPEG aux différentes images d'une séquence vidéo.

Etant donné que le M-JPEG code séparément chaque image de la séquence il permet d'accéder aléatoirement à n'importe quelle partie d'une vidéo. Ainsi son débit de 8 à 10 Mbps le rend utilisable dans les studios de montage numérique.

## Le MPEG

Dans de nombreuses séquences vidéos, de nombreuses scènes sont fixes ou bien changent très peu, c'est ce que l'on nomme la redondance temporelle.

Lorsque seules les lèvres de l'acteur bougent, presque seuls les pixels de la bouche vont être modifiés d'une image à l'autre, il suffit donc de décrire le changement d'une image à l'autre. C'est là la différence majeure entre le MPEG (Moving Pictures Experts Group) et le M-JPEG. Cependant cette méthode aura beaucoup moins d'impact sur une scène d'action.

Le groupe MPEG a été établi en 1988 dans le but de développer des standards internationaux de compression, décompression, traitement et codage d'image animées et de données audio.

Il existe plusieurs standards MPEG :

- **le MPEG-1**, développé en 1988, est un standard pour la compression des données vidéos et des canaux audio associés (jusqu'à 2 canaux pour une écoute stéréo). Il permet le stockage de vidéos à un débit de 1.5Mbps dans une qualité proche des cassettes VHS sur un support CD appelé VCD (Vidéo CD).
- **le MPEG-2**, un standard dédié originalement à la télévision numérique (HDTV) offrant une qualité élevée à un débit pouvant aller jusqu'à 40 Mbps, et 5 canaux audio surround. Le MPEG-2 permet de plus une identification et une protection contre le piratage. Il s'agit du format utilisé par les DVD vidéos.
- **le MPEG-4**, un standard destiné à permettre le codage de données multimédia sous formes d'objets numériques, afin d'obtenir une plus grande interactivité, ce qui rend son usage particulièrement adapté au Web et aux périphériques mobiles.
- **le MPEG-21**, en cours d'élaboration, dont le but est de fournir un cadre de travail (en anglais framework) pour l'ensemble des acteurs du numérique (producteurs, consommateurs, ...) afin de standardiser la gestion de ces contenus, les droits d'accès, les droits d'auteurs, ...

## Le MPEG-1

La norme MPEG-1 représente chaque image comme un ensemble de blocs 16 x 16. Elle permet d'obtenir une résolution de:

352x240 à 30 images par seconde en NTSC

352x288 à 25 images par seconde en PAL/SECAM

Le MPEG-1 permet d'obtenir des débits de l'ordre de 1.2 Mbps (exploitable sur un lecteur de CD-ROM).

Le MPEG-1 permet d'encoder une vidéo grâce à plusieurs techniques :

- Intra coded frames (Frames I, correspondant à un codage interne): les images sont codées séparément sans faire référence aux images précédentes
- Predictive coded frames (frames P ou codage prédictif): les images sont décrites par différence avec les images précédentes
- Bidirectionally predictive coded frames (Frames B): les images sont décrites par différence avec l'image précédente et l'image suivante
- DC Coded frames: les images sont décodées en faisant des moyennes par bloc

### Les frames I

Ces images sont codées uniquement en utilisant le codage JPEG, sans se soucier des images qui l'entourent. De telles images sont nécessaires dans une vidéo MPEG car ce sont elles qui assurent la cohésion de l'image (puisque les autres sont décrites par rapport aux images qui les entourent), elles sont utiles notamment pour les flux vidéo qui peuvent être pris en cours de route (télévision), et sont indispensables en cas d'erreur dans la réception. Il y en a donc une ou deux par seconde dans une vidéo MPEG.

### Les frames P

Ces images sont définies par différence par rapport à l'image précédente. L'encodeur recherche les différences de l'image par rapport à la précédente et définit des blocs, appelés macroblocs (16x16 pixels) qui se superposent à l'image précédente.

L'algorithme compare les deux images bloc par bloc et à partir d'un certain seuil de différence, il considère le bloc de l'image précédente différent de celui de l'image en cours et lui applique une compression JPEG.

C'est la recherche des macroblocs qui déterminera la vitesse de l'encodage, car plus l'algorithme cherche des «bons» blocs, plus il perd de temps...

Par rapport aux frames-I (compressant directement), les frames-P demandent d'avoir toujours en mémoire l'image précédente.

### **Les frames B**

De la même façon que les frames P, les frames B sont travaillées par différences par rapport à une image de référence, sauf que dans le cas des frames B cette différence peut s'effectuer soit sur la précédente (comme dans les cas des frames P) soit sur la suivante, ce qui donne une meilleure compression, mais induit un retard (puisque'il faut connaître l'image suivante) et oblige à garder en mémoire trois images (la précédente, l'actuelle et la suivante).

### **Les frames D**

Ces images donnent une résolution de très basse qualité mais permettent une décompression très rapide, cela sert notamment lors de la visualisation en avance rapide car le décodage «normal» demanderait trop de ressources processeur.

### **Dans la pratique...**

Afin d'optimiser le codage MPEG, les séquences d'images sont dans la pratique codées suivant une suite d'images I, B, et P (D étant comme on l'a dit réservé à l'avance rapide) dont l'ordre a été déterminé expérimentalement. La séquence type appelée GOP (Group Of Pictures ou en français groupes d'images) est la suivante:

IBBPBBPBBPBI

Une image I est donc insérée toutes les 12 frames.