

Sommaire

Introduction

Partie I : Nécessité et principes de l'encodage vidéo

- A) Pourquoi encoder la vidéo ?
- B) Principes de l'encodage vidéo : le MPEG
 - 1) Principes de base
 - 2) Où se “déroule“ l'encodage vidéo ?
 - 3) Un exemple de compression vidéo avec perte : le MPEG
 - 4) Redondances spatiales et temporelles
 - 5) Notion de bitrate
 - 6) Images I, P, B et GOP
 - 7) Niveaux (Levels)
 - 8) Profils (Profiles)
 - 9) MP@ML ?
 - 10) Les limites du MPEG-2

Partie II : Les différents formats de vidéo numérique

- A) Conteneurs et codecs
- B) Le format AVI
- C) Le format Real
- D) Le format QuickTime
- E) Le format Matroska
- F) Le format DV
- G) Le format OGM
- H) Quelques codecs bien connus

Partie III : Peut-on affirmer qu'un format vidéo est meilleur qu'un autre ?

Lexique

Bibliographie

Introduction

L'avènement de l'informatique et la popularité des médias sous forme vidéo ont apporté la nécessité d'encoder la vidéo avant de la diffuser sous forme numérique. Pour répondre à ce besoin, divers acteurs de l'informatique (entreprises, "monde libre", ...) ont conçu différentes solutions à ce problème. Peut-on affirmer que l'une de ces solutions est meilleure que toutes les autres ?

Pour répondre à cette question, nous étudierons dans un premier temps pourquoi l'encodage vidéo s'avère nécessaire. Nous nous intéresserons ensuite aux principes techniques de l'encodage en étudiant l'exemple de l'encodage MPEG. Nous continuerons par une description non exhaustive des formats vidéos les plus courants et enfin, nous nous baserons sur ces différents éléments pour tenter de répondre à notre problématique.

Partie I

Nécessité et principes de l'encodage vidéo

A) Pourquoi encoder la vidéo ?

La vidéo est constituée de deux composantes essentielles qui sont le son et l'image. En faisant défiler une série d'images à une certaine vitesse, on parvient à recréer l'impression du mouvement ; si certains films animés peuvent être visionnés à une cadence de 12 ips (images par secondes), le cinéma lui exige 24 ips, le standard vidéo PAL (Europe et Asie) en exige 25 et le standard NTSC (Amérique et Japon) en impose 30. Lors du procédé de numérisation, ce débit d'image entraîne la production d'une grande quantité de données.

Prenons un exemple afin de mieux cerner le problème : une image de qualité DVD au standard PAL a une résolution de 720 x 576 pixels, chacun de ces pixels étant codés sur 16 voire 24 bits afin d'assurer un rendu optimal des couleurs. On arrive à une taille d'environ 1,2 Mo pour une image, soit près de 30 Mo pour une seconde de vidéo.

Ajoutons-y le son : en qualité CD et sans aucune forme de compression, une seconde de son prend 172 Ko, ce qui paraît négligeable comparé au poids de la vidéo, mais atteindra tout de même 1,18 Go pour un film de deux heures. Heureusement, les diverses techniques d'encodage audios (AC3, MP3, WMA, Ogg Vorbis, etc.) sont capables de faire tomber ce chiffre à environ 280 Mo.

Ces deux chiffres paraissent toujours bien faibles par rapport aux 210 Go requis uniquement pour stocker les images sans compression. Cette quantité de données, bien que théoriquement stockable, s'avère en pratique gourmande et ruineuse. En effet, 210 Go suffisent amplement à remplir un disque dur "grand public" et équivalent à plus de 20 DVDs. Il est donc nécessaire de trouver une solution pour réduire le poids de la vidéo : l'encodage vidéo.

B) Principes de l'encodage vidéo

1) Principes de base

L'encodage vidéo repose le plus souvent sur une compression avec perte de données ("lossy compression"), c'est-à-dire une compression qui va privilégier la réduction du poids des images au détriment de leur qualité.

Il existe des algorithmes de compression vidéo et audio sans aucune perte d'informations (on parle alors de "lossless compression") mais ceux-ci n'apportent en moyenne qu'un gain de 50 % en terme de poids. Ils sont donc surtout utilisés dans le milieu professionnel (exemple : montage vidéo) et par certains puristes ou alors dans le cadre d'utilisations spécifiques (exemple : vidéos à base de captures d'écran). On citera pour exemple les codecs HuffYUV, AviZLIB et MotionSZH).

Cette compression lossless repose sur des principes similaires à ceux de la compression de fichiers quelconques (compressions ZIP, TAR.GZ, etc.), incluant l'élimination optimale des redondances dans les données.

Par exemple, si on trouve dans un fichier trente octets d'affilée ayant la même valeur, il est beaucoup plus économique de remplacer ces trente octets par un caractère spécial suivi du nombre de redondances (ici trente) et enfin de la valeur concernée. L'optimisation de ce principe pour des

L'encodage vidéo

données vidéo aboutit à un codec de compression vidéo lossless. Exemple : si une image contient 8 pixels côte à côte ayant exactement la même couleur, le 1^{er} pixel est codé normalement et il suffit ensuite d'indiquer que les 7 suivants ont la même couleur.

D'autres solutions sont envisageables pour réduire la taille d'une vidéo. En fait il suffit de réduire la valeur d'une ou plusieurs caractéristiques de cette vidéo pour cela.

Ainsi, on peut diminuer le nombre d'images par secondes ; on a ainsi moins d'images à stocker, d'où un gain de place conséquent ; en contrepartie, la lecture apparaîtrait saccadée en dessous de 16 ips pour les personnes les plus tolérantes.

Autre paramètre sensible : la résolution de la vidéo ; on peut par exemple réduire une vidéo type DVD (720x576) à la taille d'une vidéo type VCD (352x288) mais lors de la lecture, il en découlera une nécessité d'agrandir la vidéo, ce qui provoquera un effet de pixellisation des plus désagréables.

Enfin, il est possible de diminuer le nombre de couleurs utilisées de 16 777 216 à 65 536 voire 256 couleurs mais là encore la perte serait visible au niveau des dégradés.

Un codec vidéo est, sur le plan matériel, un appareil visant à transformer un signal vidéo analogique en données numériques et vice-versa ; sur le plan logiciel, il s'agit d'un pilote de compression et décompression vidéo. Une vidéo encodée avec un codec donnée ne peut être lue que si on possède le codec en question ou du moins un codec prenant en charge la décompression.

2) Où se “déroule“ l'encodage vidéo ?

Tout encodage vidéo peut être réalisé de deux manières : solution matérielle ou solution logicielle. Dans la mesure où l'encodage peut se résumer à un traitement algorithmique ne nécessitant une intervention humaine que pour sa configuration de départ, il est possible d'effectuer ce traitement de conversion au niveau d'un composant matériel. On trouve ainsi du matériel de conversion vidéo pour les particuliers (certaines cartes graphiques, cartes consacrées à l'acquisition vidéo, etc...) et pour les professionnels. Ce dernier type de matériel est capable d'effectuer le coûteux traitement que représente la conversion d'un signal vidéo en MPEG-2 en temps réel, “prouesse” qui se ressent dans les tarifs de ces appareils.



Périphérique d'acquisition vidéo gérant l'encodage de façon matérielle.

L'autre solution, bien moins onéreuse sur le plan financier mais bien plus gourmande en temps, est la compression logicielle : elle suppose que la vidéo est déjà numérisée dans un autre format (par exemple DV (Digital video)) ; il suffit alors d'utiliser des logiciels plus ou moins performants pour exploiter des processeurs plus ou moins performants afin d'encoder la vidéo.

3) Un exemple de compression vidéo avec perte : le MPEG



Le MPEG (Motion Picture Expert Group) est un groupe d'industriels fondé en 1988 qui se réunit régulièrement afin d'élaborer des standards dans le domaine de la compression numérique audio et vidéo afin d'en faciliter la transmission par divers moyens.

De leurs réflexions sont nés différents formats, regroupés sous les références MPEG-1, MPEG-2, MPEG-4 et MPEG-7 et bientôt MPEG-21. Chacun répond à un besoin spécifique :

-le MPEG-1 permet d'obtenir une qualité légèrement supérieure à celle d'une cassette VHS et est utilisé pour les VCD (Vidéo-CD) Bien que ce support ait connu un grand succès en Asie, il est passé quasiment inaperçu en Europe.

-le MPEG-2, élaboré en 1993-94, est utilisé pour les DVDs (Digital Versatil Disc) et la TVHD (TéléVision Haute-Définition) ; il fonctionne sur les mêmes principes que le MPEG-1 mais permet de conserver une qualité d'image diffusable ("broadcast"), ce qui en fait un format très apprécié des professionnels. C'est cette norme du MPEG que nous allons détailler.

-le MPEG-3, devant l'efficacité du MPEG-2 jusque dans les très hautes résolutions, n'a pas été créé. A noter que le célèbre MP3 correspond au MPEG-1 layer III.

-le MPEG-4 n'est au départ pas destiné spécifiquement à la vidéo mais plutôt à la transmission de tous types de médias. Néanmoins l'efficacité de ses algorithmes pour les médias vidéos fait que l'appellation "MPEG-4" résonne comme un standard vidéo pour la plupart des gens.

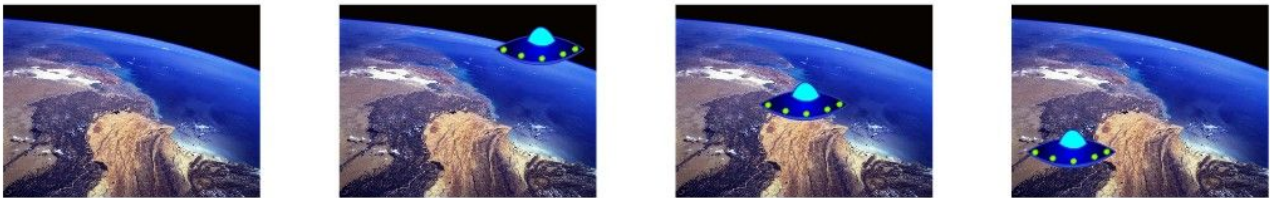
Le MPEG a d'autres projets en cours tels que le MPEG-7 et le MPEG-21. A noter que les standards MPEG sont théoriquement ouverts à des évolutions et/ou à des améliorations par le biais des fabricants de matériel et des développeurs.

4) Redondances spatiales et temporelles

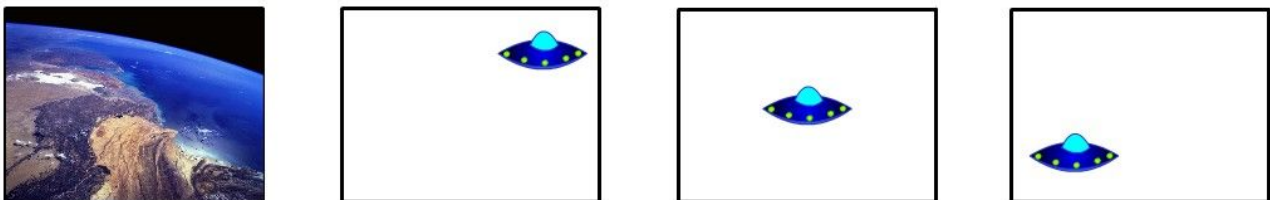
Les deux types de redondances visées par l'encodage MPEG-2 mais aussi par la plupart des codecs vidéos sont la redondance spatio-temporelle et la redondance spatiale :

-la redondance spatio-temporelle est la réapparition de certains éléments d'une image donnée dans l'image suivante. Exemple : un objet se déplace sur la longueur de l'image ; l'encodage consistera à coder la première image de la scène en entier, avec une compression classique de type JPEG (Joint Photographic Expert group) puis à ne stocker pour les images suivantes que les différences par rapport à la première image

-la redondance spatiale est la réapparition d'une partie de l'image sur cette même image ; la redondance peut être exactement la même ou alors présenter une légère différence. Selon le "bitrate" alloué à l'encodage de la vidéo, ces redondances seront plus ou moins assimilées comme étant la même.



Plutôt que de stocker intégralement les 4 images ci-dessus, on préférera stocker la première image puis ne retenir pour les suivantes que les différences par rapport à cette première image :



Une autre possibilité aurait consisté à compresser chaque image de la vidéo en JPEG puis à rassembler toutes les images ; malheureusement le gain est insuffisant et cette technique, utilisée pour le MJPEG, demande encore beaucoup d'espace mais offre la possibilité d'accéder rapidement à chaque image.

5) Notion de bitrate

Le bitrate (littéralement “taux de bit”) est le nombre de bits (et donc d'octets) qu'une seconde de vidéo devra occuper après encodage ; ce paramètre est défini par l'utilisateur avant la compression. Le bitrate s'exprime généralement en kilobits par seconde (kbps) ou Megabits par seconde (Mbps). Exemple : une vidéo d'une durée de 2 heures (7200 secondes) encodée avec un bitrate constant de 900 kbps aura une taille de 772 Mo :

$$\text{Taille finale (en kilobits)} = \text{Durée (en secondes)} \times \text{bitrate (en kbps)}$$

The screenshot shows a window titled "Nic's Mini Calc" with several sections for configuring video encoding parameters:

- Length:** Film Length: 2.00 (with a dropdown for 25.000), (e.g. Type in 1.50 for 1hr 50m) (FPS)
- Audio:** Audio Bitrate: 0 (dropdown), Audio Filesize: (input field with browse button)
- Subs Size:** Subtitle Filesize: (input field with browse button)
- Final Size:** Final Desired Size: 776 (dropdown)
- Overhead:** 1 CBR MP3, AVI Frame Overhead, 1 VBR MP3, 1 AC3, 1 (dropdown), 40ms
- Final Results:** XviD Video Size: 790315, Size in MB: 772 MB, Avg Bitrate: 899 kBit/s

Buttons: Generate, Done

Le bitrate peut être spécifié de manière constante (on parle alors de CBR pour Constant BitRate) ou de manière variable (VBR pour Variable Bitrate) ; dans le second cas, il faut généralement effectuer une première “passe” (c'est-à-dire une première phase d'encodage) pour générer un fichier de statistiques portant sur la vidéo à encoder puis une deuxième passe qui exploitera les statistiques ainsi générées pour optimiser l'utilisation du bitrate tout au long du film. Un encodage VBR est toujours plus long à réaliser qu'un encodage CBR. Dans le cas du MPEG-2, le CBR est généralement préféré au VBR.

Le bitrate étant le rapport entre la taille (donc la qualité) de la vidéo finale et sa durée, il influe lourdement sur la qualité de la vidéo et suffit généralement à la définir. Un bitrate élevé permettra à l'encodeur de créer un fichier plus lourd qui sera donc de meilleure qualité ; à l'inverse, un bitrate plus faible engendrera un fichier plus léger mais dont la qualité aura été sévèrement entamée. La principale utilité du bitrate est donc de pouvoir spécifier la taille et la qualité du fichier final avant la phase d'encodage qui s'avère généralement longue sur du matériel non professionnel

Ci-contre : un "calculateur de bitrate", c'est-à-dire une application destinée à faciliter le calcul d'un bitrate en vue d'un encodage.

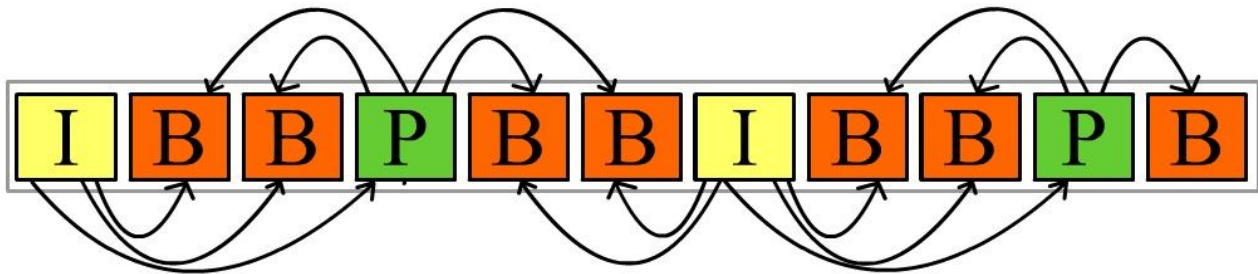
6) Images I, P, B et GOP

Lors de l'élimination des redondances spatio-temporelles, l'image de référence, c'est-à-dire la première image qui est entièrement codée, est appelée "image I" (de l'anglais Intra Frame). Les autres types d'images sont P et B. Les images P (Prédites) sont composées à partir de l'image I qui les précède. Les images B (Bidirectionnelles) sont construites à partir des images I et P qui les entourent.

Chacune de ces images est elle-même compressée avec une technique proche du JPEG. Ce type d'encodage permet de minimiser le poids des images B et P, et ce particulièrement dans les scènes statiques, où il y a peu de mouvements. En moyenne :

- une image P (Prédite) représente 50 % du poids d'une image I
- une image B (Bidirectionnelle) représente 15 % du poids d'une image I

C'est donc grâce à ces types d'images que l'on parvient à réduire le poids d'une vidéo.



Chaque image **P** est produite à partir d'une image **I**

Chaque image **B** est produite à partir d'une image **I** et d'une image **P**

Il est également possible de faire varier le nombre d'images I, B et P qui entrent dans la composition de la vidéo. Cette valeur s'appelle un GOP (Group Of Pictures = Groupe d'Images), et est fixe sur toute la longueur du film. Un GOP commence par une image de référence I et est suivi par un certain nombre d'images P et B. La taille de ce GOP influe sur le poids de la vidéo finale : plus le GOP est long (donc plus il y a d'images P et B) et plus le fichier sera léger..

On trouve des GOP d'une seule image (de type I), de deux images (sur le modèle IB ou IP), de six images (IBBPBB), de huit images (IBBPBBBP), voire de 12 images comme pour le MPEG1 (IBBPBBBPBBPBB). Il n'y a pas vraiment de règles précises dans la mesure où c'est le fabricant de l'encodeur MPEG qui définit le GOP comme bon lui semble. A noter que certains logiciels d'encodage permettent de choisir la longueur du GOP.

7) Niveaux (Levels)

Comme il a été dit dans le 1), on peut également jouer sur la résolution de l'image. Le MPEG a ainsi défini quatre résolutions d'images standards, appelés aussi Niveaux ou Levels. Ces "Levels" peuvent également porter sur le débit vidéo.

Le premier d'entre eux, appelé "Low Level" correspond à une résolution de 352x288 en PAL ou 352x240 en NTSC. Il est souvent utilisé en MPEG-1 et correspond également à la résolution d'un

VCD.

Le second, le “Main Level” (720x576 en PAL ou 720x480 en NTSC), est aussi le plus répandu, d'autant qu'il est utilisé pour la réalisation des DVD-Video.

Les deux autres formats, le “High Level” (1920x1080) et le “High 1440 Level” (1440x1152), sont destinés à la TVHD.

A partir des différents niveaux, on peut en déduire le débit maximal correspondant :

-4 Mbps pour le Low Level (abrégé LL)

-15 Mbps pour le Main Level (abrégé ML)

-60 Mbps pour le High 1440 level

-80 Mbps pour le High Level (abrégé HL)

Bien sûr, ces Niveaux ne sont que des recommandations qui n'empêchent nullement l'utilisation d'une résolution supérieure. En effet, la résolution maximale théorique du MPEG est de 16383 x 16383, c'est-à-dire de quoi remplir la surface de 256 moniteurs de PC.

8) Profils (Profiles)

Le MPEG définit également quatre types de flux MPEG-2 appelés Profils : “Simple”, “Main”, “Main+”, “Next”. Un profil détermine les caractéristiques d'un encodage, le plus courant étant le “Main Profile”. Il est compris par l'ensemble des décodeurs, utilisé pour le DVD-Video et la télévision par satellite. A lui seul, il représente plus de 95% des utilisations du MPEG-2.

9) MP@ML ?

Grâce à ces “Profiles” et ces “Levels”, on peut facilement déterminer la caractéristique d'un flux vidéo MPEG-2. Par exemple, on peut dire d'un fichier MPEG-2 qu'il est “MP@ML” (Main Profil et Main Level) ou “MP@LL” (Main Profil et Low Level). Il existe en principe 24 combinaisons possibles entre les profils et les niveaux mais toutes n'ont pas été définies.

10) Les limites du MPEG-2

De part sa structure, le MPEG-2 se limite à un format de diffusion. En effet, la plupart des images (types P et B) d'un fichier MPEG-2 devant être “calculées” avant d'être affichées, celui-ci s'avère peu utilisable pour des applications telles que le montage vidéo, où l'accès aux images doit être le plus rapide possible. Les formats destinés au montage vidéo sont donc des formats types MJPEG. Le MJPEG est un encodage relativement simple : chaque image de la vidéo est compressé en JPEG ou du moins avec une compression similaire. Il n'y a donc que des images de référence ; n'importe quelle image d'une vidéo peut être instantanément affichée sans autre calcul qu'une décompression JPEG. En contre-partie, cette forme de MPEG est beaucoup moins efficace en ce qui concerne la réduction de la taille du fichier vidéo.

Partie II

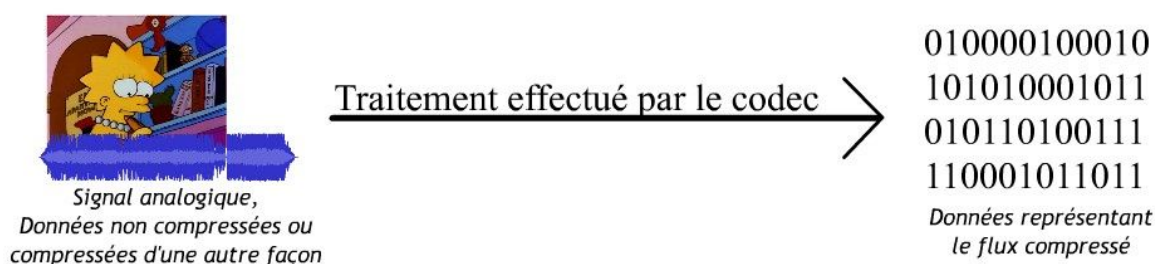
Les différents formats de vidéo numérique

Forts des principes d'encodage précédemment décrits, de nombreux formats (conteneurs et codecs) se sont développés pour la vidéo numérique. Nous tâcherons ici de nous intéresser aux plus utilisés d'entre eux, afin de cerner leurs caractéristiques, leurs avantages et leurs inconvénients.

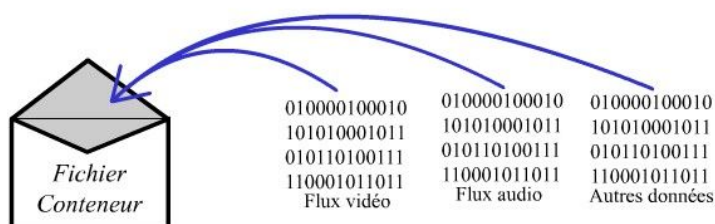
A) Conteneur et codecs

Si la plupart des codecs vidéos sont basés sur les mêmes principes de compression vidéo que ceux vus dans la première partie, la plupart des formats vidéos reposent sur une différence de conteneur ; un conteneur vidéo n'est pas un codec, il ne permet pas à lui seul de compresser une vidéo, Un conteneur est un type de fichier organisée d'une certaine manière qui peut être vue comme une enveloppe ; cette enveloppe permet de stocker dans un fichier une ou plusieurs pistes vidéos et/ou audios et parfois d'autres types de médias, comme par exemple des sous-titres, des menus, des systèmes de chapitres...

- *Un codec permet de compresser un flux audio/vidéo :*



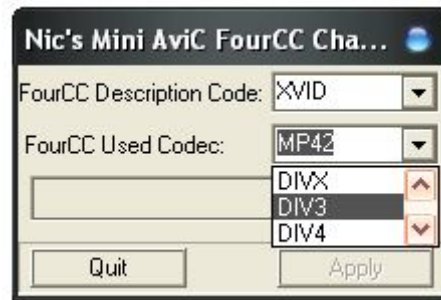
- *Un conteneur permet de stocker ces flux :*



B) Le format AVI

Le format AVI (pour Audio Video Interleave) est un container pour stocker des données audio et vidéo qui seront lues simultanément (on dit que l'image et le son entrelacés). Il a été présenté par Microsoft en novembre 1992, en tant qu'élément de son système « Video For Windows » (VFW). Dans un fichier AVI, chaque piste audio et/ou vidéo peut théoriquement être compressée par n'importe quel codec. Le codec à utiliser est défini dans le fichier en question par un code à quatre caractères appelé « FourCC ». Par exemple, une vidéo compressée avec une version 5.0 ou supérieure du codec DivX aura un FourCC égal à « DX50 ».

L'encodage vidéo



Application permettant de changer le code FourCC d'un fichier AVI donné.

Le format AVI est largement utilisé et est de ce fait lisible par de nombreux programmes sous de nombreux systèmes d'exploitation. Néanmoins dans la pratique, certains codecs (comme par exemple le codec audio Ogg Vorbis) ne sont pas intégrables dans un AVI. De plus, ce format est limité à deux pistes audios (on parle parfois de BivX pour ce genre de fichiers).

C) Le format QuickTime (MOV)



Le lecteur multimédia QuickTime d'Apple, fourni avec le système d'exploitation MacOS, est l'équivalent sur Macintosh du Lecteur Windows Media des systèmes Windows. Cela confère au format vidéo associé une solide base d'utilisateurs, ainsi qu'une assez large diffusion sur Mac. Comme le format RealVideo, QuickTime 6 s'utilise surtout en streaming, avec un lecteur téléchargeable gratuitement mais un encodeur payant (appelé QuickTime Pro).

Au chapitre des avantages, on note l'excellente intégration aux navigateurs Internet (les vidéos QuickTime sont jouées dans la fenêtre du navigateur) et les caractéristiques techniques proches des spécifications du MPEG-4. En plus de ses capacités vidéo, le lecteur QuickTime offre la possibilité de regarder des scènes en 3D (grâce au format QuickTime VR), dans une sorte de film interactif.

Les fichiers QuickTime sont généralement caractérisés par une extension *.mov. Ce type de fichier présente tout de même des inconvénients : d'une part, il n'est pas des plus performants quant à la compression vidéo (utilisation du codec 3ivX) et d'autre part il s'avère difficilement convertible en un autre format.

D) Le format Real (RM)

Le format RealVideo s'est largement imposé dans le domaine du streaming. Le streaming vidéo consiste à charger une partie de la vidéo en cache et à commencer la lecture alors que la vidéo n'est pas encore entièrement chargée. Le reste de la vidéo est téléchargée pendant la lecture elle-même. Le streaming est d'autant plus efficace que la bande passante disponible pour le chargement est élevée. Sur de faibles bandes passantes, il s'avère par contre beaucoup moins efficace, avec des pauses et des saccades d'images lors de la lecture.

L'encodage vidéo

De très nombreux sites Web ont choisi ce format pour la présentation en ligne de clips vidéo, d'émissions de télévision ou de bandes-annonces de films de cinéma. Techniquement, la dernière mouture RealVideo 9 (ou RV9) se révèle séduisante, avec des taux de compression très élevés et une excellente qualité d'image.

En contrepartie, l'encodage des vidéos est très long, même avec un ordinateur puissant. De plus, ce type de format est destiné à la vidéo en ligne et s'avère difficilement récupérable sur son propre ordinateur (à moins d'utiliser un logiciel capable d'intercepter et d'enregistrer le flux vidéo en question).



E) Le format Matroska (MKV)



Le format Matroska (correspondant à l'extension *.mkv comme Matroska Video) est un format vidéo libre. Plus exactement il s'agit d'un conteneur (le nom est en fait une référence aux poupées russes) permettant de contenir de la vidéo, du son et des sous-titres etc. dans un même fichier. Le format MKV est basé sur une structure dérivée du XML, appelée EBML (*Extensible Binary Meta Language*). Ainsi grâce au format Matroska, il est notamment possible de réaliser des fonctions de chapitrage, de créer des menus, de faire des recherches dans le fichier, de sélectionner une source

sonore ou bien de choisir un sous-titrage.

La particularité du Matroska est qu'il accepte la plupart des codecs vidéos existants à savoir RV9, RV10, DivX, XviD, MPEG, et bien d'autres. Il en va de même pour les formats audios : un fichier Matroska peut contenir de nombreuses pistes audios en Ogg Vorbis, AAC, AC3, MP3, DTS, PCM. Son système de sous-titres intégrés évite de devoir laisser un fichier *.srt ou *.sub avec le fichier principal, Ce format se targue également d'avoir un faible overhead. Par "overhead", on entend la quantité d'informations supplémentaires nécessaire au rassemblement des différents médias dans le conteneur.

F) Le format DV (AVI)

Le format DV (Digital Video = Vidéo Numérique) est principalement utilisé pour le montage vidéo amateur, particulièrement avec les caméscopes utilisant des cassettes "MiniDV" (illustration ci-contre). Il est basé sur des principes similaires au MJPEG pour l'aspect vidéo et sur une compression PCM au niveau audio, ce qui correspond à la plus simple et à la plus gourmande façon de stocker des données audios. Son ratio de compression fixe confirme qu'il a bien été conçu pour le montage vidéo et non pour la diffusion : 2 Go de fichier DV correspondent à environ 9 minutes de vidéo, en qualité DVD pour l'image, avec une perte de qualité quasi-inexistante.



L'encodage vidéo

La particularité de ce format est qu'il peut être numérisé sous forme de fichier AVI Type-1 ou Type-2. Au format Type-1, les flux vidéo et audio provenant du caméscope sont sauvegardés sous forme d'un seul flux unique de données brutes. Ces données brutes correspondent exactement au flux DV reçu par le PC lors de l'acquisition. C'est-à-dire l'audio et la vidéo multiplexées ensemble. Le programme d'acquisition se contente de rajouter une en-tête AVI au début du flux de données. Au format Type-2, le flux audio et vidéo sont séparés et sauvegardés l'un à la suite de l'autre, comme 2 flux différents au sein du fichier AVI. De cette différence découlent des avantages et des inconvénients : par exemple, le type-1, bien que pouvant être envoyé directement au caméscope, s'avère incompatible avec la plupart des éditeurs AVI. Le type-2, beaucoup plus compatible, requiert plus de puissance processeur afin d'être envoyé au caméscope

G) Le format Ogg Media (OGM)



Le Ogg Media est un format de fichier qui accepte tous les codecs vidéos mais a la particularité de n'accepter que des pistes audios en Ogg Vorbis (l'équivalent libre du MP3, voir logo ci-contre). Ce format audio est un peu plus efficace que le MP3, ce qui permet soit d'accroître la qualité audio pour une même taille de fichier soit de réduire la taille de fichier pour une même qualité audio. L'Ogg Media permet également d'ajouter des sous-titres et des chapitres à l'intérieur d'un même fichier vidéo. A noter également que, contrairement au format AVI, l'OGM supporte plus de deux pistes son. L'utilisation de l'Ogg Vorbis évite aussi quelques problèmes de décalage image / son parfois rencontrés lors de certains encodages vidéos avec une piste audio en MP3 (la solution à ce problème pour le format AVI consiste à modifier le nombre d'images lues par seconde).

H) Quelques codecs bien connus

Peut-on parler encodage vidéo sans parler du codec le plus connu, c'est-à-dire le DivX ? et son homologue libre le XviD ? Ces codecs, ainsi que quelques autres moins connus, sont l'adaptation pour les formats types AVI des principes de compression vidéo du standard multimédia MPEG-4.

Le codec DivX a été créé en 1999 par Jérôme Rota, développeur français de 27 ans, en se basant sur le codec MPEG-4 de Microsoft. Au départ on parlait du format "DivX ;-)", sans oublier le clin d'oeil à droite. Ce clin d'oeil est une référence à un ancien système de location vidéo qui n'aboutit jamais : "divx" (Digital Video Express). En 2001, Jérôme Rota fonda l'entreprise "DivX Networks" et produisit un nouveau codec propriétaire débarrassé des droits de Microsoft : DivX 4. DivX Networks intégrait une branche libre (OpenSource) appelée "OpenDivX". Lors du passage à la version 5, DivX est carrément devenu payant (une version gratuite est disponible mais accompagné d'un spyware).

Xvid fut créé en 2001, lorsque des développeurs, mécontents de l'utilisation qui avait été faite du projet OpenDivX, décidèrent de continuer leur travail dans un projet réellement OpenSource.



Partie III

Peut-on affirmer qu'un format vidéo est meilleur qu'un autre ?

Si l'on résume ce qui a été exposé précédemment, on arrive à deux grands principes :

-la plupart des codecs et formats vidéos sont basés sur les mêmes principes (en l'occurrence ceux définis par le MPEG)

-les différences entre formats se font surtout sur des critères d'utilisation possible, de taille, de compatibilité et de popularité.

Les critères d'utilisation concernent l'application pratique première d'un format : peut-on l'utiliser en streaming ? peut-on y inclure des sous-titres en polonais et une 5^{ème} piste audio en tchèque ? etc.

Les critères de taille concernant les données audio et vidéo elle-mêmes varient peu pour une même qualité d'image et / ou de son. On peut par contre s'intéresser à la taille occupée par les données complémentaires, ce qui varie selon les formats. Par exemple, pour cette question, le format MKV l'emporte sur le format AVI.

Les critères de compatibilité concernent à la fois les logiciels de lecture ou codecs requis pour lire la vidéo et les systèmes d'exploitation sur lesquels cette lecture est permise ; on peut également prendre en compte la configuration matérielle nécessaire mais de nos jours, ce paramètre a moins d'importance puisque la plupart des ordinateurs ont largement la puissance nécessaire pour décoder n'importe quel type de vidéo à vitesse réelle.

La popularité est un critère sensible parfois dépendant des autres ; sur ce plan, l'AVI l'emporte sur le plan général, mais si on se restreint par exemple au monde Mac, le format QuickTime sera sans doute plus "populaire". La popularité d'un type de fichier est facilement "mesurable" chez une personne donnée ; si le système d'exploitation est déjà configuré de manière à lire directement le type de vidéo concerné, alors on peut affirmer que ce type d'encodage est accepté par cette personne. De grandes différences de popularité sont possibles : par exemple, un AVI encodé en DivX 5, avec une piste son en MP3 a bien plus de chance d'être lisible (sans configuration supplémentaire quelconque) qu'un fichier Matroska avec la piste vidéo en RV9 et deux pistes sons en Ogg Vorbis, ne serait-ce que parce que la gestion de l'AVI est incluse dans le système d'exploitation le plus courant (Windows).

Comment, à partir de là, peut-on choisir correctement le type d'encodage qui conviendra le mieux à la diffusion de telle ou telle vidéo ? Simplement en définissant un contexte. Un codec pourra se révéler excellent dans un contexte donné et inutilisable dans un autre. Exemples : Le format QuickTime s'avère excellent pour diffuser en streaming une vidéo depuis un forum portant sur les Macintosh. Le format Matroska est parfait pour diffuser via réseau P2P un épisode de manga sous-titré en 25 langues. Mais si on inverse les deux situations (Matroska pour le streaming et QuickTime pour l'intégration des sous-titres), on aboutit à une mauvaise diffusion.

Conclusion : D'où la réponse à la problématique posée au début de cette étude : Peut-on affirmer qu'un format vidéo est meilleur qu'un autre ? OUI, si l'on définit un contexte d'application particulier. OUI si on les compare pour un type de diffusion particulier ; bien sûr, il est à priori impossible de comparer deux formats dans l'absolu ; dans la plupart des cas, l'un sera plus optimal sur le plan technique mais beaucoup moins connu donc finalement moins envisageable...

LEXIQUE

Ce lexique réunit les définitions de termes utilisés dans le dossier qui n'auraient pas été explicités ou qui auraient été considérés comme déjà acquis :

-bit : C'est l'unité binaire de quantité d'information qui peut représenter deux valeurs distinctes : 0 ou 1.

-JPEG : Joint Photographic Expert Group ; technique de compression des images fixes très largement utilisée sur Internet.

-octet : un octet est un groupe de 8 bits ; comme un bit n'est jamais seul sur un support informatique, l'octet est la plus petite unité de mesure de taille pour un fichier ; de façon approchée, un octet correspond à un caractère. Les multiples de l'octet (Kilo-octet (Ko), Mega-octet (Mo) etc.) s'expriment non pas par multiple de 1000 mais par multiples de 1024 : 1 Mo = 1024 Ko ; 1 Ko = 1024 octets.

-PCM : Type de stockage de données audios ; le format PCM est un format non compressé (donc sans perte), correspondant au type de fichier "Wave" par défaut (*.wav).

-pixel : abréviation pour "picture element" ; un pixel est le plus petit élément d'une image numérique ; il correspond à un point sur un écran.

BIBLIOGRAPHIE

Les connaissances utilisées pour rédiger ce document proviennent exclusivement de sites Internet ; voici les adresses des plus utiles d'entre eux :

<http://www.encodage-video.com/>
<http://www.lahiette.com/biboobox/?Encodage+video>
http://www.dvforever.com/article.php3?id_article=48&artsuite=2
http://www.nidinfo.com/html/repor_3_divx.html
http://www.repaire.net/navig/site/mont_choisir/essentiel/02.php
<http://cddts.free.fr/articles.php?lng=fr&pg=35>
<http://pages.videotron.ca/danjean/>
<http://www.photovideomag.com/>

Nous tenons à remercier les auteurs de ces pages pour leur clarté et l'aide qu'elles nous ont procuré.