

## **DOSSIER COMPRESSION**

### **OBJECTIFS DU DOSSIER :**

- Dans le dossier suivant, nous vous proposons de découvrir les dessous de la vidéo numérique.

Ce dossier répondra notamment aux questions suivantes :

- Qu'est-ce que la vidéo numérique ?
- Qu'est-ce que la compression vidéo ?
- Quelles sont les caractéristiques des cartes vidéos ?

### **DEMARCHE DE TRAVAIL :**

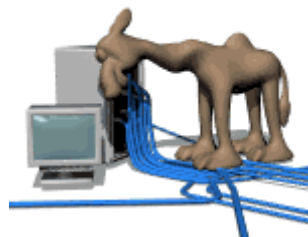
- Consultez rapidement le dossier avant de commencer
- Complétez le dossier en suivant les indications du document
- Réalisez les exercices correspondant au dossier multimédia

**BON COURAGE !**

Ce document comporte 3 dossiers :

- Un dossier sur la vidéo numérique
- Un dossier sur la compression numérique
- Un dossier sur les cartes vidéo

Lisez ces 3 dossiers et répondez aux questions qui vous seront posées...





Lisez le dossier et répondez aux questions suivantes :



Combien d'images faut-il afficher pour tromper l'œil humain et lui faire croire que l'image est animée ?



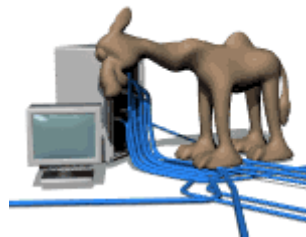
Pourquoi a-t-on décidé d'entrelacer les images dans le système PAL/SECAM ?



Combien y a-t-il d'images par seconde en PAL/SECAM ?



Combien y a-t-il d'images par seconde en NTSC et en vidéo numérique ?



## LA VIDEO NUMERIQUE

Ce document intitulé « Vidéo et imagerie numérique - Introduction à la vidéo numérique » issu de *Comment Ça Marche* est mis à disposition sous les termes de la licence *Creative Commons*. Vous pouvez copier, modifier des copies de cette page, dans les conditions fixées par la licence, tant que cette note apparaît clairement.

## INTRODUCTION A LA VIDEO NUMERIQUE

### Qu'est-ce qu'une vidéo

Une vidéo est une succession d'images à une certaine cadence. L'oeil 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'oeil et de lui faire croire à une image animée.

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

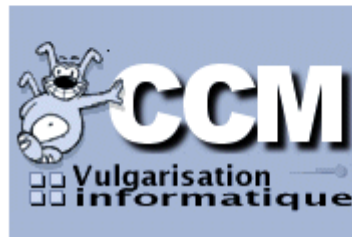
On distingue généralement plusieurs grandes familles d'"images animées"

- Le cinéma, consistant à stocker sur une pellicule la succession d'images en négatif. La restitution du film se fait alors grâce à une source lumineuse projetant les images successives sur un écran
- La vidéo **analogique**, représentant 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** consistant 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). a raison de 25 images par seconde a un format 4:3 (c'est-à-dire que le rapport largeur sur hauteur vaut 4/3).

Or a 25 images par seconde, de nombreuses personnes perçoivent un battement dans l'image. Ainsi, étant donné qu'il n'était pas possible d'envoyer plus d'information 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.



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 SECAM 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 :

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



Consignes...



Regardez le dossier multimédia « Compression » en n'oubliant pas de faire les exercices.



Lisez le dossier papier « Compression vidéo » ci-après



Répondez aux questions suivantes...



L'œil humain est-il plus sensible aux variations lumineuses ou aux variations de couleurs ?



Quelles sont les 6 lois de la compression numérique ?

1-

4-

2-

5-

3-

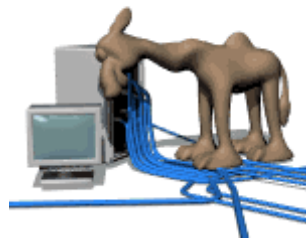
6-



Quelle est le principe de la compression temporelle ?



Quels sont les différents standards de compression (norme et codec) ?



## LA COMPRESSION VIDEO

## LA COMPRESSION VIDEO...

### Le M-JPEG

La première idée qui vient à l'esprit après s'être intéressé à la compression d'images est d'appliquer l'algorithme de compression JPEG à une séquence vidéo (qui n'est finalement qu'une suite d'images). C'est notamment le cas du M-JPEG (qui n'est pas vraiment ce que l'on appelle le MPEG) qui autorise un débit de 8 à 10 Mbps, ce qui le rend utilisable dans les studios de montage numérique, d'autant plus que chaque image étant codée séparément, on peut y accéder aléatoirement.

### Le MPEG

Dans de nombreuses séquences vidéos, de nombreuses scènes sont fixes, comme dans les feuilletons de l'Inspecteur Derrick (Oups! ça m'a échappé), cela se 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 ne décrire seulement le changement d'une image à l'autre. C'est là la différence majeure entre le MPEG et le M-JPEG.

Cependant cette méthode aura beaucoup moins d'impact sur une scène d'action (pas de craintes à avoir du côté de Derrick, vous pouvez encoder tranquillement votre épisode préféré de cette magnifique série! :)

Il existe donc 4 façons d'encoder une image avec le MPEG-1:

- 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 différence 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:

IBBPBBPBBPBB

Une image I est donc insérée toutes les 12 frames.

### Le MPEG-1

La norme MPEG-1 représente chaque image comme un ensemble de blocs 16 x 16. Il 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](#)).





Lisez le dossier « **LES CARTES VIDEO** » et répondez aux questions suivantes...



De quel élément principal dépendent les performances des cartes 2D ?

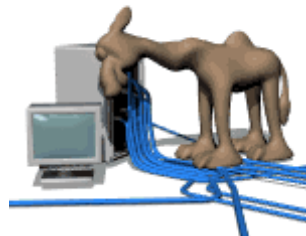


Quel élément intégrer dans les carte 3D permet de décharger le processeur du « Setup » et du « rendering » ?



Retrouvez les termes correspondants à ces différentes définitions...

	Affiche une représentation d'une scène selon 3 axes de référence (x,y, et z)
	Technique permettant de faire apparaître les pixels de façon moins crénelée.
	Image pixel par pixel



## LES CARTES VIDEOS

## LES CARTES VIDEO...

### Les cartes accélératrices 2D

Les cartes 2D n'ont pas changé de principe depuis leur création. Chaque puce possède de nombreux circuits qui permettent d'exécuter de nombreuses fonctions:

- déplacement des blocs (curseur de la souris par exemple)
- tracé de lignes
- tracé de polygones

Ainsi, les performances des cartes 2D n'évoluent plus depuis quelques temps.

Leurs performances sont tributaires du type de mémoire utilisée sur la carte (les mémoires SGRAM ou WRAM, mémoires vidéo spécifiques à 10 ns, donnent des résultats bien meilleurs que la mémoire EDO (60 ns))

La fréquence du RAM-DAC (RAM Digital Analogic Converter), ainsi que la quantité de mémoire vidéo ne permettent en rien d'avoir de meilleures performances, elles permettent juste d'avoir un meilleur taux de rafraichissement (nombre d'images par seconde) et de pouvoir accéder à des résolutions plus grandes.

### Les cartes accélératrices 3D

Le domaine de la 3D est beaucoup plus récent, donc plus porteur. On arrive à des puissances de calculs sur PC supérieures à celles de certaines stations de travail.

Le calcul d'une scène 3D est un processus qui se décompose grossièrement en quatre étapes:

- le script: mise en place des éléments
- la *géométrie*: création d'objets simples
- le *setup*: découpage en triangles 2D
- le *rendering*: C'est le rendu, c'est-à-dire le plaquage des textures

Ainsi, plus la carte accélératrice 3D calcule elle-même ces étapes, plus l'affichage est rapide. Les premières puces n'effectuaient que le rendering, laissant le processeur s'occuper du reste.

Depuis, les cartes possèdent un "setup engine" qui prend en charge les deux dernières étapes.

A titre d'exemple, un Pentium II à 266 Mhz qui calcule les trois premières étapes peut calculer 350 000 polygones par secondes, lorsqu'il n'en calcule que deux, il atteint 750 000 polygones par seconde.

Cela montre à quel point ces cartes déchargent le processeur.

Le type de bus est lui aussi déterminant. Alors que le bus AGP n'apporte aucune amélioration dans le domaine de la 2D, les cartes utilisant ce bus plutôt que le bus PCI sont beaucoup plus performantes. Cela s'explique par le fait que le bus AGP est directement relié à la mémoire vive, ce qui lui offre une bande passante beaucoup plus grande que le bus PCI.

Ces produits de haute technologie ont maintenant besoin de la même qualité de fabrication que les processeurs, ainsi que des gravures allant de 0.35 µm à 0.25 µm.

### Glossaire des fonctions accélératrices 3D et 2D

Terme	Définition
2D Graphics	Affiche une représentation d'une scène selon 2 axes de référence (x et y)
3D Graphics	Affiche une représentation d'une scène selon 3 axes de référence (x,y, et z)
Alpha blending	Le monde est composé d'objets opaques, translucides et transparents. L'alpha blending est une manière d'ajouter des informations de transparence à des objets translucides. Cela est fait en effectuant un rendu des polygones à travers des masques dont la densité est proportionnelle à la transparence des objets. La couleur du pixel résultant est une combinaison de la couleur du premier plan et de la couleur de l'arrière-plan. L'alpha a généralement une valeur comprise entre 0 et 1 calculée de la manière suivante: $\text{nouveau pixel} = (\text{alpha}) * (\text{couleur du premier pixel}) + (1 - \text{alpha}) * (\text{couleur du second pixel})$
Alpha buffer	C'est un canal supplémentaire pour stocker l'information de transparence (Rouge-Vert-Bleu-Transparence).
Anti-aliasing appelé aussi anti-crênelage)	Technique permettant de faire apparaître les pixels de façon moins crênelée.

Effets atmosphériques	Effets tels que le brouillard ou bien l'effet de distance, qui améliorent le rendu d'un environnement.
Bitmap	Image pixel par pixel
Bilinear filtering	Permet de fluidifier le passage d'un pixel d'un endroit à un autre (lors d'une rotation par exemple)
BitBLT	C'est l'une des fonctions d'accélération les plus importantes, elle permet de simplifier le déplacement d'un bloc de données, en prenant en compte les particularités de la mémoire-véo. Elle est par exemple utilisée lors du déplacement d'une fenêtre
Blending	Combinaison de deux images en les ajoutant bit-à-bit
Bus Mastering	Une fonction du bus PCI permettant de recevoir directement des informations de la mémoire sans transiter par le processeur
Correction de perspective	Une méthode pour faire du mappage (enveloppement) avec des textures (texture mapping). Elle prend en compte la valeur de Z pour mapper les polygones. Lorsqu'un objet s'éloigne de l'objectif, il apparaît plus petit en hauteur et en largeur, la correction de perspective consiste à dire que le taux de changement dans les pixels de la texture est proportionnel à la profondeur.
Depth Cueing	Baisse l'intensité des objets s'éloignant de l'objectif
Dithering	Permet d'archiver des images de qualité 24-bit dans des tampons plus petits (8 ou 16 bits). Le dithering utilise deux couleurs pour en créer une seule
Double buffering	Une méthode utilisant deux tampons, une pour l'affichage, l'autre pour le calcul du rendu, ainsi lorsque le rendu est fait les deux tampons sont échangés.
Flat shading ou Constant shading	Assigne une couleur uniforme sur un polygone. L'objet ainsi rendu apparaît de façon facettisée.
Fog	Utilise la fonction blending pour un objet avec une couleur fixe (plus il s'éloigne de l'objectif, plus cette fonction est utilisée)
Gamma	Les caractéristiques d'un affichage utilisant des phosphores sont non-linéaires: un petit changement de la tension à basse tension crée un changement dans l'affichage au niveau de la brillance, ce même changement à plus haute tension ne donnera pas la même magnitude de brillance. La différence entre ce qui est attendu et ce qui est mesuré est appelée Gamma
Gamma Correction	Avant d'être affichées, les données doivent être corrigées pour compenser le Gamma
Gouraud Shading (lissage Gouraud)	Algorithme (portant le nom du français qui l'a inventé) permettant un lissage des couleurs par interpolation. Il assigne une couleur à chaque pixel d'un polygone en se basant sur une interpolation de ses arêtes, il simule l'apparence de surfaces plastiques ou métalliques.
Interpolation	Façon mathématique de régénérer des informations manquantes ou endommagées. Lorsqu'on agrandit une image par exemple, les pixels manquants sont régénérés par interpolation.
Line Buffer	C'est un tampon fait pour mémoriser une ligne vidéo
Lissage Gouraud	Algorithme (portant le nom du français qui l'a inventé) permettant un lissage des couleurs par interpolation. Il assigne une couleur à chaque pixel d'un polygone en se basant sur une interpolation de ses arêtes, il simule l'apparence de surfaces plastiques ou métalliques.
Lissage Phong	Algorithme (portant le nom de Phong Bui-Tong) permettant un lissage des couleurs en calculant le taux de lumière en de nombreux points d'une surface, et en changeant la couleur des pixels en fonction de la valeur. Il est plus gourmand en ressources que le lissage Gouraud
MIP Mapping	C'est un mot provenant du latin "Multum in Parvum" qui signifie "Plusieurs en un". Cette méthode permet d'appliquer des textures de différentes résolutions pour des objets d'une même image, selon leur taille et leur distance. Cela permet entre autre de mettre des textures de plus hautes résolutions lorsqu'on se rapproche d'un objet.
Projection	C'est le fait de réduire un espace en 3-Dimension en un espace en 2-Dimension
Rasterisation	Transforme une image en pixels
Rendu (Rendering)	C'est le fait de créer des images réalistiques sur un écran en utilisant des modèles mathématiques pour le lissage, les couleurs ...
Rendering engine	Partie matérielle ou logicielle chargée de calculer les primitives 3D (Généralement des triangles)
Tesselation ou facettisation	Le fait de calculer des graphiques en 3D peut-être divisé en 3 parties: la facettisation, la géométrie et le rendu. La facettisation est la partie consistant à découper une surface en plus petites formes, en la découpant (souvent en triangles ou en quadrilatères)
Texture Mapping	Consiste à stocker des images constituées des pixels (texels), puis à envelopper des objets 3D de cette texture pour obtenir une représentation plus réaliste des objets
Tri-linear filtering	Basé sur le principe du filtrage bilinéaire, le filtrage trilineaire consiste à faire une moyenne de deux niveaux de filtrage bilinéaire.
Z-buffer	Partie de la mémoire qui stocke la distance de chaque pixel à l'objectif. Lorsque les objets sont rendus à l'écran, le rendering engine doit supprimer les surfaces cachées.
Z-buffering	C'est le fait de supprimer les faces cachées en utilisant les valeurs stockées dans le Z-buffer