

# Comment ça marche ?

## Les transmissions radio-numériques

### 18 – Le codage de l'image animée (1)

Par le radio-club F6KRRK

*Après avoir vu le codage des images fixes, nous allons poursuivre avec le codage des images animées en commençant par le principe général MPEG. Comme pour les précédents articles sur les codages de la parole et des images fixes, il ne s'agit pas d'exiger du lecteur une compétence en traitement du signal. Ces articles n'ont d'autres ambitions que de donner une idée générale des processus employés.*

#### Codage de l'image animée

Même compressé JPEG, un fichier d'images animées de résolution TV standard reste conséquent pour être transmis par radio à 25 images seconde. Ce procédé appelé "MJPEG" n'a guère été utilisé qu'avec des enregistrements informatiques pour faciliter le montage d'un film "à l'image près".

En 1990, le besoin de stocker sous forme numérique et de reproduire des images animées et le son associé a conduit l'ISO à former un groupe d'experts appelé "MPEG" (Moving Pictures Experts Group). Le travail de ce groupe s'est concrétisé en 1992 par une norme connue sous le nom de "MPEG-1".

Cette norme était prévue pour une application multimédia. L'objectif était d'atteindre un débit constant total de 1,5 Mb/s (débit d'un CD-ROM simple vitesse) avec 1,15 Mb/s pour la vidéo MPEG-1 et 350 kb/s pour le son en stéréo et les données auxiliaires.

Le codage vidéo est basé sur le JPEG avec pertes, auquel on rajoute de nouvelles techniques réunies dans une sorte de "boîte à outils" MPEG-1. Ces techniques dites de "prédiction avec compensation de mouvement" permettent de déduire avec un minimum d'informations additionnelles la plupart des images de celles qui la précèdent, voire qui la suivent si l'on tolère un traitement en temps différé. Ceci nécessite dans le codeur un dispositif d'estimation de mouvement qui est la partie la plus complexe. Heureusement cette fonction n'est pas nécessaire dans le décodeur.

Par ailleurs, la nécessité d'un temps de synchronisation relativement court et un temps acceptable pour l'accès aléatoire à une séquence (0,5 s maxi) limite le nombre d'images pouvant dépendre d'une même première à une douzaine pour un système à 25 im/s.

#### *Les types d'images en MPEG*

MPEG définit trois types d'images qui s'enchaînent selon la figure 1 <sup>[1]</sup>.

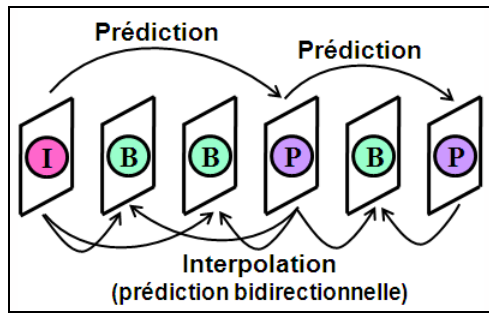


Figure 1 : Enchaînement des trois types d'images MPEG

### 1 - Images I (Intra)

Elles sont codées sans référence à d'autres images, comme en JPEG avec un taux de compression comparable. Elles contiennent tous les éléments nécessaires à leur reconstruction et sont le point d'entrée obligatoire d'une séquence.

### 2 - Images P (Prédites)

Elles sont codées par rapport à l'image de type I ou P précédente, grâce aux techniques de prédiction avec compensation de mouvement. Comme celle-ci n'est pas parfaite, on ne peut multiplier indéfiniment le nombre d'images P entre deux images I, car étant utilisées pour coder d'autres images P ou B, elles propagent les erreurs en les amplifiant. Leur taux de compression est nettement plus important que celui des images I.

### 3 - Images B (Bidirectionnelles)

Elles sont codées par interpolation entre deux images de type I ou P précédente et suivante qui les encadrent. Elles ne propagent pas les erreurs de codage et elles ont le taux de compression le plus élevé.

Deux paramètres M et N définissent la manière dont les images I, P et B s'enchaînent :

- M est la distance (en nombre d'images) entre deux images P (Prédites) successives.
- N est la distance entre deux images I (Intra) successives.

Dans le cas du CD-ROM, M=3 et N=12.

## Décomposition en couches d'une séquence vidéo MPEG

MPEG définit une hiérarchie de couches (layers) à l'intérieur d'une séquence d'images, illustrée par la figure 2 <sup>[1]</sup>.

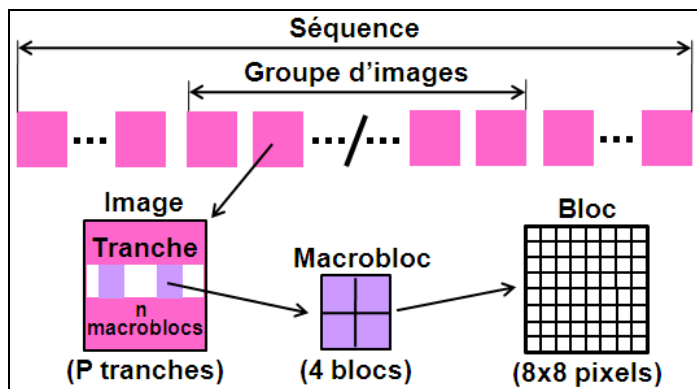


Figure 2 : Hiérarchie des couches MPEG, de la séquence au bloc

On trouve les couches suivantes :

- 1 – **Séquence**. C'est la couche la plus élevée. Elle détermine le contexte dans lequel cette séquence est définie (paramètres de base, tels que norme de balayage, etc.).
- 2 – **Groupe d'images (Group Of Pictures)**. Elle permet l'accès aléatoire à la séquence avec une première image de type I. Dans le cas du CD-ROM, le groupe comporte 12 images.
- 3 – **Image (picture)** de type I, P ou B. C'est la couche d'affichage élémentaire.
- 4 – **Tranche (slice)**. Elle est définie comme une suite de macroblocs contigus. L'ensemble des "slices" doit couvrir l'image entière, sans chevauchement. En pratique une "slice" couvre une rangée H complète de macroblocs.
- 5 – **Macrobloc (macroblock)**. De taille 16×16 pixels, le macrobloc est formé de 4 blocs de luminance et de 2 blocs de chrominance 1×Cr et 1×Cb, chacun de 8×8 pixels. C'est la couche où s'effectue la prédiction avec compensation de mouvement.
- 6 – **Bloc (block)**. Sa taille est de 8×8 pixels. C'est, comme en JPEG, la couche où s'effectue la DCT. En raison du découpage de l'image en un nombre entier de macroblocs, la définition H du MPEG-1 est réduite à 352 pixels (22 macroblocs) et la définition V à 288 pixels (18 macroblocs).

### **Estimation de la compensation de mouvement <sup>[1]</sup>.**

Entre deux images successives, les objets en mouvement entraînent une différence entre les zones de départ sur la première image et les zones d'arrivée sur la deuxième. L'estimation de mouvement consiste à trouver un vecteur assurant la correspondance des zones de départ (1<sup>ère</sup> image) et d'arrivée (2<sup>ème</sup> image). Cette recherche se fait au niveau du macrobloc (16×16 pixels) en déplaçant à l'intérieur d'une fenêtre un macrobloc de l'image en cours au voisinage de cette position dans l'image précédente, de manière à trouver le plus ressemblant (block matching), ce qui donne un vecteur de mouvement qui s'applique à toutes les composantes du macrobloc (Y, Cb, Cr). Seuls les macroblocs différents d'une image à l'autre auront besoin d'être codés d'où une réduction de la quantité d'information à transmettre.

La comparaison se faisant entre une image P et une image I ou deux images P, les vecteurs de mouvement peuvent être d'amplitude relativement importante. Pour les images B, les vecteurs intermédiaires sont obtenus par interpolation compte tenu de leur position temporelle entre les images I et/ou P.

On calcule ensuite la différence entre le bloc à coder et le bloc que l'estimation de mouvement a trouvé le plus proche et cette différence est transmise après un codage similaire à celui des blocs des images I.

La taille moyenne des objets étant supérieure au macrobloc, il y a généralement corrélation entre les vecteurs de mouvement associés à deux macroblocs contigus. D'où l'intérêt d'utiliser une méthode de codage différentielle (DPCM) avec codage entropique pour transmettre ces vecteurs.

### **Régulation de débit <sup>[1]</sup>.**

Pour s'adapter au canal de transmission et au buffer d'entrée spécifié pour le décodeur MPEG de référence, le débit doit généralement être constant en sortie du codeur. Ceci n'est pas garanti par le processus que nous avons vu, étant donné la nature très diverse des images à coder et la variabilité des mouvements qu'elles contiennent. Pour réguler le débit, on utilise en fin de processus une mémoire tampon (FIFO) dont on surveille l'état de remplissage de manière à le maintenir entre certaines limites en jouant sur la résolution des coefficients de quantification dont l'influence sur le débit est majeure. On obtient ainsi un débit constant avec une résolution qui sera d'autant meilleure que l'image sera plus calme, soit une bonne adaptation physiologique à la vision humaine.

Voir sur la figure 3 un schéma synoptique du codeur / décodeur MPEG-1 qui ne donne qu'une faible idée de sa complexité réelle.

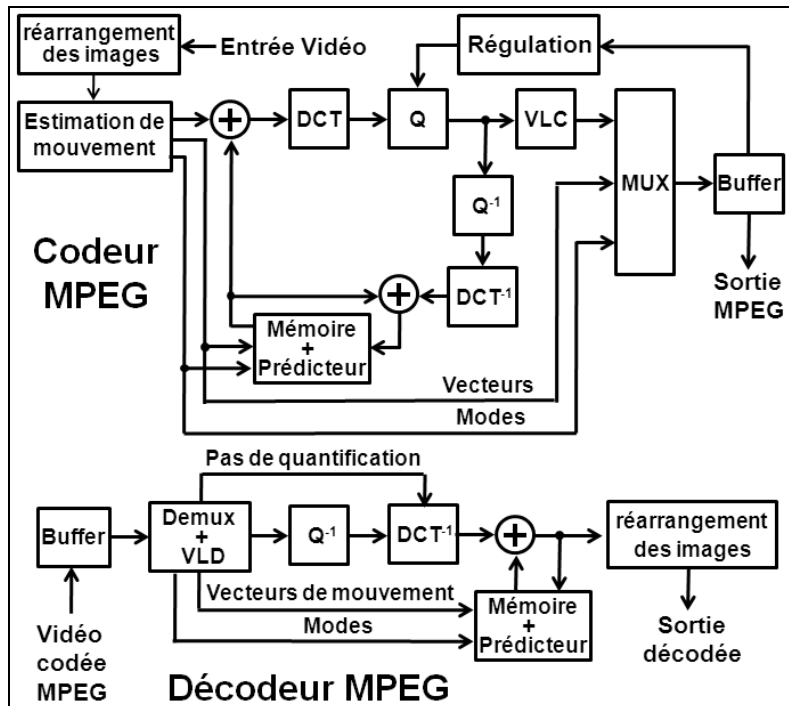


Figure 3 : Schémas simplifiés d'un codeur et d'un décodeur MPEG-1

Le décodeur n'ayant pas à accomplir d'estimation de mouvement est beaucoup plus simple, ce qui était l'objectif essentiel lors de la définition de la norme.

Dans le prochain "Comment ça marche", nous poursuivrons avec les normes MPEG-2 et la suite.

**La Rubrique "Comment ça marche" est une activité collective du radio-club F6KRK (<http://www.f6krk.org>). Pour une correspondance technique concernant cette rubrique : "f5nb@orange.fr".**

### Bibliographie

Nous nous sommes principalement servi de cette référence :  
 [1] – La TELEVISION NUMERIQUE par Hervé Benoit (DUNOD)