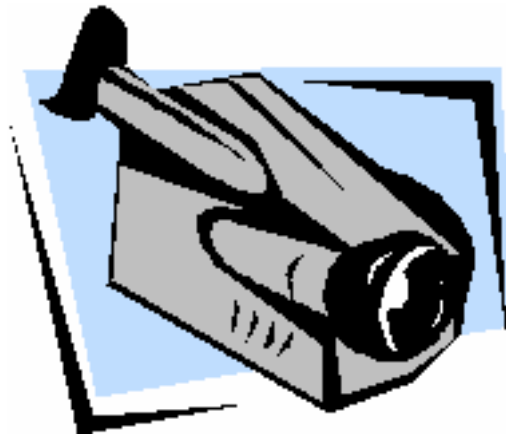


# INTRODUCTION À LA VIDÉO NUMÉRIQUE



Gilles Boulet  
<http://gillesboulet.ca>  
Septembre 2001

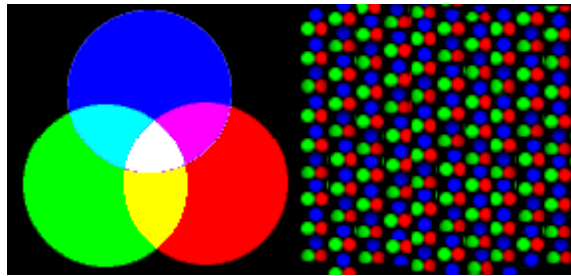
# 1. La vidéo

## 1.1 La vision

Dans le phénomène de la vision humaine, un flux lumineux composé de photons frappe l'œil. Dans l'œil, les bâtonnets vont réagir à l'intensité de la lumière (la luminance ou Y) et les cônes vont réagir à la couleur (la chrominance ou C). Le cerveau effectue la synthèse de ces informations pour composer une image. L'œil humain est davantage sensible à l'intensité lumineuse (Y) qu'à la couleur (C).

## 1.2 La couleur

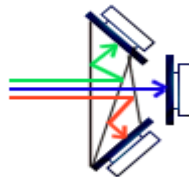
La vidéo utilise ce qui est identifié comme la synthèse additive de la couleur. Dans le système colorimétrique additif, les trois couleurs primaires sont le rouge, le vert et le bleu (RVB ou RGB dans la terminologie anglophone). En combinant ces trois couleurs primaires, il est possible de reproduire tout le spectre visible par l'œil humain. L'écran de visualisation vidéo sera donc composé d'une série de triades rouge-vert-bleu. L'activation de l'ensemble de ces triades formera l'image.



## 1.3 Le signal

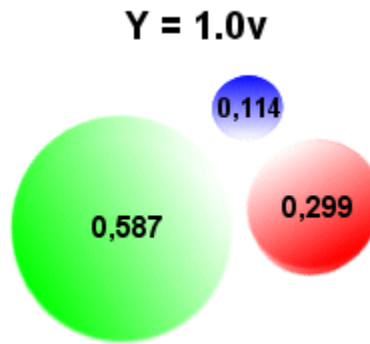
En vidéo, c'est la caméra qui transforme l'information lumineuse (photons) en signal électrique (électrons). En vidéo analogique, l'intensité de ce signal électrique varie de façon continue. Le processus de transformation de l'information lumineuse en signal électrique est le suivant :

- L'élément optique de la caméra, l'objectif, sépare la lumière en trois composants le rouge, le vert, le bleu. Cette opération est réalisée en faisant passer le flux lumineux par une succession de filtres dichroïques réfléchissant certaines couleurs et en laissant passer d'autres.
- En vidéo professionnelle, les trois images sont projetées sur trois capteurs photosensibles distincts formés chacun de centaines de milliers de points, généralement entre 400 000 et 700 000. Ces capteurs sont nommés CCD (charged coupled device) ou dispositifs à transfert de charge. Les caméras domestiques ne sont généralement équipées que d'un seul capteur CCD.



- Pour chacun des points de chacun des capteurs, l'énergie lumineuse sera transformée en énergie électrique. Ainsi, à la sortie des capteurs, trois signaux électriques d'intensité variable, un signal pour chacune des trois composantes. La lumière blanche est formée par la somme des trois composantes RVB. Toutefois, elle n'est pas composée des trois couleurs primaires en quantité égale. Les proportions sont les suivantes :

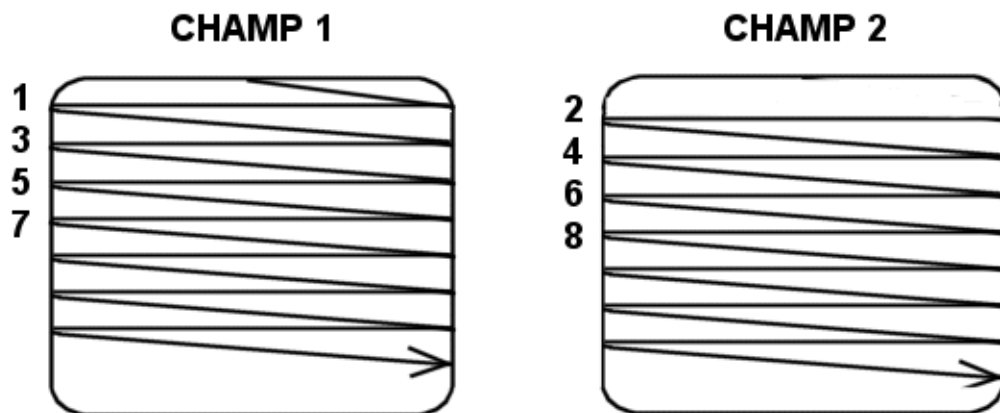
$$100\% Y = 29,9\%(R) + 58,7\%(V) + 11,4\%(B)$$



- Une information supplémentaire devra être ajoutée à ces signaux. Comme nous l'avons dit, une valeur doit être établie pour chacun des points des capteurs. C'est cette valeur pour chacun des points de chacun des pixels qui permettra de recomposer l'image sur un écran de visualisation, chacun des points de la surface d'affichage pouvant alors prendre la même valeur qu'au moment de la capture par la caméra. La restitution de l'information pour chacun des pixels se fait par balayage ligne par ligne de la surface de l'écran de reproduction, les valeurs enregistrées pour chacun des pixels au moment de la capture étant transférées aux pixels de la surface de reproduction. Pour synchroniser ces balayages, un signal de fin de ligne est ajouté pour chacune des lignes. Un signal de fin d'image est également inséré.

#### 1.4 L'affichage

La technique de la composition d'image par balayage a une conséquence : contrairement à l'image cinéma, l'image vidéo ne s'affiche pas d'un seul coup, elle s'affiche ligne par ligne, en commençant par le haut à gauche pour finir en bas à droite. Afin d'assurer une uniformité dans la luminosité de cet affichage, en d'autres termes afin que l'intensité lumineuse des pixels du haut de l'écran ne commence pas à diminuer alors que ceux du bas viennent d'être activés, une technique spécifique a été développée. Il s'agit de la technique d'entrelacement. Un premier balayage est effectué sur les lignes impaires, puis un deuxième sur les lignes paires. Il faut donc deux passages pour recomposer une image complète. La fréquence d'entrelacement s'appuie sur la fréquence du réseau électrique. Comme en Amérique du Nord la fréquence des réseaux est de 60Hz, le système de balayage est donc de 60 demi images, soit 30 images par seconde. De façon précise, le taux de rafraîchissement est plutôt de 29,97 images par seconde.



La norme Nord-Américaine en télévision est le NTSC (National Television Standard Committee). L'image NTSC est formée de 525 lignes. De ces lignes, seulement 486 sont utilisées pour afficher l'image. Les deux autres principales normes sont européennes. Il s'agit des normes PAL (Phase Alternation by Line) et SECAM (Séquentiel Couleur à Mémoire). L'image PAL ou SECAM est formée de 625 lignes dont 576 sont utilisées pour l'image. Comme la fréquence des réseaux électriques européens est de 50 Hz, le système de balayage est de 50 demi images, soit 25 images par seconde.

### 1.5 Le son

Le son fera l'objet d'un traitement distinct. Le signal sonore sera distinct du signal vidéo. Il sera également enregistré sur des pistes distinctes sur le ruban vidéo.

### 1.6 Les formats d'enregistrement

En vidéo analogique, les signaux vidéo et audio seront enregistrés sur ruban magnétoscopique. Il existe différents formats d'enregistrement.

- Le format RVB  
(Format DVD vidéo)  
Les trois informations pour chacune des trois couleurs sont encodées distinctement, véhiculées sur des câbles séparés et accompagnées d'un signal de synchronisation.
- Le format composante (ou YUV ou Y, B-Y, R-Y ou Y, Cr, Cb)  
(Formats Betacam, M-II)  
Les informations de luminance (noir et blanc) (Y) et les informations de chrominance (couleur, R-Y et B-Y) sont enregistrées sur des pistes différentes. Lors de l'enregistrement, on diminue la quantité d'information concernant la couleur. Pour y arriver, puisque nous savons que la luminance (Y) est composée de la somme des 3 chrominances (rouge, vert et bleu), il est donc possible de supprimer une des composantes couleur et de la reconstruire à partir de la valeur Y et des deux autres valeurs, le bleu (B) et le rouge (R) :

$$Y = (0,299 \times R + 0,587 \times V + 0,114 \times B) \times 1$$

$$B - Y = (-0,299 \times R - 0,587 \times V + 0,886 \times B) \times 0,306 = U$$

$$R - Y = (-0,299 \times R - 0,587 \times V + 0,886 \times B) \times 0,877 = V$$

- Le format composite  
(Formats VHS, 8mm, U-Matic ou 3/4 po)  
Les informations de luminance (noir et blanc) et de chrominance (couleur) sont combinées en un seul signal, transportées sur un seul câble et enregistrées sur une même piste magnétique.
- Le format S-Video (Y/C)  
(Formats S-VHS, Hi-8)  
Format également composite, il transmet les informations de luminance (Y) sur un câble et combine en un seul signal les informations de chrominance (U, V). Les deux signaux seront, par la suite, enregistrés sur une même piste magnétique.

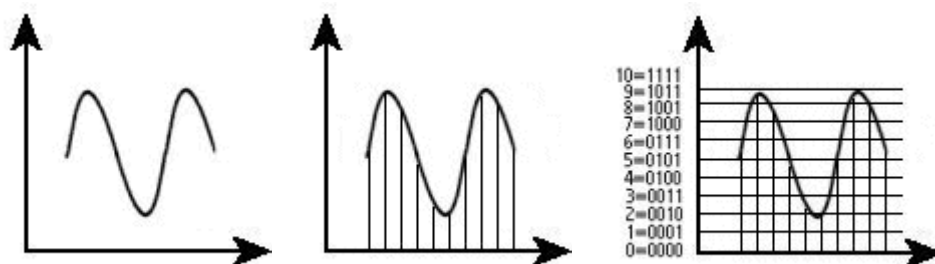
## 1.7 La vidéo numérique

Le processus de captation de l'image vidéo en mode numérique est essentiellement le même que pour la vidéo analogique. Un système optique sépare la lumière en trois composants. À la sortie toutefois, plus d'enregistrement d'un signal électrique mais plutôt l'enregistrement d'une valeur numérique définie pour chacune des trois couleurs de chacun des pixels (RVB). Le processus de transformation d'un signal électrique et valeurs numériques s'appelle la numérisation ou l'échantillonnage. Deux paramètres servent à échantillonner un signal électrique : la fréquence d'échantillonnage et la quantification de l'échantillonnage ou le nombre de bits utilisés pour le codage.

- La fréquence d'échantillonnage

Il existe un théorème appelé le théorème de Shannon. Ce théorème stipule que « l'information véhiculée par un signal dont le spectre est à support borné n'est pas modifiée par l'opération d'échantillonnage à condition que la fréquence d'échantillonnage soit au moins deux fois plus grande que la plus grande fréquence contenue dans le signal. La reconstitution du signal original peut être effectuée par un filtre passe-bas idéal de fréquence de coupure égale à la moitié de la fréquence d'échantillonnage.»

La bande passante d'un signal vidéo peut atteindre 6Mhz. Pour numériser ce signal, une fréquence d'échantillonnage de 13,5Mhz sera donc utilisée. En d'autres termes, le signal vidéo sera mesuré 13 500 000 fois par seconde. Comme nous savons que l'œil est plus sensible à la lumière (luminance) qu'à la couleur (chrominance), souvent, la fréquence d'échantillonnage de la chrominance sera moins importante. Dans les faits, l'on utilise un échantillonnage à 13,5Mhz pour Y (luminance) et 6,75 pour C (U et V). Cet échantillonnage est donc un échantillonnage à 16 bits par pixel.



- La quantification de l'échantillonnage

Un certain nombre de bits serviront à coder chacune des mesures. Nous savons que l'œil humain discerne un maximum de 256 niveaux différents de luminance. Ainsi, un codage à 8 bits est suffisant pour reproduire l'ensemble des variations perceptibles ( $2^8 = 256$  combinaisons possibles) Comme chacune des composantes couleur de l'image sera quantifiée sur 8 bits (8 bits Rouge, 8 bits Vert, 8 bits Bleu) ( $2^{24} = 16\,777\,216$  combinaisons possibles) un total de 24 bits (3 octets) seront utilisés pour coder les informations relatives à un seul pixel d'une image.

## 1.8 Encombrement

Une image vidéo en format NTSC (National Television Standard Committee) a une résolution horizontale de 720 pixels et une résolution verticale de 486 pixels. Si nous utilisons un échantillonnage 4:2:2 de 16 bits par pixel, nous constatons qu'une seule minute de vidéo numérique non compressée nécessitera près de 1,2 Go d'espace disque.

- **Petite mathématique :**

En mode binaire 1 kilo = 1024 ( $2^{10}$ )

1 kilo octet (Ko) ou kilo byte (Kb) = 1024 bits

1 méga octet (Mo) ou mega byte (MB) =  $2^{20}$  octets = 1,048,576 octets = 8 388 608 bits (1 octet = 8 bits)

1 giga octet (Go) ou giga byte (GB) =  $2^{30}$  octets = 1,073,741,824 octets = 8,589,934,592 bits

En informatique, le préfixe multiplicatif peut avoir deux sens, selon la base utilisée, une base 10 ou une base 2. La base 10 correspond au Système International (SI).

Kilo  $2^{10} = 1024$  octets

Méga  $2^{20} = 1\,048\,576$  octets

Giga  $2^{30} = 1\,073\,741\,824$  octets

Tera  $2^{40} = 1\,099\,511\,627\,776$  octets

Peta  $2^{50} = 1\,125\,899\,906\,842\,624$  octets

Exa  $2^{60} = 1\,152\,921\,504\,606\,846\,976$  octets

Zetta  $2^{70} = 1\,180\,591\,620\,717\,411\,303\,424$  octets

Yotta  $2^{80} = 1\,208\,925\,819\,614\,629\,174\,706\,176$  octets

### **Vidéo :**

Résolution - 720 x 486 x 29.97 images par seconde (ips) (720 x 243 x 59.94 champs par seconde)

Échantillonnage- 4:2:2 ( 2 Y : 1Cr, 1Cb )

Luminance (Y) 720 x 486 x 29.97 Ips = 10,487,102.4 octets/sec x 8 bits par octet = 83,896,819.2 bits/sec

Chrominance R (Cr) 360 x 486 x 29.97 Fps = 5,243,551.2 octets/sec x 8 bits par octet = 41,948,409.6 bits/sec

Chrominance B (Cb) 360 x 486 x 29.97 Fps = 5,243,551.2 octets/sec x 8 bits par octet = 41,948,409.6 bits/sec

Total = 20,974,204.8 octets/ sec = 167,793,638.4 bits/sec

Mo/sec 20,974,205 / 1,048,576 = 20.00256062 = 20 Mo/sec

20Mo/sec x 60 sec = 1200Mo/min

1024Mo par giga octet = 1,17Go/min

Ainsi, une minute de vidéo aura un poids de 1,17giga octets

### **Audio :**

Échantillonnage 16 bits (2 octets) sur 2 pistes audio

Fréquence d'échantillonnage 44,100 (qualité CD audio)

Un canal- 2 octets par échantillon x 44,100 échantillons par seconde = 88,200 octets/sec

88,200 octets/sec x 2 canaux = 176,400 octets/sec

0.17 Mo/s X 60 sec = 10Mo/min

1024Mo par giga octet = 0.01Go/min

### **Vidéo et audio :**

20Mo/sec + 0,17Mo/sec = 20,17Mo/sec

1,17Go/min + 0,01Go/min = 1,18Go/min

## 1.9 Débit de données

Alors que, au niveau de la taille des fichiers, la quantification repose sur des données binaires, lorsque l'on parle de débit ou de taux de transfert, alors ces données sont décimales. Ainsi un mégabit par seconde signifiera un million de bits par seconde et un mégaoctet signifiera un million d'octets ou 8 millions de bits puisqu'il y a 8 bits dans un octet.

Les abréviations généralement reconnues sont les suivantes :

Kbps kilobits par seconde = 1000 bits  
Kops kilo octets par seconde = 8000 bits  
Mbps mégabits par seconde = 1 000 000 bits  
Mops mégaoctets par seconde = 8 000 000 bits

Sans entrer dans une discussion trop technique sur le débit et le taux de transfert, mentionnons simplement qu'un signal vidéo NTSC plein écran non compressé et échantillonné 4:2:2 nécessite un débit d'un peu plus de 20 mégaoctets par seconde (20Mops) pour pouvoir être affiché correctement. Comme point de comparaison, soulignons qu'un disque DVD vidéo a un débit maximal possible de 9,8 Mops, un cédérom 36X un débit de 5 Mops, un lien Internet par câble-modem un débit de 2Mops en réception, une liaison LNPA un débit de 1Mops alors que le débit d'une liaison modem par ligne téléphonique pourra varier de 14 à 56 kbps, en fonction bien sûr de la vitesse des modems en communication.

Il est donc aisé de comprendre qu'il sera nécessaire de traiter le signal vidéo, en fonction du mode de diffusion prévu. Il existe une cascade de traitements possibles. Nous les verrons un peu plus loin. Une constante toutefois, plus l'on réduit la quantité d'information, plus la qualité objective puis subjective du document est affectée.

## 2. Le format DV

Il est aisé de comprendre que, si la capture a été réalisée à l'aide d'une technologie vidéo numérique, le fichier pourra être directement transféré sur disque rigide d'ordinateur pour être traité.

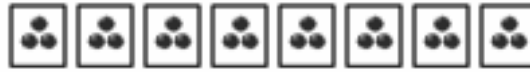
Si la capture a été réalisée à l'aide d'une technologie analogique, il faudra alors numériser le signal, soit traduire l'information électrique en information numérique. Pour réaliser cette opération, il faut équiper l'ordinateur d'un convertisseur analogique-numérique aussi identifié comme une carte d'acquisition vidéo.

### 2.1 Le format DV :

Afin de permettre que les fichiers et les supports pour les enregistrer demeurent de taille raisonnable, les formats d'enregistrement vidéo numérique effectuent une opération de compression des données dans la caméra, au moment de la prise de vue. Il existe plusieurs formats de vidéo numérique. Le format sans le plus répandu en ce moment est le format DV pour *Digital Video*. Ce format est fondé sur une norme internationale créée par un consortium de 10 sociétés en 1993. Il existe des variantes du format DV, par exemple le format DVCAM ou DVCPRO. Mais les différences sont davantage liées au processus d'enregistrement des données sur un support physique qu'au processus de traitement et de compression du signal.

L'enregistrement se déroule selon les séquences suivantes :

- L'élément optique de la caméra sépare la lumière vers trois composants : rouge, vert et bleu (RVB) et transmet l'information à un convertisseur. Le taux de transfert de la donnée est alors de 248Mops.



4:4:4

- Un convertisseur transforme le signal RGB en signal YUV (Y pour la luminance, U et V pour la couleur). Comme l'oeil humain perçoit mieux la lumière que la couleur, lors de l'échantillonnage chaque valeur Y d'un pixel est échantillonnée quatre fois, alors que les valeurs U et V ne le sont que deux fois, ce que l'on nomme la compression YUV 4:2:2. Ce premier échantillonnage réduit la taille de la donnée d'un tiers sans dégradation perceptible par l'oeil. Le taux de transfert est alors réduit à 165Mops.



4:2:2

- Afin de réduire encore davantage le volume d'information, on utilise un convertisseur YUV 4:2:2 - YUV 4:2:0/YUV 4:1:1. Chaque pixel du signal résultant garde sa même valeur de lumière (Y), alors que quatre pixels se partagent la même valeur de couleur (U et V). Le débit est alors de 124Mops.



4:1:1

- A ce stade, la donnée est réduite de deux tiers de sa taille initiale. On applique alors la compression DV au signal. Cette compression réduit le flux des quatre cinquièmes ( compression 5:1 ) jusqu'à un taux de transfert de 25Mops. C'est ce dernier résultat qui sera enregistré sur la bande magnétique.



### 3 Grammaire

L'image est l'unité de base en cinéma et en vidéo.

Le plan est, quant à lui, constitué d'un enchaînement d'images

La scène est un ensemble de plans situés dans un même temps, dans un même décor, et constituant un fragment de récit cohérent.

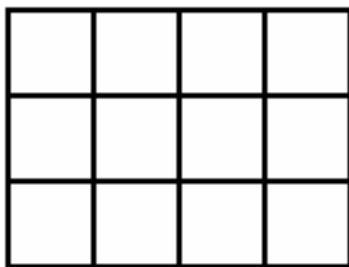
La séquence est une suite de scènes qui peuvent se situer dans des temps et des décors différents, mais qui constituent un sous-ensemble cohérent du récit.

#### 3.1 L'image :

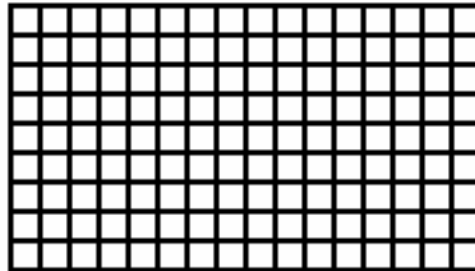
Une image vidéo possède un ratio d'affichage et une résolution couleur.

- Ratio d'affichage  
En vidéo, le format d'affichage le plus universel est le format 4:3 soit 4 unités de largeur par 3 unités de hauteur. L'écran est le plus fréquemment désigné par sa diagonale. Ainsi un écran de 15 pouces aura 12 pouces de largeur et 9 pouces de hauteur.

Un autre ratio existe. Celui de la télévision haute définition qui est de 16 unités de largeur par 9 unités de hauteur (16:9). Il s'agit d'un format qu'on dit plus proche de la vision humaine normale. Il s'agit du format retenu pour la télévision à haute définition. Ce format a aussi été retenu car il s'agit d'un multiple du format 4:3.



Format 4:3



Format 16:9

- Résolution des couleurs :  
Il s'agit du nombre de couleurs qui sont affichées simultanément à l'écran. En vidéo analogique, le nombre de couleurs affichable est infini puisque le signal est un flux électrique constant. En vidéo numérique, le nombre de couleur affichable est un nombre fini puisqu'il dépend du nombre de bits qui sera consacré à chacun des pixels de l'image. Ainsi, dans un agencement de couleurs RVB (rouge, vert, bleu), un codage 8 bits par pixel entraînera une résolution de 256 couleurs alors qu'un codage 24 bits permettra d'obtenir une résolution de plus de 16,7 millions de couleurs. L'affichage 24 bits est parfois appelé affichage en couleurs réelles car l'oeil humain ne distingue qu'environ 6 millions de teintes chromatiques différentes, soit moins que n'en offre un ensemble de couleurs de 24 bits. Évidemment, à mesure que le nombre de couleurs augmente, la taille du fichier s'accroît considérablement.

Codage 8 bits =  $2^8$  possibilités de valeurs, soit 256 couleurs possibles

Codage 16 bits =  $2^{16}$  possibilités de valeurs, soit 65 536 couleurs possibles

Codage 24 bits =  $2^{24}$  possibilités de valeurs, soit 16 777 216 couleurs possibles

### 3.2 Le plan :

Un plan est composé d'une série d'images fixes qui sont projetées à un rythme suffisamment rapide pour donner au spectateur l'illusion de mouvement en se fondant sur le phénomène de la persistance rétinienne de l'œil. Les images constitutives du plan sont donc projetées à un certain rythme. On parle alors de rythme d'enchaînement ou cadence de défilement. À moins de 18 images par seconde, l'œil commence à percevoir le scintillement et les mouvements commencent à apparaître saccadés. Le cinéma utilise une cadence de 24 images par seconde. Le format NTSC 525/60 (lignes /trames) utilisé avec le codage couleur est, quant à lui, fondé sur une cadence de 59.94 trames, soit à 29.97 images par seconde - une différence de une sur mille.

Il existe, en télévision, une convention d'identification de chacune des images d'une vidéo. Cette convention est le code temporel (time code) SMPTE (Society of Motion Pictures and Television Engineers). Ce code, lui identifie 30 images/sec. Le "Drop-frame time code" établit une compensation en omettant de numéroter deux images à chaque minute, sauf la dixième. Cette convention est identifiée comme celle du drop-frame. Il est à noter que le système PAL est, quant à lui, exact et ne nécessite pas de "drop-frame".

Pour créer un mouvement ralenti il faut enregistrer un plus grand nombre d'images dans une seconde, puis les projeter à rythme normal, par exemple enregistrer 60 images d'un mouvement en une seconde pour par la suite les projeter au rythme de 30 images secondes. Inversement, si la capture est effectuée à 15 images secondes puis projetée à 30, il y aura accéléré.

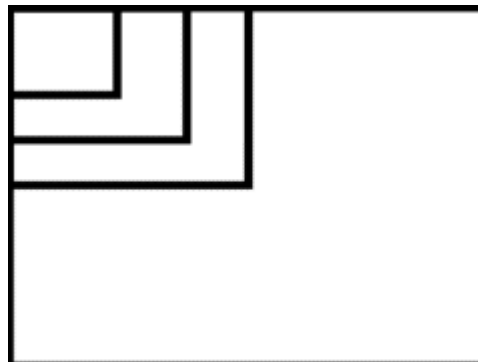
## 4. Ajustement du débit au média de diffusion

La séquence vidéo enregistrée pourra être diffusée sur support physique (cédérom ou dévédérom) ou sur réseau (LAN, Internet). Nous avons vu que les capacités des techniques des différents médias varient grandement. Pour diffuser de la vidéo numérique, il est donc nécessaire d'ajuster le débit au canal de diffusion c'est à dire de diminuer au maximum le nombre de bits ou d'octets utilisés pour représenter une image. Il existe différentes solutions pour ce faire.

La première solution est de **réduire la taille de la fenêtre** ce qui aura pour conséquence de réduire le nombre de pixels.

Les principales tailles d'écran utilisées sont :

640 x 480  
320 x 240  
240 x 180  
160 x 120



Une fenêtre de 320 x 240 occupera le quart de l'espace d'une fenêtre de 640 x 480. Conséquemment la taille du fichier et le taux de transfert nécessaire pour sa lecture seront considérablement diminués.

La deuxième solution est de **diminuer le nombre d'images par secondes**. La vidéo est projetée à une cadence de 30 images par seconde. En diminuant la cadence, on réduit encore la taille du fichier et le taux de transfert. À moins de 15 images par secondes toutefois, l'œil commencera à capter une saccade dans le mouvement.

La troisième possibilité est de **coder moins d'informations couleur**, en utilisant par exemple une palette de 256 couleurs. Cette solution occasionnera, elle aussi, dégradation de la qualité objective et subjective de l'image.

La solution la plus efficace au problème de débit est toutefois apportée par la **compression**, aussi appelée **réduction du débit binaire** (Bit Rate Reduction).

## 5. La compression

La vidéo numérique est l'un des applications informatiques les plus gourmandes tant en terme d'espace requis pour son enregistrement que de volume de débit et de bande passante pour sa diffusion. C'est pourquoi tout type d'application vidéo numérique requiert l'application d'algorithmes de compression.

La compression est le processus par lequel des fichiers volumineux sont réduits en taille en supprimant de ces fichiers toutes les informations redondantes. La compression va donc consister à déterminer ces redondances et à les éliminer tout en préservant la capacité de reproduire les images originales intactes ou avec le moins de dégradation possible. Puisque la vidéo est un média dynamique, c'est à dire que les images défilent les unes après les autres dans le temps, on peut distinguer deux types de redondances :

### 5.1 La redondance spatiale

La compression par redondance spatiale consiste à éliminer les informations similaires dans des zones de l'image proches l'une de l'autre. Par exemple, dans le cas d'une personne interviewée, l'arrière scène ou le fond de l'image pourra être fixe et d'une seule couleur. Par conséquent les valeurs d'une grande partie des pixels de l'image seront semblables les uns aux autres. Dans un tel cas, plutôt que d'assigner une valeur à chacun des pixels, le codec généralisera l'information à une région de l'image plutôt qu'à chaque pixel individuellement. C'est pourquoi, au moment de la production, plus l'arrière scène est fixe et uniforme, plus la compression sera efficace. Par voie de conséquence, un panoramique sur une forêt de feuillus d'automne aux arbres agités par le vent sera à peu près impossible à compresser.

### 5.2 La redondance temporelle

Alors que la redondance spatiale s'applique à éliminer les informations redondantes à l'intérieur d'une même image, la redondance temporelle travaille sur un ensemble d'images. Dans un plan, de nombreuses informations telles par exemple les formes ou les couleurs, sont redondantes d'une image à l'autre. Pour effectuer la compression, l'algorithme détermine une image clé (image I) et, par rapport à cette image, l'algorithme ne décrit que les informations divergentes dans les images situées après l'image clé (image P ou image prédite) et d'une image bi-directionnelle (image B), recomposée à partir des images I et P qui l'entourent.

### 5.3 MPEG

Le Moving Pictures Expert Group a été fondé en 1987. Ce groupe a pour objectif de définir des normes internationales pour la compression numérique des images vidéo. Situation particulière, le MPEG est à la fois une norme, un codec et un format de diffusion.

## 5.4 Principe de compression

La compression *MPEG* est basée sur l'extrapolation des informations. Elle utilise la redondance spatiale et la redondance temporelle. L'algorithme de compression enregistre tout d'abord le fond de l'image, et ce, toutes les 12 images. Ensuite, il n'enregistre que les déplacements des objets dans les inters images. Elle le fait de la façon suivante :

- l'image I (*Intra-frame*):  
à toutes les 12 images, une image est enregistrée. Elle représente l'image de référence. Elle est codée indépendamment comme une image fixe. Elle est compressée suivant au format *JPEG*.
- l'image P (*Predicted frame*):  
est une image calculée comme étant la différence entre l'image originale correspondante et l'image P ou I précédente.
- l'image B (*Bi-directional frame*):  
est une image intermédiaire fortement compressée calculée comme étant la différence entre l'image originale correspondante et la moyenne entre deux images P ou I, dont une est située en avant et l'autre en arrière.

## 5.5 MPEG 1 à MPEG 4

- MPEG 1  
Après 5 ans d'études, la norme MPEG 1 fut lancée en 1992. Elle est principalement utilisée pour la compression du son et de l'image sur support cédérom. Grâce à leur technique, les concepteurs de MPEG 1 réussissent à atteindre un taux de compression de 26:1. La norme convient assez bien pour des vidéos qui peuvent se contenter de débits peu élevés, entre 1 et 1.5Mbits/s.
- MPEG 2  
La norme MPEG 2 est apparue en 1994. Elle permet d'obtenir une qualité d'image professionnelle. La norme permet de supporter des débits allant de 4 à 20 Mbits/s. Elle est utilisée pour la diffusion par satellite, la télédistribution, la télévision haute définition et le DVD.
- MPEG 3  
La norme MPEG 3 devait s'appliquer à des médias nécessitant des débits de 20 à 40Mbits/s. Elle visait les applications relatives à la télévision numérique. Les paramètres de la norme MPEG 3 ont été englobés dans la norme MPEG 2. La norme MPEG 3 ne sera jamais développée.
- MPEG 4  
La norme MPEG 4 s'applique à des médias pouvant être diffusés à faible débit, c'est à dire un débit inférieur à 64Kbits/s.

D'autres versions de la norme MPEG sont actuellement en développement, entre autre la norme MPEG 7 qui permettra l'indexation du contenu de la séquence dans le signal vidéo lui-même.

## 6. Distribution et diffusion

La vidéo numérique peut être distribuée ou diffusée.

Pour la **distribuer**, elle est généralement enregistrée sur support physique tel un disque cédérom ou dévédérom. Mais, même compressés, les fichiers vidéo sont, par nature, lourds : ils accaparent beaucoup d'espace disque sur les serveurs et sont particulièrement gourmands en bande passante. De plus, la compression a ses limites. Toute compression dégrade la qualité du signal original. Cette dégradation est, jusqu'à un certain niveau de compression, subjectivement imperceptible à l'œil. Cependant, lorsque la compression devient trop importante, la qualité objective de l'image et du son se dégrade rapidement jusqu'à devenir inacceptable pour l'utilisateur. C'est pourquoi, dans des applications où le volume vidéo est important, les fichiers sont enregistrés sur un support physique et distribués. Un niveau de qualité minimale est alors garanti. Un précieux espace sur les serveurs et les disques rigides des utilisateurs est également sauvegardé.

La vidéo peut également être **diffusée** sur les réseaux informatiques. Malgré les limites qu'impose ce choix, de diffuser les contenus sur réseau, par exemple dans le cas d'informations devant être mises à jour fréquemment ou encore d'informations devant être indexées et recherchées. Il existe deux modes de diffusion de la vidéo sur les réseaux : le téléchargement, la diffusion en continu, mieux connue sous l'appellation anglophone de «streaming». La diffusion en continu peut quant à elle être faite soit en direct, l'événement capté étant numérisé en temps réel et diffusé sur les réseaux, soit en différé, le fichier numérisé en compressé étant alors déposé sur un serveur. Il existe trois principales architectures de diffusion vidéo sur Internet.

## 7. Architectures, codecs et formats de diffusion

Une architecture est une solution matérielle et logicielle permettant de produire, d'enregistrer et de diffuser des fichiers informatiques de médias dynamiques. Une architecture comporte deux principaux éléments: un serveur et un client. Les trois architectures les plus courantes pour l'audio et la vidéo sont Quicktime, RealSystem et Windows Media. Les trois architectures possèdent leur logiciel client spécifique. Ainsi, le client Quicktime devra être installé sur le poste de l'utilisateur pour que celui-ci puisse lire des fichiers QuickTime (.mov). Certains logiciels clients peuvent toutefois lire plus d'un format. Ainsi le client Windows Media Player peut lire les formats .aif, .aifc, .aiff, .asf, .asx, au, .avi, .cda, .ivf, .m1v, .m3u, .mid, .midi, mp2, .mp3, .mpa, .mpe, .mpeg, .mpg, .mpv2, .rmi, .snd, .wax, .wav, .wma, .wms, wmv, .wvx, .wmz.

Un codec peut être un logiciel, une pièce d'équipement ou combinaison des deux. Un codec (compresseur-décompresseur) est utilisé pour réduire la taille et le débit de fichiers qui seraient autrement trop encombrants ou trop lourds. Un codec est donc utilisé pour compresser des fichiers qui pourront être diffusés dans différents formats. Pour pouvoir lire ces fichiers, l'utilisateur devra avoir le même codec installé sur le logiciel client de son poste de visualisation. Cinepak, Sorenson ou Indeo sont des codecs. Les codecs peuvent compresser les fichiers et les sauvegarder sous différents formats de diffusion. Par exemple, un fichier compressé avec le codec Sorenson pourra être sauvegardé en format QuickTime (.mov) ou en Audio Video Interleaved (.avi). Certains Codecs sont partie intégrante d'une architecture donnée. Ainsi les architectures Windows Media et RealSystem possèdent leurs codecs spécifiques qui compressent les fichiers dans leurs formats propriétaires, respectivement .asf (Advanced Streaming Format) pour Windows Media et .ra ou .rm (Real Audio ou Real Media) pour RealSystem.

Un format est, quant à lui, une organisation des informations d'un fichier informatique. Dans un environnement Windows, un format se reconnaît à son extension. Ainsi, une extension .wav

identifie un fichier Audio Wave et une extension .mp3 identifie un fichier audio MPEG Layer III audio.

## 7.1 Architectures

### Quicktime

Une des toutes premières architectures de vidéo numérique. D'abord développée pour distribution sur cédérom, l'architecture a peu à peu évolué et permis une diffusion sur Internet d'abord en mode téléchargement, puis en transmission progressive pour enfin permettre la transmission en continu (*streaming*).

Le serveur de transmission en continu Quicktime est gratuit. Il peut être installé sur différentes plates formes (Mac OS, Solaris, Linux, Windows NT).

L'encodage Quicktime à un niveau de qualité professionnel nécessite des outils spécialisés, par exemple l'encodeur Sorenson Video ou Sorenson Broadcaster ou encore des encodeurs audio complémentaires. Quicktime supporte la norme SMIL (Synchronized Multimedia Integration Language) qui permet de synchroniser la séquence vidéo avec des pages HTML.

Un des grands avantages du format Quicktime est sa parfaite compatibilité avec les systèmes Mac OS et Windows. Cette caractéristique est particulièrement utile dans le cas de cédéroms bi-systèmes.

### Real System

RealSystem développé par RealNetworks est une architecture de transmission en continu. Elle nécessite un serveur dédié. Les coûts de licence pour l'installation sur le serveur sont calculés en fonction du nombre de canaux simultanés supportés. Real supporte le SMIL (Synchronized Multimedia Integration Language) qui permet de synchroniser la vidéo avec des pages HTML. Real permet la transmission en continu normale ou intelligente. Real est avant tout une solution de transmission en continu et c'est pourquoi on ne retrouve pratiquement jamais le format Real sur cédérom.

Le logiciel Real possède son propre codec vidéo, le Real G2. L'encodeur doit aussi être acheté directement de RealNetworks. Le client de base est gratuit mais la version disposant des fonctionnalités complètes doit être achetée.

### Windows Media

Les fichiers Windows Media s'installent sur un serveur Windows NT. L'encodeur et le client sont gratuits.

Pour encoder, Windows Media utilise une version modifiée par Microsoft de la norme MPEG 4 appelée MPEG-4 v3. Cette norme est légèrement différente de la norme MPEG 4 originale qui est quant à elle basée sur le format Quicktime. L'encodage produit des fichiers au format .asf (Advanced Streaming Format).

L'architecture Microsoft fonctionne évidemment en parfaite harmonie avec le système Windows et Internet Explorer. Dans un environnement MacOS et Netscape certaines limites s'imposent.

## 7.2 Quelques codecs

### CODECS VIDÉO :

Cinépak  
Indeo  
Real G2  
Sorenson  
MPEG-4 v3

### CODECS AUDIO :

Liquid Audio  
MP3 (MPEG Layer III Audio)  
Qdesign Music Codec  
Real Audio  
Windows Media Audio

## 7.3 Les formats

### FORMATS AUDIO

.aif / Audio Interchange Format  
.mp3 / MPEG Layer 3 Audio  
.ra / Real Audio  
.wav / Wave Audio

### FORMATS VIDÉO

.asf / Advanced Streaming format  
.avi / Audio Video Interleaved  
.mov / Quicktime Vidéo  
.mpg / MPEG 1 Video  
.rm / RealVideo

## 8. La vidéo sur réseau

Malgré la compression, la diffusion vidéo sur Internet est toujours confrontée au problème de la limitation de la bande passante et du débit. Un fichier vidéo, même compressé, peut avoir une taille de plusieurs mega octets, ce qui implique de longues minutes ou heures de téléchargement pour un utilisateur relié au réseau à l'aide d'un modem téléphonique de faible capacité. C'est pour pallier ce problème que la solution de la transmission en continu (*streaming*) a été développée ;. Plutôt que d'obliger l'utilisateur à attendre téléchargement complet du fichier avant d'en débiter le visionnage dès que la connexion au serveur est établie. Il existe deux modes de transmission en continu.

### Transmission en continu à débit constant

méthode de transmission du contenu par laquelle le contenu multimédia situé sur un serveur, puis lu grâce à l'envoi des données sur un réseau en flux continu. La transmission de données en continu permet d'écouter ou de visionner une séquence audio et vidéo immédiatement plutôt que d'attendre le téléchargement complet du fichier.

### Transmission en continu intelligente

Type de transmission en continu qui détecte les conditions du réseau et ajuste les propriétés d'un flux vidéo ou audio afin d'obtenir une qualité maximale.

## 9. La diffusion Web : une esthétique différente

La télévision et l'ordinateur sont des médias fort différents. Les réseaux télévision et l'Internet sont aussi différents. Un utilisateur d'Internet et un téléspectateur auront, au moment où ils utilisent l'un ou l'autre des médias, une attitude et un comportement fort différents. En effet, si le téléspectateur s'assoit généralement pour écouter des blocs de 30 minutes de façon continue, la session de navigation de l'Internaute est, selon les résultats d'une enquête Nielsen / NetRatings de mars 2001, d'une durée moyenne 31 minutes et 16 secondes alors que le temps moyen d'arrêt sur une page est de 54 secondes.

La télévision est un média dynamique à défilement linéaire. Généralement, une « bonne » émission de télévision est une émission qui divertira ou informera, qui racontera une histoire, qui aura une structure narrative. L'Internet, quant à lui, est avant tout interactif et réticulaire. L'utilisateur navigue sans nécessairement avoir de plan préalable précis, il construit lui-même ses enchaînements, son cycle d'attention est beaucoup plus court.

En ce sens, la diffusion vidéo sur Internet s'effectue dans un contexte sensiblement différent de celui de la diffusion sur les réseaux télévisuels traditionnels. Ainsi, il est à peu près inutile, d'un point de vue d'efficacité communicationnelle, d'encoder une émission au format Real (.ra) ou MediaPlayer (.asf), de la déposer sur serveur et de croire que les utilisateurs s'y précipiteront et resteront les yeux rivés sur leur écran cathodique pendant une longue période de temps. Une des «lois» de l'Internet est que le doigt sur la souris est plus rapide que le doigt sur la télécommande... Le Web est interactif, l'utilisateur est naturellement porté à cliquer, à bouger, à copier, à fusionner, à réorganiser plutôt que de rester au même endroit et d'écouter ou regarder l'histoire qu'on lui propose. L'utilisateur vient au Web pour faire quelque chose ou chercher quelque chose plutôt que pour se laisser bercer par quelque chose.

Ainsi, la grammaire de la Web vidéo est différente. Les séquences destinées à être diffusées sur Internet doivent être courtes, les contenus segmentés. D'un point de vue davantage technique la compression vidéo doit être bien comprise et appliquée. Les algorithmes de compression sont conçus pour réduire les redondances spatiales et temporelles. Ainsi, au moment de la capture, la complexité spatiale et temporelle doivent être réduites dans toute la mesure du possible. Les mouvements de caméra, les changements de plan fréquents, les transitions dynamiques sont exigeantes pour les algorithmes. La taille de l'affichage et les contraintes de résolution font aussi qu'il est préférable de construire les séquences en utilisant davantage de gros plans et peu ou pas de plans grand ensemble. Toujours dans le but de minimiser la taille des fichiers et le débit nécessaire, le nombre de couleurs dans les séquences doit être réduit au maximum.

D'autres fonctionnalités peuvent également être ajoutées à la vidéo utilisée dans les environnements interactifs, par exemple des possibilités de recherche par contenu de séquence. Le développement des bases de données vidéo est un des exemples d'une telle application. La norme MPEG 7 supportera cette possibilité d'indexation.



# Glossaire

## **AIFF**

Initiales de *Audio Interchange File Format*, un format audio développé par Apple Computers.  
Extension de nom de fichier : .aif

## **Analogique :**

Mode de représentation de l'information par une valeur continue. On parle de communication analogique lorsque l'information à transmettre est une grandeur physique concrète, comme le signal électrique généré par un microphone ou une caméra.

## **Algorithme**

Tout ou partie d'un programme qui effectue un calcul, ou une suite d'opérations. Par exemple, un algorithme de compression réduit la taille d'un fichier ; un algorithme d'affichage transfère les informations depuis une mémoire vers un écran.

## **AVI :**

Fichier vidéo de Microsoft *Windows* (audio et vidéo entrelacés, enregistré sans compression ou avec une compression minimale). Extension de nom de fichier : .avi

## **ASCII**

Initiales de *American Standard Code for Information Interchange*. Codage informatique d'origine américaine qui définit la correspondance entre les caractères de l'alphabet et un ensemble de valeurs binaires. Par exemple la séquence binaire 1010000 correspond à la lettre «P» majuscule alors que la séquence 1110000 correspond à la lettre «p» minuscule.

## **Assemblage :**

Montage élémentaire qui colle bout à bout diverses séquences vidéo et qui précède le montage proprement dit.

## **Balayage**

Parcours effectué par le faisceau d'électrons que le tube cathodique envoie sur les lignes paires et impaires de diodes lumineuses recouvrant un écran de télévision.

## **Bande passante**

Capacité de transfert de données d'un système de communication numérique, tel que le réseau Internet ou un réseau local. La bande passante est généralement exprimée en nombre de bits qu'un système est capable de transférer en une seconde. Les notions de bande passante élevée ou de large bande font référence à un réseau doté d'un taux de transfert de données élevé.

Elle exprime également la capacité d'un canal à transmettre un flot d'informations.

La largeur de bande passante s'exprime généralement en kilohertz (kHz) ou en mégahertz (MHz). La bande passante en télévision est de 5,5 MHz. Celle du signal haute-fidélité en audio est de 20 kHz, ce qui correspond à la gamme des fréquences audibles.

## **Bit**

Unité élémentaire d'information dont l'abréviation internationale est bit (**BI**nary **digiT**). Il peut prendre deux valeurs codées 0 ou 1.

## **Câble**

Matériel de transmission d'informations qui relie les composants d'un système. Un câble est composé d'un ou plusieurs fils entourés d'une ou de plusieurs gaines isolantes. Il se termine à chaque extrémité par un connecteur. Un câble se caractérise par ses composants (nombre de fils qu'il contient, matériaux utilisés, type de gainage...), son débit et sa longueur. Les câbles métalliques les plus utilisés sont les câbles coaxiaux et les câbles à paires torsadées. Les câbles à paires torsadées sont constitués d'une ou plusieurs paires torsadées. Ces câbles transmettent à la fois en analogique et en numérique. Les câbles coaxiaux permettent de transmettre avec de hauts débits sur de longues distances : sur une longueur de 1 kilomètre, des débits supérieurs à 100 Mbits/s peuvent être atteints. Le câble coaxial est utilisé pour la communication large bande à un coût encore significativement inférieur à celui de la fibre optique... Les câbles à fibres optiques utilisent des fils de verre très fins (1/10 mm) et transmettent, avec un minimum de pertes, les informations codées sous forme de flux lumineux. Ils se caractérisent par un important débit - plusieurs centaines de Mbits/s jusqu'à plusieurs Gbits/s. Ils sont utilisés de préférence pour la transmission d'images et de sons.

## **Canal**

1. ensemble de fréquences attribuées à une chaîne de télévision.
2. support de transmission d'informations, par exemple, un câble, une liaison hertzienne, une bande magnétique ou une simple feuille de papier.

## **Capture :**

Transformation d'un signal analogique sur ruban en un fichier numérique sur disque rigide. Cette opération est réalisée par la carte de capture (ou d'acquisition) vidéo.

## **Carte d'acquisition vidéo**

Carte permettant de transformer la source analogique en numérique et de compresser les données pour que l'ordinateur soit capable de les traiter. La qualité de capture est tributaire de la puissance globale de l'ordinateur et de ses périphériques. Les cartes d'acquisition vidéo intègrent aussi généralement un circuit audio pour une acquisition audio totalement synchronisée.

## **Codec**

Abréviation de COMPressor/DECOMPRESSOR. Les codecs sont des algorithmes informatiques que l'on applique aux fichiers images, audio ou vidéo afin de réduire leur taille. Ainsi, les fichiers requièrent moins d'espace de stockage. Ils requièrent également une bande passante moins importante pour la transmission en continu sur un réseau. Pour décompresser un fichier, un utilisateur doit posséder le même codec que celui utilisé pour le compresser.

## **Composante (YUV ou Y, R-Y, B-Y ou Y,Cr,Cb)**

Technique de composantes couleurs utilisant les voies de chrominance du rouge ( $R-Y = 0,7 R - 0,59 V - 0,11 B$ ) et bleu ( $B-Y = 0,89 B - 0,3 R - 0,59 V$ ), la luminance étant restituée par ( $Y = 0,3 R + 0,59 V + 0,11 B$ ).

## **Composite**

Signal unique contenant les informations de luminance (noir et blanc), chrominance (couleur) et informations de synchronisation.

## **Compression**

Réduction de la taille d'un ensemble d'informations.

## **Cryptage**

Opération qui consiste à rendre illisible une information en la modifiant suivant un algorithme. Le cryptage est utilisé par les chaînes de télévision à péage et, en informatique ou en transmission, pour la sécurité et l'inviolabilité des données. En vidéo et en télécommunication, on préférera le terme de brouillage.

## Débit numérique

Nombre d'éléments unitaires d'information (bits) transmis par seconde par un système numérique. Il s'exprime généralement en Mégabits par seconde (Mbits/s)

## Débit numérique vidéo

Débit de données en acquisition vidéo. Exprimé en nombre de bits par seconde, il dépend du nombre d'images par seconde, de la taille de l'image et de la profondeur de couleur (nombre de couleurs différentes possible que peut prendre un pixel). Le débit numérique choisi pour la capture vidéo doit tenir compte des performances du disque rigide.

## DIB (device independent bitmap)

Fichier image universel non compressé qui enregistre chaque point de l'image (pixel) et qui peut être lu par n'importe quel logiciel de dessin. Ce format de fichiers, gourmand en mémoire, s'utilise pour les importations / exportations entre logiciels. En principe, il est destiné à être compressé, sauf s'il doit servir de référence pour d'autres travaux : le plus souvent un fichier **dib** ou **bmp** servira de "master" pour la compression gif ou jpeg.

## Drop Frame

Le format 525/60 (lignes /trames) utilisé avec le codage couleur NTSC n'a pas une cadence exacte de 60 trames par seconde mais bien de 59.94, soit à 29.97 images par seconde - une différence de une sur mille. Le time code SMPTE donne attribue un numéro spécifique à chacune des images, au rythme de 30 numéros à la seconde. Le "Drop-frame time code" établit une compensation en sautant deux numéros d'image par minute, sauf la dixième. Il est à noter que le système PAL, lui, est exact et ne nécessite pas de "drop-frame".

## DV

Système d'enregistrement numérique grand-public présenté pour la première fois en août 1995 à la Funkausstellung de Berlin. Principal avantage: la multi génération possible sans dégradation de la qualité audio et vidéo. De plus la qualité d'image et de son se révèle supérieure à ce que l'on peut attendre des équipements analogiques grand-public, et ce, malgré la compression de 1:5. La luminance est échantillonnée à 13,5 MHz et la chrominance à 6,75 MHz (4:2:0). Le son peut être enregistré au choix en stéréo sur 16bits/48KHz ou, toujours en stéréo sur 12bits/32KHz. On conservera alors deux voies 12bits/32KHz pour un doublage son ultérieur sur un magnétoscope numérique. Autres caractéristiques techniques: les bandes possèdent des pistes de 10um et une vitesse de défilement de 18,8 mm/s. Tous les appareils DV sont compatibles, quelle que soit la marque. Une cassette enregistrée sur un modèle Panasonic sera parfaitement lisible sur un Sony, etc. La durée d'enregistrement des petites cassettes (mini DV), exploitées par la quasi totalité des caméscopes grand-public est limitée à 60 minutes. Des cassettes de grande taille (de 120 à 270 minutes) ont été développées pour les magnétoscopes. Au nombre des atouts du DV: l'enregistrement systématique d'une time code, les sub codes, le gel d'image, le mode photo, la possibilité lors d'un montage ultérieur d'insertion d'image seule.

## DVCAM

Système d'enregistrement numérique professionnel. Si tous les produits DV grand-public présentent les mêmes caractéristiques, il n'en est pas de même pour dès que l'on touche à l'univers pro. Ainsi, le DVCam signé Sony se révèle incompatible avec le DVCPPro de Panasonic. A noter: un magnétoscope DVCam peut lire une cassette DV. Certains rares équipements DV, comme le magnétoscope Sony DHR-1000 peuvent cependant lire des cassettes DVCam et les enregistrer en DV. Mais en règle générale un caméscope DV ne saurait lire du DVcam. Caractéristiques de la bande: des pistes de 15um et une vitesse de défilement de 28,2 mm/s. Structure d'échantillonnage chrominance: 4:2:0. Durée maximal des grandes cassettes: 184 minutes. Pour le reste on retrouve les mêmes spécifications que pour le DV.

## **DVCPRO**

Système d'enregistrement numérique professionnel. C'est la version Panasonic du DV Pro. Sa spécificité: il utilise des bandes Métal Particules, présente de pistes de 18µm et une vitesse de défilement de 33.8 mm/s. Sa structure d'échantillonnage de la chrominance est de 4:1:1. Durée maximal des cassettes de taille moyenne: 63 minutes et des grandes cassettes: 123 minutes. N'accepte pas de petites cassettes. Côté compatibilité: un magnétoscope DVCPRO peut lire des enregistrements DV ou DVCam quelle que soit la taille de la cassette. L'inverse n'est pas possible et seul un lecteur DVCPRO peut décrypter une cassette DVCPRO. Pour le reste, on retrouve les spécifications du DV.

## **DVCR**

Magnétoscope numérique se présentant sous forme de disque dur amovible.

## **DVD**

Initiales de Digital Video Disc ou Digital Versatile Disc. Norme de vidéodisque comportant des sons, des images numériques et des données. Les principaux acteurs de l'électronique grand-public, l'information et l'industrie du divertissement se sont entendus sur les normes du DVD. Les fondateurs, qui se sont déclarés le 24 janvier 1994, ne sont autres que Matsushita, Toshiba, Thomson, Hitachi, MGA, MGM/UA, Time Warner, Pioneer... Se sont notamment ralliés à cette alliance, les sociétés Sony et Philips qui dans un premier temps prévoyaient de faire bande à part en commercialisant leur propre format de DVD. Le vidéodisque se présente sous la forme d'une galette de 12cm de diamètre, d'une capacité de 4,7 Giga-Octets, soit une contenance 7 fois supérieure à celle d'un CDROM sur une seule face. Il peut accueillir jusqu'à 2h15 minutes de vidéo, qualité 500 points/ligne (MPEG2) et un son multicanal. Il devrait également offrir des possibilités de multilinguisme (jusqu'à 8 langues), de sous-titrage (32), un choix de rapport d'écran (4/3, Letterbox, Panscan et 16/9). Devraient également être commercialisés dans le futur des disques de 8,5 Go (4h) double couche, toujours sur simple face, ou encore des 9,4 (4H30) et 17Go (9H) double face/double couche. Interviendront aussi à l'horizon 2000, des disques enregistrables une fois (3,8Go), voire réinscriptibles (2,6Go). Dès 1997 tous les majors de l'électronique devraient sortir leurs lecteurs DVD. Les modèles des différents constructeurs seront compatibles.

## **Echantillonnage**

Pour être numérisé, le signal vidéo ou audio modulé doit d'abord être découpé selon des intervalles de temps précis, selon une certaine fréquence d'échantillonnage d'ailleurs très élevée (13,5MHz en numérique DV, par exemple). En effet, les échantillons unitaires doivent être assez brefs pour reproduire fidèlement les variations les plus fines et rapides de la modulation.

## **Fibre optique**

Fibre constituée de silice ou de matière plastique utilisée pour le transport de rayons optiques. Un système à fibre optique est un système de transmission d'informations par câbles à fibre optique utilisé en téléphonie, télématique et télévision.

## **Format d'écran**

1. Taille de l'écran exprimée généralement par la dimension de sa diagonale (par exemple : 14 pouces, 17 pouces, 55 centimètres).

2. Rapport entre la largeur et la hauteur d'un écran (par exemple : 4/3, 16/9).

Les écrans les plus courants en télévision ou en informatique ont un format 4/3, c'est-à-dire qu'un écran de 15 pouces (diagonale) a une largeur de 12 pouces et une hauteur de 9 pouces. La télévision à haute définition utilisera le format 16/9, plus proche du format du champ visuel et du format utilisé au cinéma.

## **GIF (graphics interchange format, © CompuServe) :**

Compression d'images non photographiques (dessins comportant des aplats de couleur). Le gif produit des images en 256 couleurs, sans dégradation. Une variante de ce format permet d'enregistrer des animations image par image (animated gif). Extension du nom de fichier : gif.

## **Interactivité**

1. Type de relation entre deux systèmes qui fait que le comportement d'un système modifie le comportement de l'autre.
2. Propriété d'un message qui donne la possibilité à l'opérateur qui le reçoit d'intervenir sur le déroulement du programme qui l'a émis. S'oppose à linéarité.
3. Ensemble des techniques et des procédures mises en œuvre pour conférer cette propriété à un message.

Les différents degrés d'interactivité dépendent du support utilisé et de l'émetteur du message : une conversation peut être très interactive ou au contraire tomber dans le monologue, un livre peut être un support interactif (ex : Bottin). Les supports électroniques permettent différents niveaux d'interactivité. Comme dans un livre, mais de manière plus souple encore, le spectateur peut naviguer à son choix, sauter d'une séquence à l'autre, revenir en arrière, demander des compléments d'informations sur un point particulier... Ce qui est nouveau par rapport au livre, grâce à l'informatique, c'est que le programme peut suggérer ou imposer au spectateur différents déroulements du récit, en fonction de données précises (identité du spectateur, chemin déjà parcouru, réponses fournies par le spectateur à des questions explicites ou implicites) ou en fonction d'un choix aléatoire effectué par la machine...

## **Internet**

Réseau mondial reliant plus de 20 000 réseaux de toutes tailles et environ vingt millions d'utilisateurs répartis dans tous les pays. Internet est le précurseur des autoroutes de l'information. Il utilise la commutation de paquets. Le transfert des informations se fait à un débit variant de 240 caractères par secondes (à partir d'un PC relié à une ligne téléphonique) à 150 000 caractères par seconde. A ce débit, on peut échanger de la parole, des documents graphiques, des photos couleurs.

## **JPEG :**

Norme de compression des images photographiques (en 16 ou 24 bits) basée sur les travaux d'un groupe de chercheurs "Joint Photographic Experts Group". La compression s'effectue par perte de qualité Extension : .jpg

## **MJPEG (Motion Jpeg) :**

Norme de compression de la vidéo dérivée du Jpeg. Chaque image est compressée individuellement, ce qui se prête bien au montage image par image et aux effets spéciaux. Le taux de compression s'exprime par une fraction. Ex. 1:3 : l'image a été compressée d'un facteur 3. Entre 1:4 et 1:7 on parlera de vidéo de qualité S-VHS ; au-delà de qualité VHS... 1:1 exprime une compression nulle.

## **Modulation**

1. Transmission analogique : la modulation d'un signal (appelé onde porteuse) consiste à transformer une de ses caractéristiques (amplitude, fréquence ou phase) par l'effet d'un autre signal contenant l'information utile pour faciliter le transfert de ce dernier (transmission, enregistrement ou reproduction).
2. Transmission numérique : la modulation désigne la technique de modulation par impulsion et codage et ses dérivés.

## **Multimédia**

Technique de communication qui tend à rassembler sur un seul support l'ensemble des moyens audiovisuels (graphisme, photographies, dessins animés, vidéos, sons, textes) et informatiques (données et programmes) pour les diffuser simultanément et de manière interactive.

## **Multiplexage**

Assemblage de plusieurs signaux en un seul, appelé signal composite ou multiplex, pour faciliter la transmission. Le multiplexage est un procédé réversible : il est fait de telle façon que les signaux composants puissent être à nouveau séparés lors de la réception du signal composite.

## **Montage**

Action qui consiste à créer une séquence vidéo définitive à partir de séquences brutes ("rushes") ou d'autres vidéos. Le montage consiste à couper, coller, insérer des plans ou séquences...

## **Montage linéaire**

En reliant 2 magnétoscopes (copie A-B) on peut, à partir de bandes contenant les rushes, créer une "première" bande appelée "master" produite par l'assemblage de séquences mises bout à bout. Il est à noter que la production du "master" est déjà soumise à la dégradation de la vidéo lors de la copie.

## **Montage virtuel ou non linéaire**

L'ordinateur gère des directives de montage à partir desquelles on peut produire autant de "masters" qu'on le souhaite. Il est donc très facile de revenir sur la durée d'une séquence, d'interpoler des écrans, de substituer une séquence à une autre même si leur durée est différente etc. sans avoir à tout refaire car l'accès aux images n'est pas séquentiel mais aléatoire. Les fichiers obtenus sur disque dur permettent de mener plusieurs projets de montage, ce qui est impensable en montage "cut".

## **Morphing :**

Déformation progressive d'une image 1 vers une image 2. Le "morphing" produit une séquence animée. Ne pas confondre avec le "warping", simple déformation d'une seule image.

## **MOV (movie) :**

Nom des fichiers de vidéo numérique de *QuickTime*.

## **MPEG (Moving Pictures Experts Group, fondé en 1987) :**

Groupe d'experts chargé d'élaborer des normes de compression d'images animées et sonorisées. La norme MPEG, adoptée en 1990, vise à réduire la taille d'un fichier d'images de 100 à 200 fois. Elle a été retenue par Philips pour la numérisation de l'image animée plein écran du CD-I. La norme MPEG2, destinée à la télévision numérique, adoptée en 1994, réduit la taille des fichiers image dans un rapport 1/15 à 1/40, ce qui permet de disposer d'une qualité améliorée par rapport à MPEG1. La particularité de MPEG2 est de couvrir l'ensemble des applications multimédia (télécommunications, audio, images fixes et animées, applications interactives), et les supports associés (bandes magnétiques, disques compacts, ordinateurs, télévisions). En images animées, les normes MPEG se basent sur une continuité du mouvement pour ne retenir que quelques images dans le temps, qui seront effectivement captées, les images intermédiaires étant calculées. Extension du nom de fichier : mpg.

## **MPEG 1**

Mis au point en 1992, avec comme référence technique d'alors le lecteur de cédérom simple vitesse et son débit de 150 Ko par seconde. Cette contrainte a nécessité la réduction de l'image vidéo avant sa compression. Le Mpeg1 ne traite que des images de 352 x 240/288 pixels (format SIF, Standard Image Format) non entrelacées. Le taux de compression maximale est de 26.

## **MPEG 2**

Proposé en 1994, reprend la totalité des algorithmes de Mpeg1 qu'il optimise avec quelques compléments. Il peut traiter des images de 752 x 480 pixels. Cette compression a été pensée avec comme référence la télévision numérique et le lecteur de DVD (débit de 1200 Ko par seconde au minimum).

**N. B.:** Il n'existe pas de Mpeg3. Les fichiers mp3 (fichiers son compressés de qualité CD Audio) appartiennent en fait à la norme Mpeg1 pour l'audio (Mpeg Audio Layer3).

### **MP3 (Mpeg Audio Layer3) :**

Norme de compression du son de qualité CD Audio. Un fichier compressé à cette norme occupe environ 1/12è de l'espace de son équivalent en wav. 3 minutes de musique occupent 30 Mo en wav et seulement 2 Mo en mp3. Ce format a d'ores et déjà un énorme succès car il permet de diffuser de la musique sur Internet. Extension du nom de fichier : mp3.

### **NTSC (National Television Standards Committee) :**

Procédé de télévision couleur utilisé aux U.S.A. et au Japon. 525 lignes, 60 trames par seconde, 30 images par seconde. Voir aussi PAL, SECAM.

### **Numérique (vidéo numérique) :**

La vidéo est codée par une suite de 0 et de 1 qui décrivent précisément chaque image. Le numérique est inusable (on peut lire un fichier numérique sans altérer son support) et n'est pas soumis à la perte en cas de copies successives, contrairement aux copies de bandes à bandes. L'énorme flux de données nécessaire rend indispensable la compression.

### **PAL (Phase Alternative Lines) :**

Procédé de télévision couleur utilisé en Europe. 625 lignes. 50 trames par seconde, 25 images par seconde.

### **QuickTime (© Apple) :**

Vidéo numérique développée par Apple Computers pour le Macintosh puis portée sous *Windows*. Extension du nom de fichier : mov.

## **Réseau**

Ensemble des moyens de transmission et éventuellement de commutation permettant l'échange ou la diffusion d'informations. Un réseau Multimédia transporte des images, des données et des sons et permet à ses utilisateurs de réagir de manière interactive aux programmes qu'ils reçoivent. Une grande similitude existe entre tous les réseaux, quelle que soit la nature de ce qu'ils véhiculent.

### **Réseau hertzien**

Réseau utilisant les ondes hertziennes comme support de transmission. Il est aussi appelé réseau de faisceau hertzien. Il s'oppose au réseau câblé et au réseau satellitaire. Un réseau hertzien est constitué de stations relais installées au sol pourvues d'antennes et réparties sur l'ensemble d'une zone. Aux États-Unis, on ne compte que 4000 émetteurs/réémetteurs contre 17000 pour la France, pourtant 20 fois moins étendue.

### **Réseau Numérique à Intégration de Services (RNIS)**

Réseau plus connu sous son sigle RNIS disponible sur l'ensemble du territoire français depuis 1991 et capable de transporter des images, des sons et des données informatiques. Il est principalement utilisé dans le monde professionnel mais a été conçu pour équiper le grand public. La plupart des pays industrialisés mettent en place de tels réseaux, créant ainsi un réseau à l'échelle mondiale à caractère multimédia.

### **Real audio / real video (© RealNetworks )**

Procédé de compression du son et de la vidéo permettant le "streaming" c'est à dire la consultation du contenu d'un fichier au fur et à mesure de son téléchargement sur Internet. Extension du nom de fichier : rm.

### **Rush :**

Séquence vidéo brute qui vient d'être tournée. On appelle "dérushage" le premier travail de montage qui consiste à couper les éléments indésirables de ces séquences brutes.

## **SECAM (SEquenciel Couleurs Avec Mémoire) :**

Procédé de télévision couleur développé en France. 625 lignes. 50 trames par seconde, 25 images par seconde.

## **S-Vidéo :**

Evolution de la vidéo composite qui sépare les signaux de luminance (Y) et de chrominance (C)

Pour le matériel (magnétoscopes, caméscopes) on parle de S-VHS.

## **Télévision**

Littéralement " vision à distance ", la télévision recouvre l'ensemble des actions et des techniques mises en œuvre pour produire, transmettre et reproduire des images animées sonorisées.

## **Télévision par câble**

Télévision utilisant des réseaux câbles pour la diffusion de ses programmes, par opposition à la télévision traditionnelle utilisant les faisceaux hertziens. La télévision par câble permet à l'utilisateur de recevoir un plus grand nombre de chaînes avec une meilleure qualité d'image. C'est elle qui permettra la télévision interactive.

## **Télévision à haute définition**

Nouvelles caractéristiques de diffusion de télévision permettant d'améliorer fortement la qualité de l'image (en se rapprochant de la qualité du cinéma), et de profiter d'un son stéréophonique.

## **Télévision hertzienne**

Télévision dont la diffusion est faite par faisceaux hertziens relayés à partir de stations émettrices et réémettrices.

## **Télévision interactive**

Type de télévision dans lequel le téléspectateur a une possibilité d'action sur le programme qu'il regarde : par exemple, la possibilité de choisir parmi plusieurs angles de vue lors d'une retransmission sportive ou de demander des détails sur telle ou telle information pendant une émission ou de choisir sur catalogue un film ou une émission (vidéo à la demande).

L'interactivité est encore très limitée par rapport à celle que peut offrir un système informatique. Les principaux champs d'application de la télévision interactive sont pour l'instant : les jeux, le téléachat, la vidéo à la demande.

Un système de télévision interactive est composé, en principe, des éléments suivants : un terminal décodeur désembrouilleur, un système de commande, un serveur multimédia de forte puissance, un système de navigation gérant l'accès aux chaînes, un réseau permettant une voie de retour des commandes des téléspectateurs vers le serveur d'images.

## **Télévision numérique**

Type de télévision où les signaux sont traités sous forme numérique d'un bout à l'autre de la chaîne de l'image, du capteur de la caméra jusqu'au récepteur du téléspectateur : la production, la transmission, la diffusion, la réception et le stockage sont numériques.

## **Transmission en continu**

Méthode de transmission du contenu, dans laquelle le contenu multimédia est situé sur un serveur, puis lu grâce à l'envoi des données sur un réseau en flux continu. La transmission de données en continu permet d'écouter ou de visionner une séquence audio et vidéo immédiatement plutôt que d'attendre le téléchargement complet du fichier.



### **Transmission en continu intelligente**

Type de transmission en continu qui détecte les conditions du réseau et ajuste les propriétés d'un flux vidéo ou audio afin d'obtenir une qualité maximale.

### **Vitesse de transmission**

Vitesse à laquelle un contenu vidéo et audio numérique est transmis en continu depuis une source, telle qu'un fichier, afin d'être correctement restitué par un lecteur, ou vitesse à laquelle un contenu binaire est transmis en continu sur un réseau. La vitesse de transmission se mesure généralement en kilobits par seconde (Kbits/s), par exemple 28,8 Kbits/s. La vitesse de transmission d'un fichier ou d'un flux de données en direct est déterminée au cours du processus de codage, lors de la création du flux. La bande passante correspond à la capacité de vitesse de transmission totale d'un réseau. Pour qu'un contenu audio et vidéo soit correctement restitué lors de sa transmission en continu sur un réseau, la bande passante du réseau doit être suffisamment élevée pour prendre en charge les vitesses de transmission de tous les contenus transmis simultanément.

### **WAV**

Format de fichiers son introduit par Microsoft dans *Windows*. Peu ou pas compressé, ce format produit des fichiers volumineux. Extension du nom de fichier : wav.

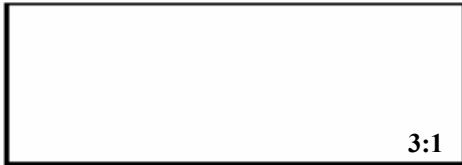
### **Y/C**

Norme vidéo qui code les signaux de luminance et de chrominance séparément . Est utilisée par le format S-VHS.



### Academy

Créé en 1910. Restera le «standard» jusqu'en 1927. D'abord ratio 1,37:1, sera réduit à 1,33:1 de 1927 à 1953 pour laisser de la place à la piste son optique.



### Cinérama

Le plus large des formats jamais employé. Développé par la Paramount et utilisé dès 1952, il nécessitait l'emploi de 3 caméras et 3 projecteurs. L'écran courbe accentuait l'effet panoramique.



### Camera 65

Développé par la MGM. L'image est anamorphosée par une lentille cylindrique. Il n'y eut qu'une dizaine de films tournés dans ce format.



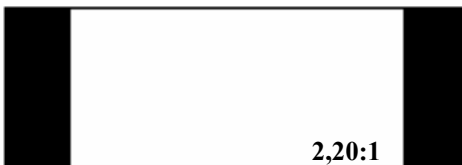
### Cinemascope

Utilisé pour la première fois par la Fox en 1953. Utilise aussi une lentille anamorphique compressée sur la largeur comme pour le format Panavision. Le format Cinemascope n'est plus utilisé depuis 1965.



### Panavision

Utilise aussi l'anamorphose pour la projection. C'est le standard le plus utilisé aujourd'hui.



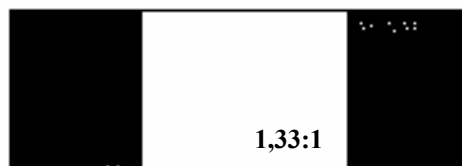
### Super Panavision 70

N'est plus utilisé. Quelques films célèbres tels, par exemple, *My Fair Lady* et *Exodus* ont été tournés dans ce format.



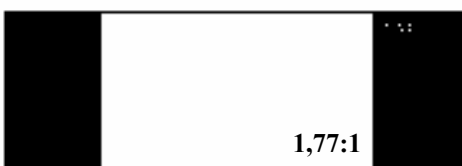
### Academy Flat

Le grand standard actuel pour les films tournés en 35mm.



### Télévision 4:3

Standard TV pour toutes les émissions hors cinéma. Il est aussi utilisé en photographie 35mm. Correspond exactement au format Academy de 1,33:1



### Télévision 16:9

Très proche du Academy Flat de 1,85:1. Standard d'affichage adopté pour la HDTV. Le choix du 16:9 tient aussi au fait qu'il s'agit d'un multiple de 4:3 ( $4:3 \times 4:3 = 16:9$ )