

LES FORMATS

NUMERIQUES

TROISIEME VERSION

Dossier repris de :

- Alexandre Polasek Bourgougnon

➤ mcfly2@voilà.fr

- Nicolas Voisin

➤ nicolas_voisin@caramail.com

Complété par
Vanessa Louiset
Élodie Korecki

D'après le support: Les formats Numériques
Première version (avril 2000)
Réalisé par: Matthieu Coupez
Johann Henry

BTS AUDIOVISUEL - option Image
Lycée Henri MARTIN - Saint Quentin (02 100)

Septembre 2001

LES FORMATS NUMERIQUES

SOMMAIRE

I. COMMENT OBTENIR UN SIGNAL BINAIRE

1. Echantillonnage
2. Quantification
3. Codage

II. LA CAPTURE DES IMAGES.

1. Signaux analogiques et numériques.
2. Définition et résolution.
3. Les capteurs Mono et Tri C.C.D..
4. La définition limitée par le format vidéo.

III. EXPLICATIONS TECHNIQUES.

1. Le brassage des données et corrections d'erreurs.
2. Spécificités du signal vidéo.
 - a) L'enregistrement hélicoïdale.
 - b) L'azimut.
 - c) Les pistes longitudinales.

IV. LA COMPRESSION NUMERIQUE.

1. Pourquoi une compression ?
2. Comment fonctionne la compression numérique ?
3. Les redondances de l'image vidéo.
 - a) Redondance spatiale.
 - b) Redondance temporelle
 - c) Redondance subjective.
 - d) Redondance statique.
4. Quelques chiffres.
5. Les normes de compression.
 - a) JPEG.
 - b) MPEG-1.
 - c) MPEG-2.
 - d) MPEG-2 MP@ML.
 - e) MPEG-2 422P@ML.
 - f) « The Mole »
 - g) MPEG- 4
 - h) MPEG- 7
 - i) MPEG-IMX
 - j) SDTI - SDI

V. LES FORMATS NUMERIQUES.

1. Les premiers formats numériques.
 - a) D1.
 - b) D2.

- c) D3.
- d) D5.
- e) D5 HD.
- f) D6.
- 2. La famille Betacam.
 - a) Digital Betacam.
 - b) Betacam SX.
- 3. HDCam.

- 4. La famille DV.
 - a) DV.
 - b) DVCAM.
 - c) DVCPRO 25.
 - d) DVCPRO 50.

- 5. Digital S : D9.
- 6. Le Digital 8.
- 7. Le D-VHS.
- 8. Le superVCD
- 9. Le MP3
- 10. Le DAB
- 11. Le Div-X

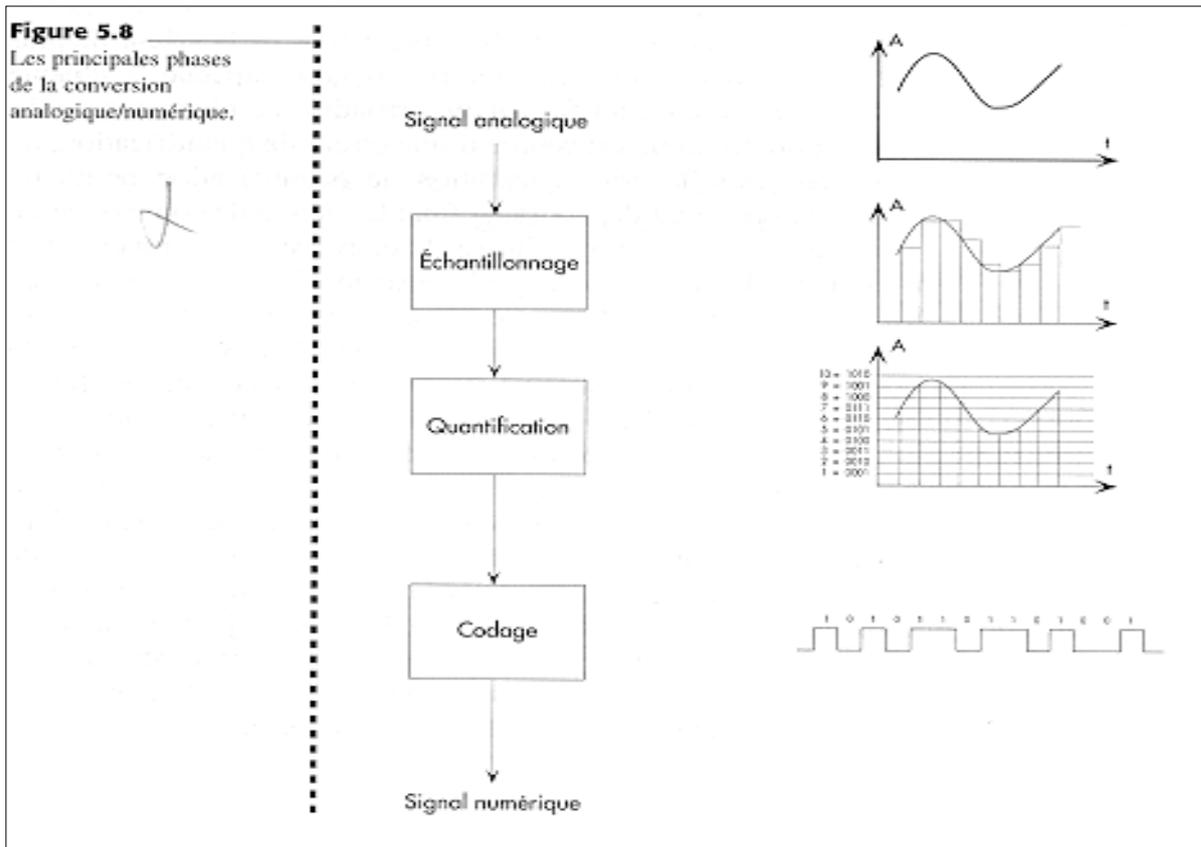
VI. TABLEAUX RECAPITULATIFS.

VII. LEXIQUE.

Sources :

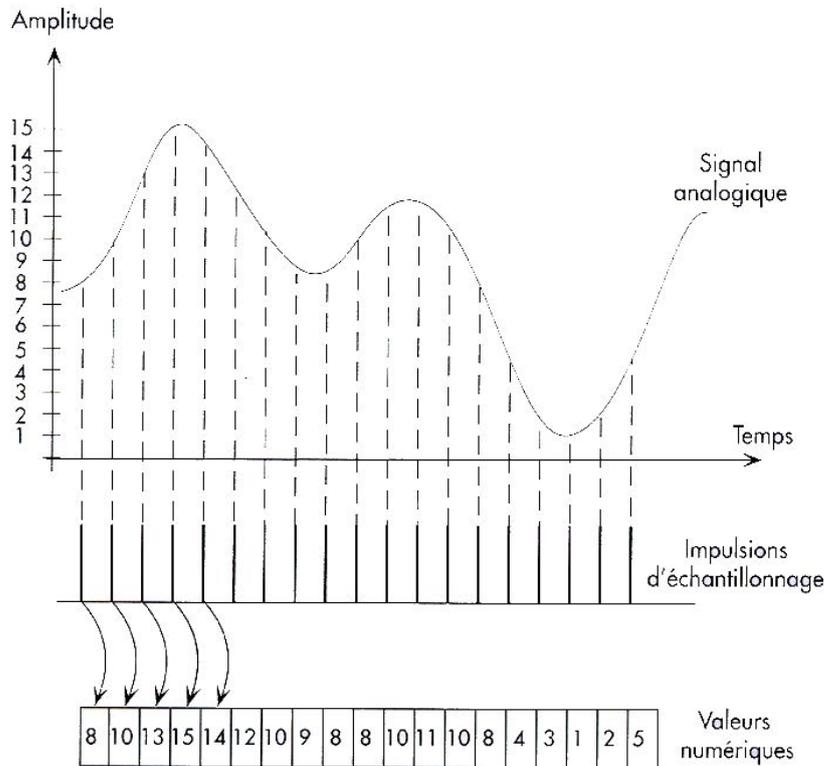
Les secrets de l'image vidéo, Ed. Eyrolles
Video Pratique, n°58 sept-oct 1999
Le technicien Film et Video, n°498
www.pro.sony-europe.com
www.panasonic.com
www.cst.fr

I. COMMENT OBTENIR UN SIGNAL BINAIRE



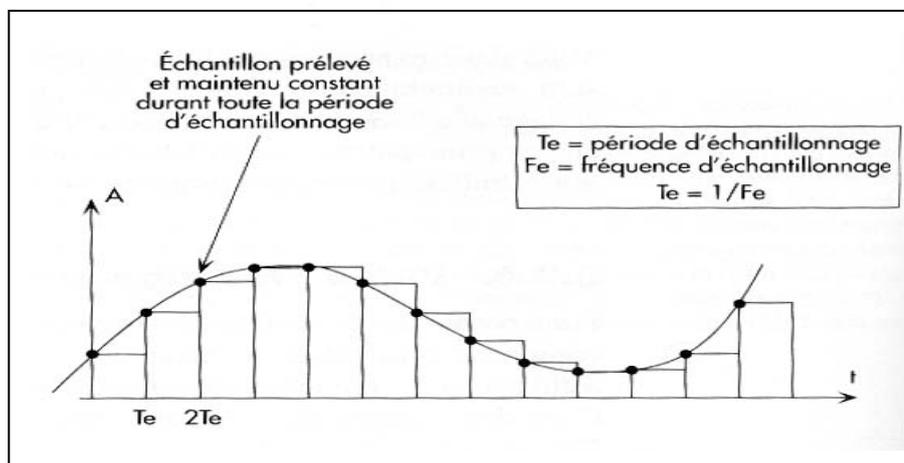
Signal analogique, signal numérique.

En vidéo, la caméra est chargée de la transduction de la lumière en courant électrique ou, plus simplement, elle convertira une information optique en information électrique. A la sortie du capteur de la caméra (capteur CCD) nous nous trouvons en présence de trois signaux fondamentaux: le rouge, le vert et le bleu à partir desquels il est possible de reconstituer toute image en couleur. Ces signaux fournis par la caméra sont analogiques; pour convertir un signal analogique en signal numérique, nous procéderons en trois étapes:



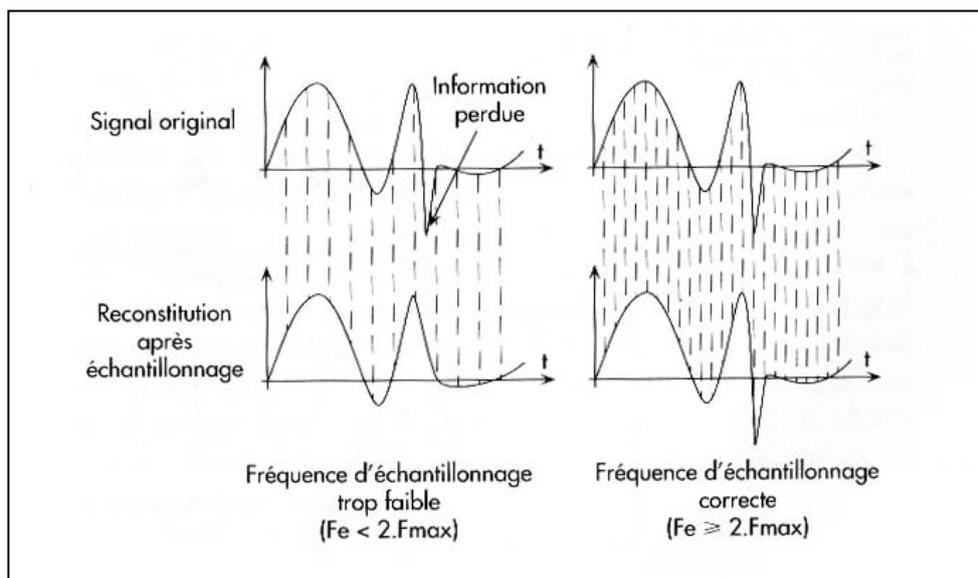
1) l'échantillonnage

Préserver l'information: Echantillonner un signal variant de façon continue consiste à le découper en tranches temporelles et prélever ponctuellement ses valeurs à des instants réguliers.



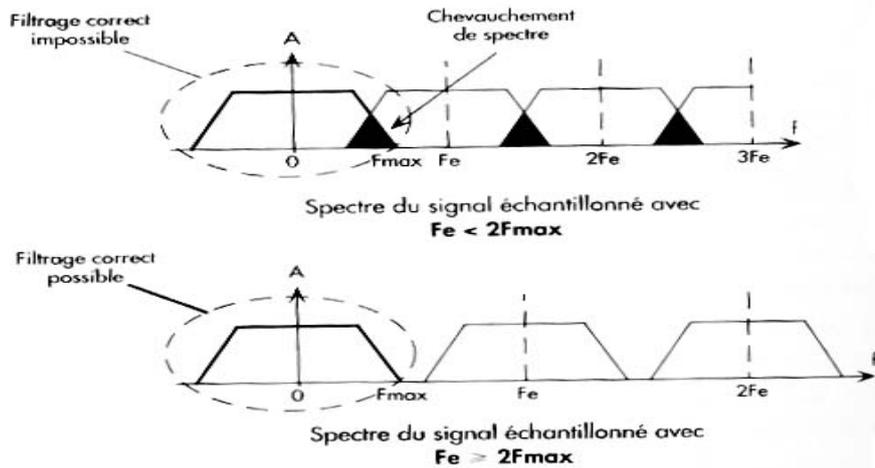
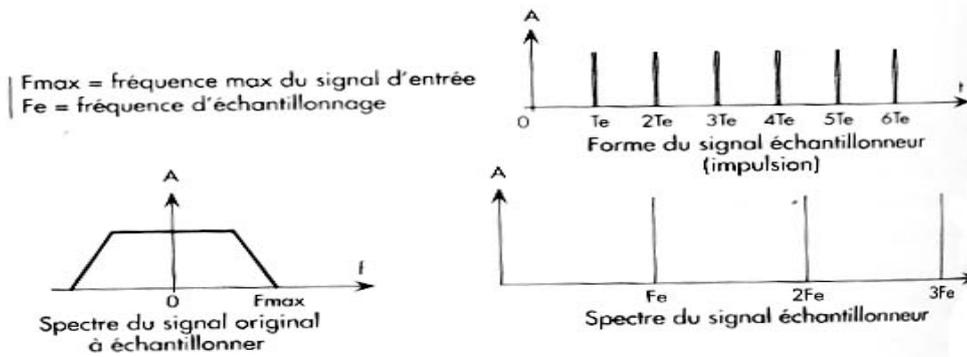
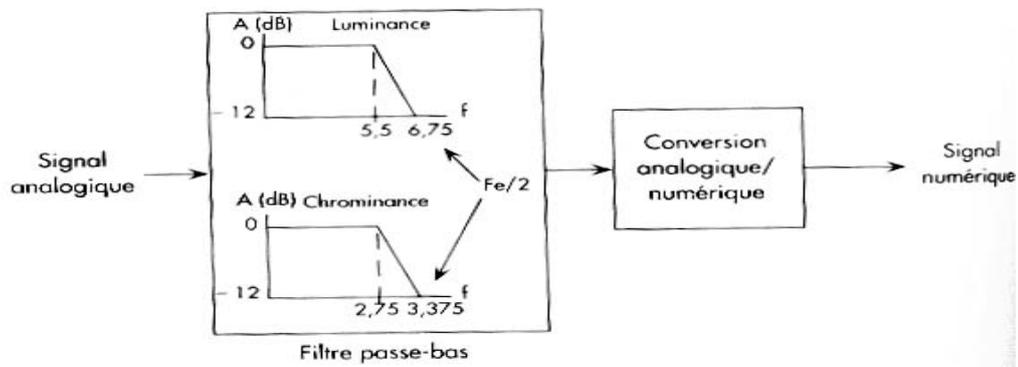
On pourrait à priori penser que la succession d'échantillons ainsi générée est déficitaire en information par rapport au signal d'origine. En fait, si on respecte un ensemble de conditions concernant la bande passante du signal et l'espace maximal entre les échantillons, l'information portée par ces derniers sera une représentation intégrale du signal analogique. Tout doit être fait pour que le rythme de découpe soit suffisamment rapide pour être à l'affût de la variation la plus brève du signal d'origine. Si la distance entre deux échantillons est supérieure à la plus rapide des variations du signal analogique, cette dernière passera inaperçue et sera définitivement ignorée. Il faut comprendre qu'entre deux échantillons, on ignore tout du signal d'origine. Autrement dit, pour échantillonner un signal en préservant toute son information, il est nécessaire de connaître au préalable la fréquence la plus élevée à laquelle il est susceptible de

varier. On admet que cette fréquence maximale est de 6 MHz pour la vidéo et de 20 KHz pour l'audio. Une loi mathématique établie par Shannon permet de déterminer la fréquence d'échantillonnage minimale à choisir pour ne pas manquer la plus petite (en terme de durée) des informations à saisir: le signal échantillonné sera la représentation exacte du signal original si la fréquence d'échantillonnage F_e est au moins supérieure à deux fois la fréquence maximale du signal. Si ce critère n'est pas vérifié, les composantes spectrales répétitives du signal échantillonné ne sont pas assez espacées et se chevauchent. La zone d'interférence, appelée zone de repliement ou "Aliasing", donne naissance à des fréquences aberrantes dans la bande utile. C'est pourquoi un filtre passe-bas, dit anti-repliement ou anti-aliasing, à pente très abrupte, est impérativement requis en amont du processus de conversion.



Le filtre anti-repliement

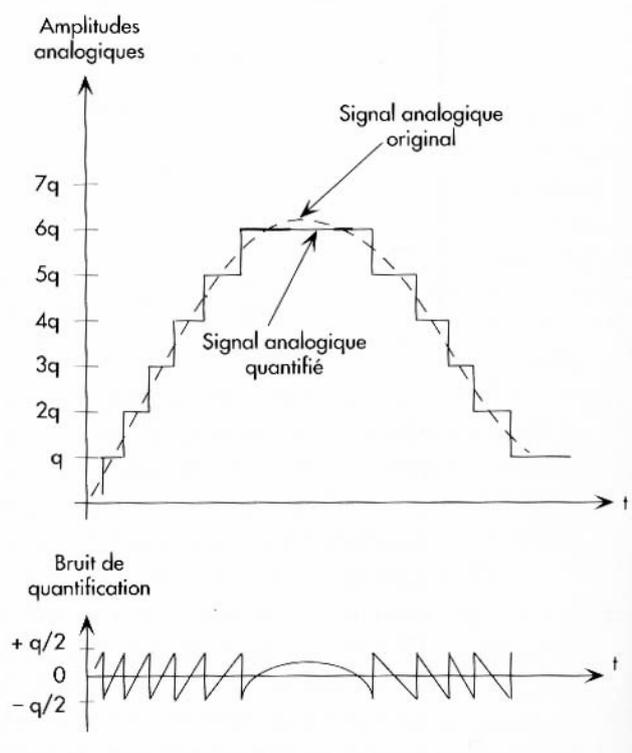
Il a pour but de rejeter toutes les fréquences indésirables du signal d'entrée qui sont supérieures à la moitié de la fréquence d'échantillonnage choisie. Il élimine les détails extrêmement fins de la scène captée par l'objectif, réduisant ainsi le phénomène d'Aliasing. Ce filtre (condensateur en bas) doit être caractérisé par un gain constant dans la bande utile et par un affaiblissement rapide aux fréquences supérieures à la demi-fréquence d'échantillonnage.



2) Quantification

Principe de base: Elle a pour but de faire correspondre à chaque amplitude mesurée un nombre entier exprimé en base 2, dans laquelle n chiffres permettent de coder $N = 2^n$ valeurs distinctes. Au cours de cette étape de quantification, le signal analogique, qui peut prendre une infinité de valeurs, est converti en un signal constitué d'un nombre fini N de valeurs numériques codés sur n numériques. On comprend alors que des erreurs d'arrondis par défaut ou par excès se produiront inévitablement. Car à plusieurs valeurs proches mais différentes, du signal analogique, correspondra une seule et même valeur numérique binaire. En fonction de la plage d'amplitudes à quantifier, on définit une échelle constituée d'un nombre fini d'intervalles " q " appelés "pas de quantification" ou "échelons de quantification", ou encore "quantums". A chaque échelon $q, 2q, 3q, \dots$ est associée une valeur numérique. A un instant t , l'amplitude du signal se trouvant à l'intérieur d'un échelon est remplacée par la valeur de cet échelon; la valeur exacte de l'amplitude n'est pas prise en considération. Il est évident que plus les échelons sont petits, plus ils sont nombreux sur une plage donnée, et plus la précision du signal quantifié est grande. Quoiqu'il en soit, l'erreur de quantification est toujours inférieure au quantum q , qui est en fait la plus petite valeur mesurable par le convertisseur. De manière générale, plus le nombre utilisé

pour la quantification est élevé, plus la précision du système de conversion est grande, mais le temps de traitement est long. L'expression mathématique du quantum en fonction de l'amplitude maximale du signal d'entrée ($V_{max}-V_{min}$) et du nombre de bits de quantification n est la suivante: $q = (V_{max}-V_{min}) / 2^n$



La correction d'erreurs: si toute la chaîne de traitement vidéo fonctionnait - idéalement- sans aucune perturbation, les signaux numériques pourraient être enregistrés, traités et transmis tels quels avec une totale fidélité. Dans la pratique, il faut cependant tenir compte du comportement des supports d'enregistrement et des voies de transmission, qui présentent parfois des défaillances. Des erreurs peuvent alors être introduites dans le flux de données numériques et modifier le contenu du message. Sur un magnéto, les principales sources d'erreurs sont les irrégularités du mécanisme d'entraînement (gigue), les fluctuations du niveau du signal, mais aussi la poussière, un défaut d'enduit magnétique, un léger froissement de la bande, etc....

En transmission, les dégradations pouvant affecter le signal sont essentiellement dues au bruit et aux distorsions: écho, évanouissement, diaphonie, inter-modulation, parasite, etc... Une erreur en numérique, c'est soit une inversion de valeur binaire dans un mot, soit une absence fugitive de données. Une erreur est d'autant plus gênante qu'elle affecte les bits les plus significatifs (MSB).

Il existe aujourd'hui des systèmes de correction très performants, capables de détecter et de corriger un grand nombre d'erreurs. Le problème est en fait de repérer quand un "1" a pris la place d'un "0", et vice versa, parce qu'une fois le bit faux détecté, sa correction est immédiate. Pour permettre un contrôle de la validité des infos transmises, des données supplémentaires sont ajoutées aux données utiles lors du codage, répondant à une loi connue du codeur et du décodeur. Au cours du décodage, à chaque fois que cette loi n'est pas vérifiée, un processus de détection, puis de correction des bits erronés est déclenché. (figure).

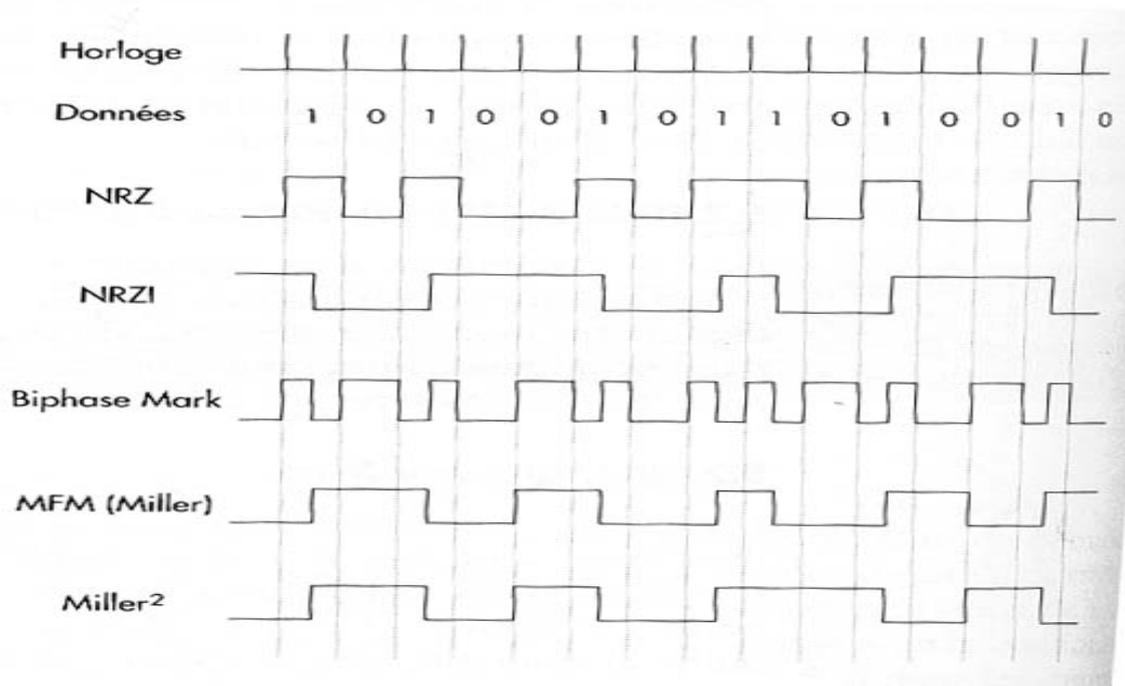
3) Le codage:

La valeur quantifiée est écrite en un nombre binaire 0 ou 1.

(A l'origine, les signaux quantifiés sous 8 bits ont fait apparaître des problèmes de luminance; on a donc étendu le système à 10 bits: 2^{10} états possibles).

Le codage de canal: Le codage de canal a pour but de moduler le flux de données numériques pour l'adapter aux caractéristiques du canal de transport ou d'enregistrement. Par exemple, les longues suites de 1 ou de 0 du message numérique créent des composantes continues, impossibles à relire une fois enregistrées. Il faut donc les rompre par un codage particulier dont l'unique but est de modifier la forme du signal, sans bien sûr toucher à son contenu. Il existe plusieurs codes, chacun ayant ses avantages et ses inconvénients qui les rendent plus appropriés à tel ou tel type d'application: réduction de la composante continue, mais aussi recouvrement de l'horloge, distribution spectrale, etc....

- *NRZ (non retour à zéro)*: c'est le plus simple de tous. Une donnée binaire "1" engendre un niveau haut du signal et une donnée binaire "0", un niveau bas. Ce code est caractérisé par une importante composante continue, ainsi que par une absence du signal d'horloge, qu'il faut donc régénérer à la réception.
- *S-NRZ (Scrambled NRZ= NRZ embrouillé)*: Il s'agit d'une variante du code NRZ dans laquelle le signal est mélangé avec une séquence binaire pseudo-aléatoire. Cela a pour effet d'une part de générer des transitions fréquentes en brisant les longues suites de symboles identiques, et d'autre part de réduire la valeur de la composante continue, ce qui est essentiel pour le transformateur du tambour de têtes du magnétoscope.
- *NRZI (non retour à zéro inversé)*: Un "1" détermine une transition au milieu de la demi-période d'horloge, un "0" n'a aucun effet. Ce code, qui présente l'avantage d'être insensible aux inversions de polarité, et caractérisé par une faible composante continue. Il est notamment utilisé dans les liaisons série 4 : 2 : 2.
- *Biphase Mark* : un "0" provoque une transition et un maintien du niveau pendant toute la période d'horloge, tandis qu'un "1" entraîne une transition et un changement de niveau à la moitié de la demi période d'horloge. Ce code présente une composante continue nulle et contient tous les fronts d'horloge; il est dit autosynchroniseur. Il est employé par le code temporel longitudinale LTC des magnétoscopes.



II. LA CAPTURE DES IMAGES.

1. Signaux analogiques et numériques :

Un signal vidéo ou audio analogique représente les variations de chaque élément d'une image ou d'un son. Un signal analogique qui est échantillonné, puis codé en éléments binaires, sous la forme de 0 et 1, devient un signal numérique. Les échantillons sont prélevés à une certaine cadence, appelée fréquence d'échantillonnage, et codés sur un nombre d'éléments binaires.

2. Définition et résolution :

Ces deux termes sont souvent confondus, alors qu'ils ne désignent pas la même chose. Prenons l'exemple d'un capteur CCD de caméscopes ; son rôle est de recueillir l'image formée par les rayons lumineux focalisés par l'objectif. Il contient un ensemble de cellules sensibles qui constituent autant de points d'analyse de l'image (les pixels) qui stockent des électrons proportionnellement au nombre de photons reçus (lumière). Le nombre de points contenus sur la cible d'analyse de l'image constitue la résolution du capteur CCD. Aujourd'hui, sur certains appareils, les capteurs CCD contiennent jusqu'à 800 000 points.

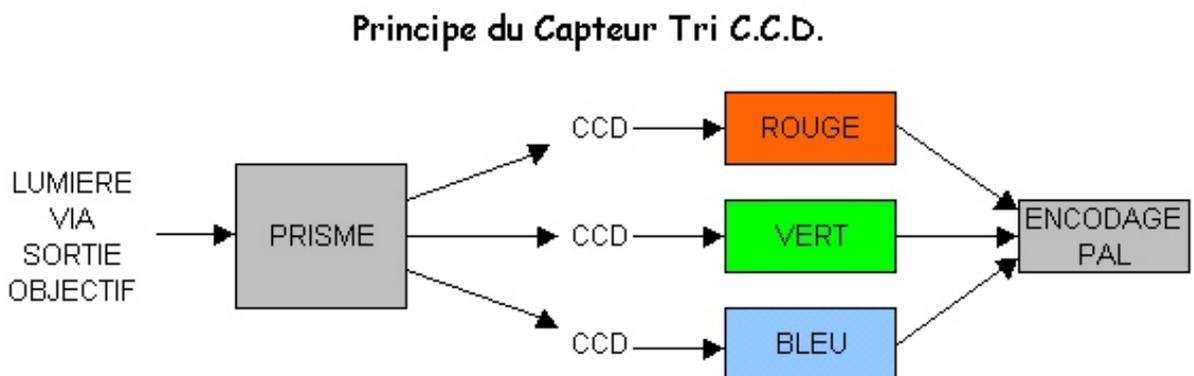
La définition dépend de la qualité de l'optique, du format d'enregistrement et de la résolution du moniteur. Elle dépend donc des résolutions de chaque maillon de la chaîne. En vidéo, cette définition est caractérisée par le nombre de points par ligne reproduit par l'équipement, souvent appelé « nombre de ligne » ; attention, cette dernière dénomination concerne des lignes verticales et ne doit pas être confondue avec le nombre de lignes horizontales qui dépend du standard utilisé (625 lignes en PAL/SECAM et 525 en NTSC).

3. Les capteurs Mono et Tri C.C.D. :

Une image formée par un capteur CCD est analysée pixel par pixel, en fonction de l'intensité lumineuse des trois couleurs primaires, rouge, vert et bleu. Pour extraire ces signaux, le capteur mono -CCD contient un filtre mosaïque qui recouvre chaque pixel le découpant ainsi en sous - pixel affecté à chacune des couleurs du filtre, de telle manière que l'on puisse obtenir les signaux rouges, vert et bleu. En revanche, sur un caméscopes tri - CCD, un capteur est dédié à chacune de ces trois couleurs primaires. Le capteur tri - CCD garantit une meilleure reproduction des couleurs, notamment au niveau des contours qui ne bavent plus, mais il ne triple pas la définition de l'image reproduite.

4. La définition limitée par le format vidéo :

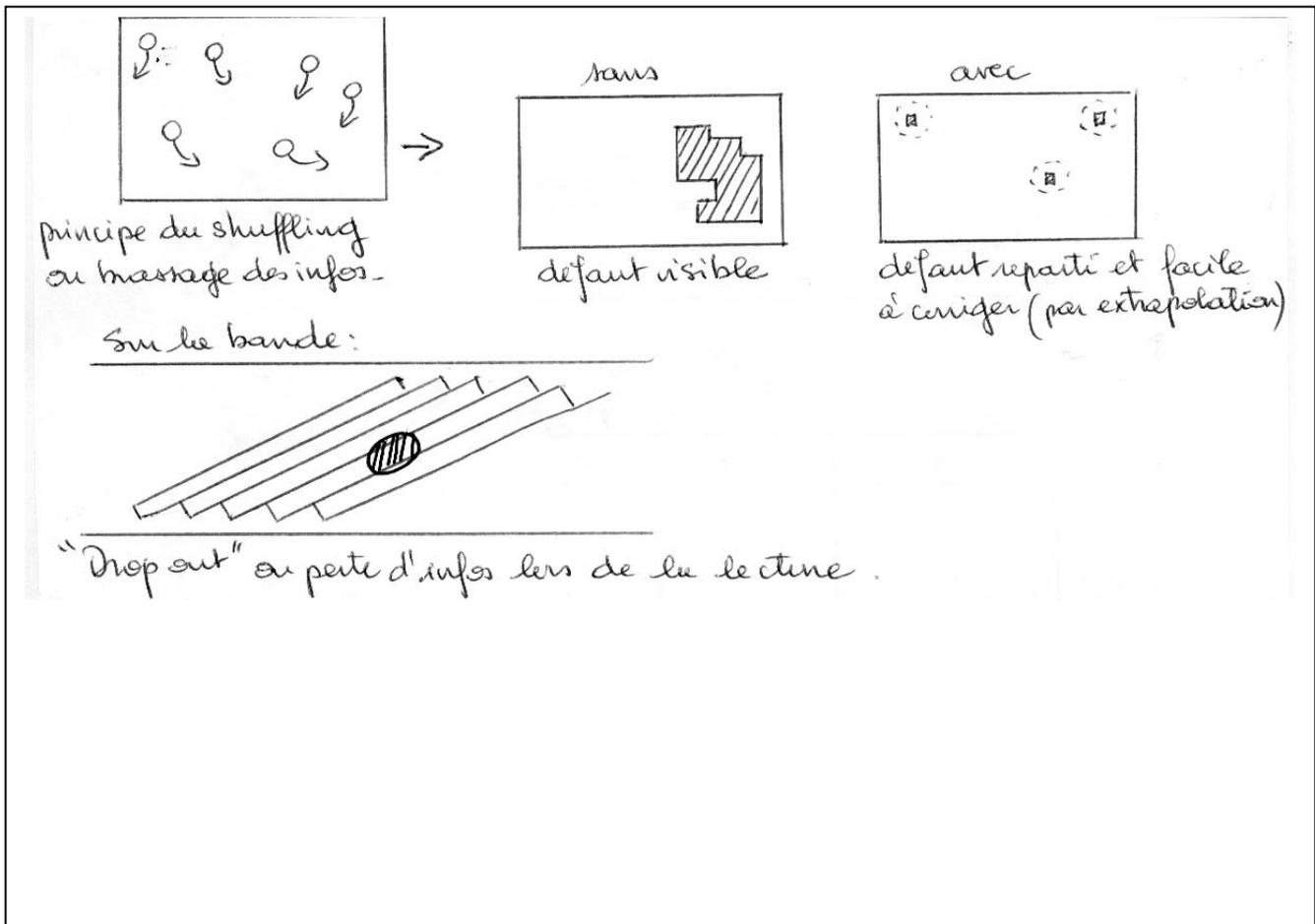
Intuitivement, nous comprenons que la finesse des images reproduites va croître avec le nombre de pixels du capteur ; il existe pourtant un seuil au - delà duquel ils n'apportent plus d'amélioration à la qualité de l'image reproduite. Ce seuil dépend de l'optique, de l'électronique et surtout du format vidéo utilisé qui fixe une limite au-delà de laquelle la définition ne peut plus être améliorée



III. EXPLICATIONS TECHNIQUES :

1. Le brassage des données et corrections d'erreur :

L'une des principales qualités des magnétoscopes numériques est l'absence de drop out ou parasites dans l'image. Ces parasites sont dus à un défaut lors de la lecture, poussière sur la tête, défaut de magnéto du support ou rayure de la bande magnétique. Dans les magnétoscopes numériques, les informations qui correspondent à des zones voisines de l'image sont enregistrées à des endroits différents pour fractionner et répartir dans l'image un éventuel défaut. C'est le brassage des données ou « shuffling ». D'autre part, des codes de correction d'erreurs permettent de retrouver des données perdues ou erronées. L'ensemble de ces techniques donne aux formats numériques une grande fiabilité lors de l'exploitation.

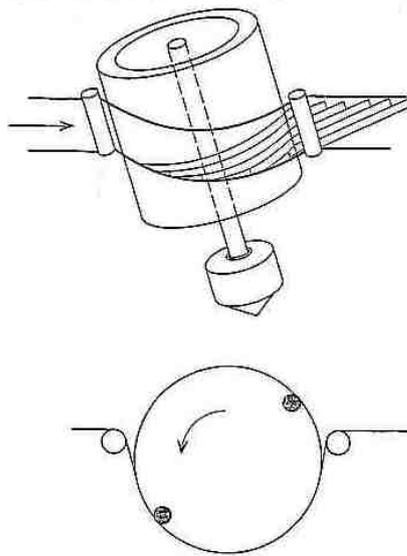


2. Spécificité du signal vidéo :

a) L'enregistrement hélicoïdal :

La solution découle en fait directement de la nature même du signal vidéo, qui délivre une description découpée en lignes et en trames de l'image. Les intervalles de suppression offrent en effet l'opportunité de fractionner l'enregistrement en plusieurs pistes, le passage de l'une à l'autre pouvant s'effectuer pendant les instants de suppression durant lesquels aucune information utile n'est présente. Ces pistes vidéo sont inscrites en diagonale par rapport à l'axe de la bande, par des têtes montées sur un tambour rotatif incliné. Ce dernier fournit l'essentiel de la vitesse relative tête/bande nécessaire. Le défilement de la bande peut ainsi s'effectuer à vitesse raisonnable, son rôle étant ici limité à décaler régulièrement les pistes.

Dans un tel système d'enregistrement hélicoïdal, la bande s'enroule en biais autour du tambour portant plusieurs têtes qui se relaient.



La rotation du tambour peut s'effectuer dans la même direction ou dans la direction opposée à celle du défilement de la bande, ce qui est le cas dans la plupart des machines broadcast. Si V_t est la vitesse de rotation du tambour, V_b la vitesse de défilement de la bande, D le diamètre du tambour, la vitesse relative tête/bande, ou vitesse d'écriture, V_r est :

$$V_r = (\pi * D * V_t) \pm V_b$$

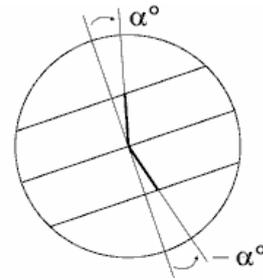
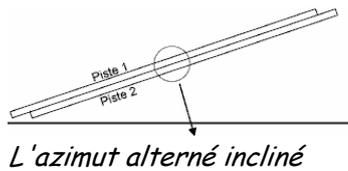
L'avantage d'un tambour de grand diamètre est qu'il peut accueillir un nombre relativement important de têtes, et par conséquent tourner moins vite. Cependant, les tambours de taille réduite sont les seuls à permettre la fabrication de caméscopes ultra compacts.

b) L'azimut :

Pour qu'un enregistrement soit reproduit dans des conditions optimales, il est primordial que l'angle que forme la tête de lecture avec l'axe de la piste soit exactement le même que celui de la tête d'enregistrement. Si cette condition n'est pas respectée, il se produit un affaiblissement rapide des hautes fréquences qui est fonction de l'angle d'azimut.

Sur un magnétoscope, l'entrefer des têtes doit toujours être strictement perpendiculaire à l'axe de la piste. Dans le cas d'un enregistrement multipistes, les pistes doivent être suffisamment espacées, afin d'éviter qu'une tête lise les informations d'une piste voisine. Soulignons cependant que les intervalles séparant les pistes - dits intervalles de garde - occupent un espace supplémentaire sur la bande, qui ne peut être utilisé pour enregistrer les signaux utiles.

Sur un magnétoscope, l'effet d'azimut est exploité efficacement pour enregistrer les pistes obliques sans intervalle de garde. Il suffit de tracer les pistes contiguës avec des azimuts opposés (azimut alterné incliné). Par exemple, si une piste est tracée avec un azimut de $+15^\circ$, les pistes qui l'entourent sont tracées avec un azimut de -15° . Ainsi, lorsqu'une tête de lecture empiète un peu sur une piste voisine, elle ne récupère que des signaux très atténués. Aucun espace n'est plus nécessaire pour séparer les pistes ; l'utilisation de la surface de la bande est ainsi optimale. Le choix de la valeur de l'angle d'azimut résulte d'un compromis : l'angle doit être suffisamment grand pour assurer une bonne protection contre l'intermodulation de piste à piste, mais pas trop grand, car la vitesse d'écriture - vitesse relative tête/bande - est multipliée par le cosinus de l'angle d'azimut.



c) Les pistes longitudinales :

Dans la plupart des formats vidéo, les bords supérieurs et inférieurs de la bande sont occupés par des pistes longitudinales utilisées pour enregistrer l'audio, le code temporel, ainsi que des impulsions de synchronisation. Ces dernières permettent d'asservir la phase de rotation du tambour avec la phase de transport de la bande pendant la lecture et l'enregistrement. Ainsi, les pistes sont balayées exactement de la même façon à la lecture et à l'enregistrement.

d) Les codes temporels LTC et VITC

Un code temporel (time code en anglais) est une adresse absolue permettant de repérer chaque image enregistrée sur la bande par sa position en heure, minute, seconde, et numéro d'image. Il existe deux types de codes temporels: le LTC (Longitudinal Time Code) et le VITC (Vertical Interval Time Code).

Le code LTC est enregistré par une tête fixe sur une piste longitudinale. Il est difficilement détecté lorsque la lecture de la bande s'effectue à une vitesse ralentie (le niveau du signal étant trop faible), et totalement illisible en mode arrêt sur image (la bande étant fixe devant une tête fixe). En revanche, il est parfaitement détecté en rembobinage très rapide.

Le code VITC est, quant à lui, enregistré par des têtes rotatives dans le signal vidéo, sur deux lignes situées dans l'intervalle de suppression trame. Sa lecture est par conséquent possible à des vitesses très lentes, et même sur une image fixe - la bande ne bouge pas, mais le tambour de têtes tourne. Par contre, il est illisible en rembobinage très rapide.

3. Les pertes et défauts de l'enregistrement

a) Effet d'éloignement

Si, pour une raison quelconque, un espace d est introduit entre l'entrefer et la bande, il se produit une atténuation du niveau de sortie, qui dépend à la fois de cet éloignement d et de la longueur d'onde λ du signal. Cette atténuation croît très rapidement pour les fréquences élevées. D'où l'importance de vérifier régulièrement que les têtes ne sont pas encrassées et d'éviter tout risque d'introduction de poussière dans la machine. L'affaiblissement en dB est donné par la formule:

$$A \text{ (dB)} = 55.d / \lambda$$

b) Effet d'azimut

Une erreur d'azimut se produit lorsque l'entrefer de la tête de lecture n'a pas exactement la même inclinaison que celui de la tête ayant réalisé l'enregistrement. Soit A l'angle de différence d'inclinaison. λ la longueur d'onde du signal et h la hauteur de la piste. L'atténuation en dB du niveau de lecture est donnée par la formule:

$$A \text{ (dB)} = 14.3 \times (h.A / \lambda)^2$$

Cet effet, comme celui d'éloignement, affecte davantage les fréquences élevées que les basses fréquences.

c) Les pertes de niveau, ou drop out

Malgré tout le soin pouvant être apporté à sa fabrication, une bande magnétique n'est pas à l'abri de Disparités locales dans la distribution des particules ferromagnétiques. Cela peut engendrer des chutes de niveau plus ou moins importantes sur le signal lu, se traduisant sur l'image par la perte d'une ou plusieurs lignes. Ce défaut est communément appelé drop out. Les magnétoscopes analogiques détectent ces pertes de niveau sur la bande et remplacent les informations manquantes par d'autres récupérations sur les lignes précédentes. Sur un magnétoscope numérique, les méthodes de correction d'erreurs parviennent à lutter dans une large mesure contre les absences fugitives de signal.

4. Différents types d'enregistrement

a) Les méthodes d'enregistrement vidéo analogiques

Les formats analogiques mettent en œuvre différentes méthodes pour enregistrer le signal vidéo. Certains enregistrent le signal composite directement en modulation de fréquence; d'autres transposent la chrominance dans le bas du spectre pour s'adapter à une largeur de bande réduite; d'autres encore effectuent un traitement séparé sur les signaux de luminance (Y) et de chrominance (C) avant de les combiner à l'enregistrement; d'autres enfin enregistrent sur des pistes différentes les composantes Y et Dr/Db.

- *L'enregistrement direct du signal composite (2", 1"B, 1"C)*

Les premiers magnétoscopes à bandes broadcast enregistraient directement le signal Composite (bande passante 5.5 MHz) en utilisant la modulation de fréquence. Pour chaque format, un standard de modulation était défini par une fréquence f1 correspondant au fond de synchro du signal vidéo (0 V), et une fréquence f2 pour le blanc (1 V). L'excursion en fréquence f2 - f1 était d'environ 2 MHz. Une telle méthode d'enregistrement n'a été rendue possible qu'au prix d'une vitesse relative tête/bande très élevée - de l'ordre de 20 m/s. Ces machines très encombrantes ont aujourd'hui presque totalement disparu au profit de magnétoscopes à cassettes bien plus compacts et plus pratiques. Le principe de la modulation de fréquence a cependant été conservé par tous les formats analogiques qui ont suivi.

- *Le procédé " Under Color " (U-Matic, VHS, 8 mm)*

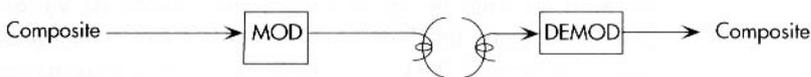
Ce procédé a été inventé à l'origine pour l'U-matic 3/4", le format à cassettes le plus répandu dans les années 1970 sur les marchés institutionnels et broadcast. Il a été repris ensuite par les formats grand public VHS et 8 mm. La technique "under color" est en fait nécessaire quand la vitesse relative tête/bande n'est pas assez élevée pour enregistrer l'intégralité de la bande passante du signal vidéo sans dégrader ses hautes fréquences, c'est-à-dire la zone où se trouve notamment la chrominance. La technique utilisée consiste à transposer le signal de chrominance - 4.43 MHz en PAL - dans le domaine des basses fréquences, en dessous de 1 MHz. Cette opération n'est évidemment pas transparente: elle engendre une réduction des bandes passantes de la luminance - pas plus de 3 MHz - et de la chrominance. A l'enregistrement, la luminance modulée en fréquence est additionnée à la chrominance transposée. A la lecture, un filtrage sépare les deux signaux et la chrominance retrouve sa place à 4.43 MHz. Après démodulation, la luminance est ajoutée à la chrominance pour former le signal vidéo composite.

- *Le traitement séparé Y/C (S-VHS, Hi8)*

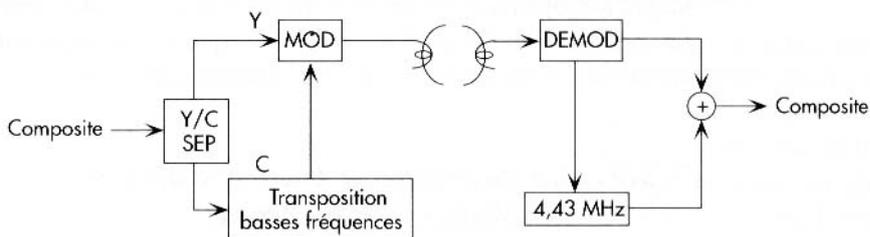
Le signal vidéo Y/C se situe à mi-chemin entre le signal composite et le signal en composantes pures (dérivées de R, V, B).

Le signal Y/C est caractérisé par un traitement séparé des signaux de luminance et de chrominance. La coupure classique du signal composite "passe-bas" pour Y et "cloche" pour C est ainsi évitée. La bande passante de la luminance est pour sa part élargie, ce qui améliore de manière significative la résolution de l'image. Il faut cependant préciser que les signaux Y et C sont au final recombinaés pour être enregistrés par la même tête; ils n'occupent donc pas des pistes séparées comme dans les systèmes en composantes. Sur les machines travaillant en mode Y/C, les signaux de luminance et de chrominance sont cependant disponibles séparément en entrée/sortie sur des connecteurs appelés S-vidéo.

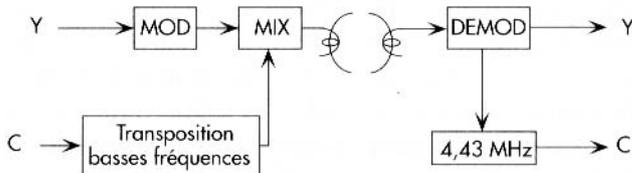
Composite direct en modulation de fréquence (2", 1")



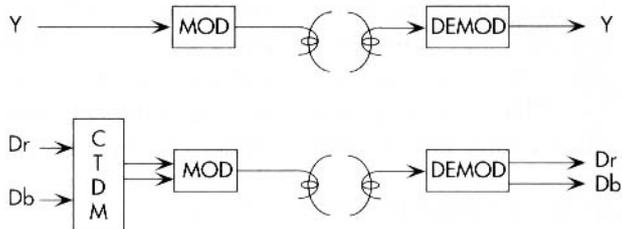
Under color (U-Matic, VHS, 8mm)



Y/C (S-VHS, Hi8)



Composantes (Betacam, Betacam SP)



- **L'enregistrement en composantes (Betacam, Betacam SP)**

Fin du fin dans le domaine de l'enregistrement du signal vidéo analogique, les formats en composantes enregistrent séparément les signaux Y et Dr/Db. Des têtes différentes sont en effet allouées au signal de luminance et aux signaux de différence de couleurs. Ces têtes tracent ainsi séparément et simultanément pour chaque trame une piste pour la luminance et une piste pour les signaux Dr/Db.

b) Les méthodes d'enregistrement vidéo numériques

On distingue aujourd'hui trois grandes catégories de formats d'enregistrement vidéo numériques:

- *Les formats en composantes numériques (D1, D5)*

Ils enregistrent le signal en composantes 4:2:2 à plein débit, avec une quantification sur 8 ou 10 bits. Ils sont totalement compatibles 525/625 lignes et sont caractérisés par un excellent rapport S/B de l'image, ainsi que par une totale transparence en multigénération (plus de 100 sans dégradation). Les puissants systèmes de correction d'erreurs corrigent de manière très efficace les pertes d'informations sur la bande, si bien que le signal relu est exempt de drop out.

- *Les formats composites numériques (D2, D3)*

Ils enregistrent le signal composite PAL ou NTSC numérisé à 4 fois la fréquence de la sous-porteuse de la chrominance du système analogique, avec une quantification sur 8 bits. Comme pour les formats en composantes numériques, le rapport S/B est très élevé et les problèmes de drop out sont inexistantes. En revanche, la notion d'universalité apportée par les formats en composantes disparaît, puisque les machines PAL et NTSC sont incompatibles entre elles. De plus, les limitations inhérentes au signal composite sont conservées.

- *Les formats compressés (Digital Betacam, famille DV, Betacam SX, D9, MPEG-IMX...)*

Ils enregistrent, selon le cas, un signal en composantes 4:2:2, 4:2:0 ou 4:1:1, numérisé sur 8 ou 10 bits et soumis à une réduction de débit de type M-JPEG ou MPEG-2. Ces nombreux formats ont permis une véritable démocratisation de l'enregistrement numérique, aussi bien dans le domaine broadcast qu'auprès du grand public.

IV. LA COMPRESSION NUMERIQUE :

1. Pourquoi une compression ?

Ceci n'est pas une idée nouvelle. Depuis les débuts de la télévision, on a toujours utilisé les limites de perception de l'œil humain pour réduire le nombre d'information à transmettre (réduction du nombre d'images par seconde, du nombre de lignes, des informations de chrominances inutiles dans les détails fins où l'œil n'est pas capable de les discerner...)

La numérisation des images est apparue en 1980 et la norme 4:2:2 produit des fichiers énormes, donc à des débits extrêmement élevés compte tenu des supports de stockage et de transmissions classiques. En effet 1 seconde d'images (en ne tenant compte que de la partie utile) en 4:2:2 quantifiée sur 8 bits occupe 21 Mo (Méga Octets) soit 15 disquettes 3,5" (ce qui fait qu'un CD-ROM ne peut contenir que 30s d'images). On peut donc comprendre la nécessité de la compression des signaux numériques. En plus de ce côté d'économie d'espace, la compression des images a ouvert de nouvelles fonctionnalités comme l'accès aléatoire à n'importe quelle image d'un programme, l'accès aux données par plusieurs utilisateurs au même moment, le transfert plus rapide qu'en temps réel.

2. Comment fonctionne la compression numérique ?

La compression numérique consiste en la suppression de certaines informations de l'image, d'en simplifier d'autres tout en s'assurant que les modifications ne seraient pas ou peu perceptibles par l'œil. On peut déjà remarquer que dans une image il y a des plages qui sont plus au moins uniformes (plusieurs pixels consécutifs identiques), on a donc des données redondantes qu'il n'est pas nécessaire de transmettre intégralement, on ne transmettra qu'une de ces données et on déduira les autres de celle-ci. Dans le cas d'une compression vidéo animée, on pourra exploiter le fait qu'on retrouve une grande similitude entre plusieurs images successives. Là aussi on peut économiser des données. L'œil humain est également beaucoup moins sensible aux détails fins qu'aux plages uniformes, ainsi qu'à la couleur par rapport à la luminosité. Ce sont donc beaucoup d'informations qu'on peut réduire ou éliminer pour compresser le signal vidéo.

Selon les compressions qui seront appliquées au signal, on distinguera les algorithmes de compression « lossless » (sans perte) qui effectuent un traitement totalement transparent, permettant de retrouver sans aucune perte les données d'origines après décompression. Mais ce type de traitement n'apporte qu'un faible taux de compression.

On distinguera également les algorithmes de compression « lossy » (avec pertes) où les taux de compressions sont nettement plus importants mais qui s'accompagne de la perte d'informations (négligeables par rapport à la perception de l'œil humain). Si ces pertes passent inaperçues, on parle alors de compression virtuellement transparente.

En cas de compression trop poussée, on distinguera une perte de définition, une saccade dans les mouvements, un flou de certaines parties de l'image...

3. Les redondances de l'image vidéo.

On distingue quatre types de redondances :

a) Redondance spatiale :

Une plage uniforme sur une image est constituée de pixels consécutifs identiques. Il suffit alors de ne transmettre que deux données, une pour la valeur du pixel et une autre pour le nombre de pixels. Cette redondance est mise en évidence par la Transformée en Cosinus Discrète (DCT).

b) Redondance temporelle :

Dans une séquence vidéo, des images successives sont souvent très semblables. Au moyen de techniques d'estimation de mouvement, on peut coder une image par rapport à celle qui la précède ou qui la suit en ne transmettant que les données relatives au déplacement des zones identiques. Cette élimination de redondances temporelles apporte de grands taux de compression. Mais comme le codage s'applique alors à un groupe d'images et non plus image par image, ce type de compression reste approprié à la diffusion mais pas au montage, surtout si les groupes d'images sont longs.

c) Redondance subjective :

Elle tire profit des faiblesses de la vision humaine. Elle consiste à coder avec un nombre de bits réduit les éléments de l'image jugés les moins significatifs. Cette opération entraîne des pertes de résolutions irréversibles.

d) Redondance statistique :

Cette redondance s'applique aux codes eux-mêmes. Certains sont plus employés que d'autres, on leur réservera alors les mots les plus courts Cette opération est dite codage entropique et n'entraîne aucune perte.

4. Quelques chiffres :

Dans une image fixe on peut supprimer 70% (taux 3,3 :1) de ses informations sans la dégrader, on dit alors que ce taux est totalement transparent.

En acceptant quelques pertes peu perceptibles, le taux de compression peut atteindre 10 :1 (90% des infos sont éliminées).

En diffusion les taux de compressions sont compris entre 15 :1 à 40 :1 (soit 93 à 97,5% des infos éliminées). Par contre les exigences en studio imposent une limite à 5 :1. Quant aux applications multimédias, elles se contentent d'une qualité inférieure au VHS où plus de 99% des informations sont supprimées.

Avant compression, le débit initial du signal vidéo est déjà nettement allégé. En ne prenant compte que la partie utile de l'image, un débit de 270 Mbits/s passe à 207 Mbits/s et il suffit de passer de 10 à 8 bits pour encore abaisser ce taux à 166 Mbits/s avant compression.

Ces chiffres sont donnés pour un signal en 4 :2 :2, c'est à dire 720 points de luminance et 360 pour chaque chrominance. Mais il suffit encore de passer en 4 :1 :1 ou en 4 :2 :0 pour que le taux s'abaisse à 124 Mbits/s.

Après compression on obtient :

En 4 :2 :0, 8 bits on a un taux de 3 à 10 Mbits/s (diffusion)

En 4 :2 :0 ou 4 :1 :1 en 8 bits un taux de 25 Mbits/s (DV)

En 4 :2 :2, 8 ou 10 bits un taux de 50 Mbits/s (postproduction)

5. Les normes de compressions :

4:2:2	720 576 Y	+ 576 Cr	+ 576 Cb		x 10	x 25		= 207 Mbits/s
	(10 bits)	(720 x 576	+ 360 x 576	+ 360 x 576)	x	bits/pixel	images/secon	
		pixels x lignes	pixels x lignes	pixels x lignes				
4:2:2	720 576 Y	+ 576 Cr	+ 576 Cb		x 8	x 25		= 166 Mbits/s
	(8 bits)	(720 x 576	+ 360 x 576	+ 360 x 576)	x	bits/pixel	images/secon	
		pixels x lignes	pixels x lignes	pixels x lignes				
4:2:0	720 576 Y	+ 288 Cr	+ 288 Cb		x 8	x 25		= 124 Mbits/s
	(720 x 576	+ 360 x 288	+ 360 x 288)	x	bits/pixel	images/secon		
		pixels x lignes	pixels x lignes	pixels x lignes				
4:1:1	720 576 Y	+ 576 Cr	+ 576 Cb		x 8	x 25		= 124 Mbits/s
	(720 x 576	+ 180 x 576	+ 180 x 576)	x	bits/pixel	images/secon		
		pixels x lignes	pixels x lignes	pixels x lignes				

Débits avant compression en ne tenant compte que de la partie utile de l'image (576 lignes sur les 625)

Remarque: Y: signal de luminance; Cr: signal de chrominance rouge; Cb: signal de chrominance bleu.

L'équation de luminance établit après études est: $Y=0.3R +0.59V +0.11B$.

On observe que la couleur verte vaut pratiquement 60% de la luminance. Il n'est pas nécessaire de récupérer ce signal puisqu'on récupère déjà la luminance. De plus, le signal du vert est facilement retrouvé grâce à l'équation: $V= 1/0.59(Y-0.3R -0.11B)$.

La norme la plus utilisée est le 4:2:2 car en plus de respecter le théorème de Shannon, elle est compatible aux standards mondiaux: PAL 625 lignes et NTSC 525 lignes.

a) JPEG (Joint Photographic Experts Group) (principe de codage intra-image) :

Apparu en 1989, JPEG est une norme de compression des images fixes. Cette norme a été développée à la base en ce qui concerne l'informatique. Le JPEG exploite les redondances spatiales.

Comme les circuits de codage JPEG étaient déjà disponibles, les fabricants d'équipements vidéo ont utilisé ce standard en partant du principe qu'une séquence vidéo est une succession d'images fixes. Mais chaque fabricant a développé son propre format dérivant du JPEG (appelé M-JPEG pour Motion JPEG). Les constructeurs ne s'étant pas mis en accord pour ce format et la synchronisation du son, les fichiers générés par les équipements de marques différentes sont très souvent incompatibles entre eux. Ces solutions sont par conséquent vouées à disparaître au profit de systèmes normalisés. Le codage du JPEG et du M-JPEG est intra-image.

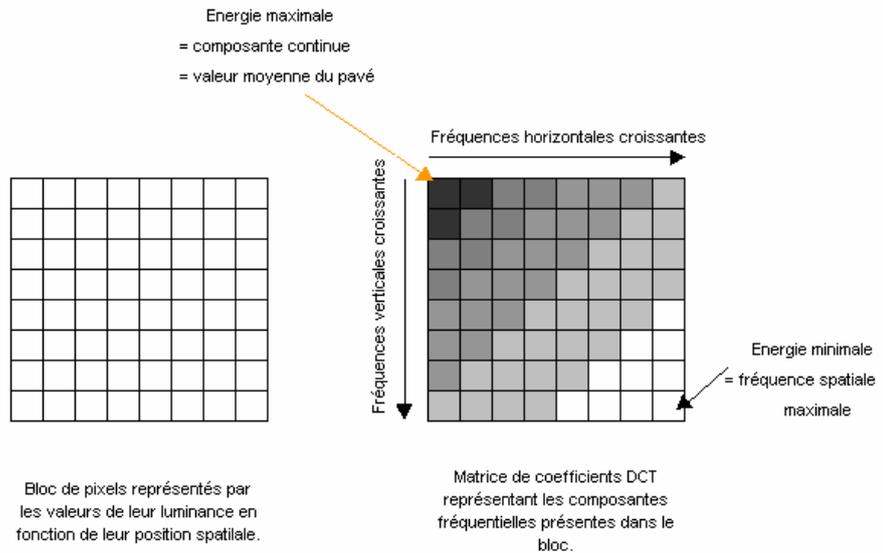
L'algorithme du JPEG comporte six étapes :

➤ **N°1 : Décomposition en blocs :**

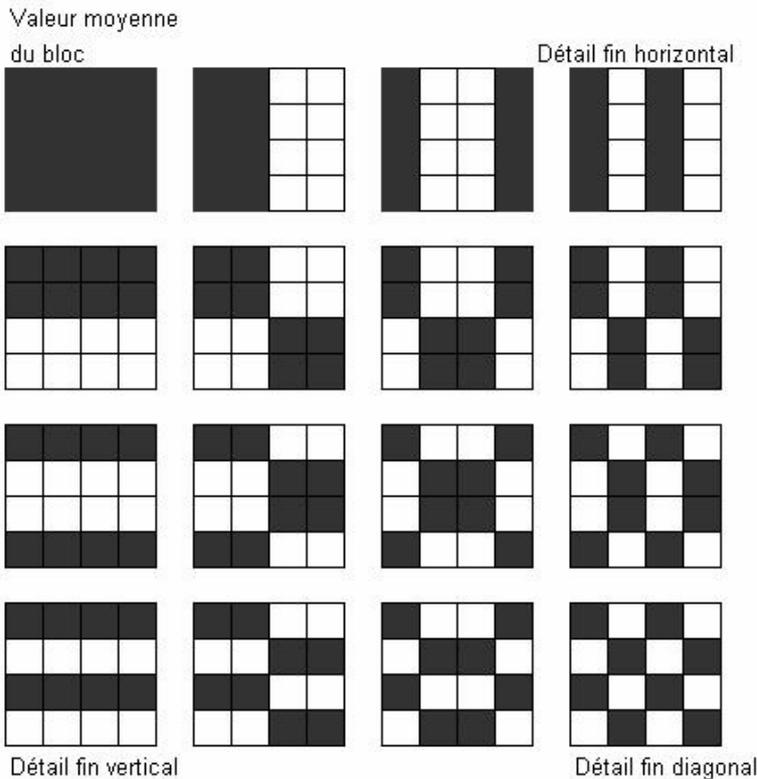
La DCT n'est pas effectuée d'un trait sur l'ensemble de l'image, cela impliquerait une quantité énorme d'informations et un temps de calcul important. Il faut donc découper l'image en blocs, en générale de 8x8 pixels.

➤ **N°2 : La Transformée en Cosinus Discrète (DCT) :**

La DCT a pour rôle de convertir les valeurs de luminances et de chrominance des pixels en coefficients fréquentiels. Une image ne contient que très peu de détails fins, par conséquent il y a très peu d'informations dans le haut du spectre, la quasi-totalité des informations se trouvent dans le bas du spectre. La DCT permet donc de regrouper les informations essentielles et de les traiter avec plus de précision.



La première case en haut à gauche contient le coefficient le plus élevé représentant la valeur moyenne du bloc. Plus on s'éloigne de cette case en se dirigeant vers celle située en bas à droite plus les motifs rencontrés sont fins. Cette transformée purement mathématique préserve l'intégralité de l'information du bloc et n'induit aucune perte.

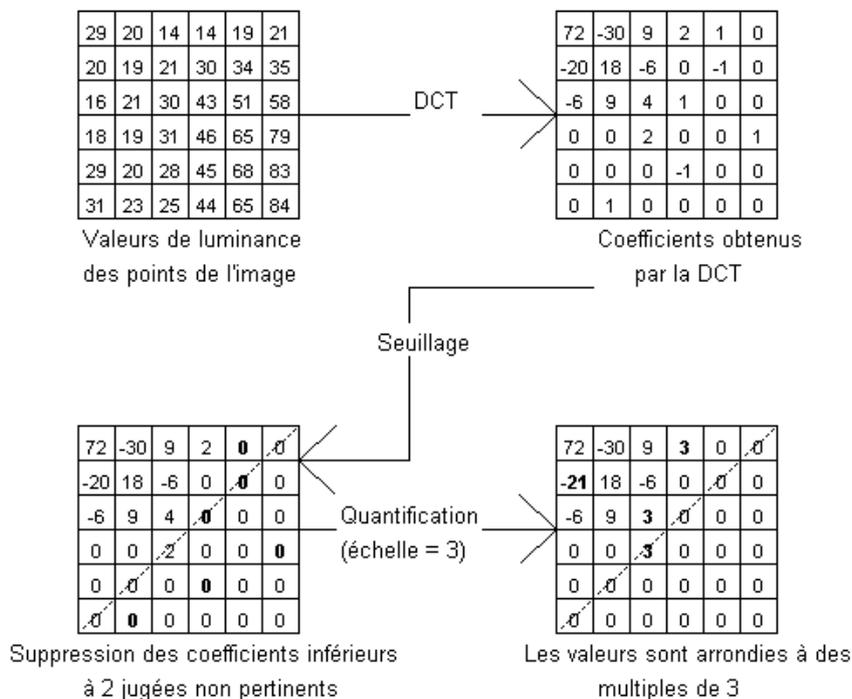


Voici l'exemple d'une bibliothèque de motifs que l'on peut obtenir avec un bloc 4x4 pixels. Les fréquences spatiales horizontales augmentent de gauche à droite, les fréquences spatiales verticales de haut en bas.

➤ **N°3 : Quantification sélective des coefficients de la DCT :**

Du fait que l'œil humain accorde moins d'importance aux détails fins qu'aux plages uniformes, on pourra coder les hautes fréquences (détails fins) avec moins de précision que les basses fréquences. Les valeurs situées dans les hautes fréquences vont être sous quantifiées (arrondies) et codées sur peu de bits. De plus certains coefficients inférieurs à un certain seuil seront même éliminés (fonction de seuillage).

C'est dans cette étape de quantification que les dégradations de codage JPEG interviennent.

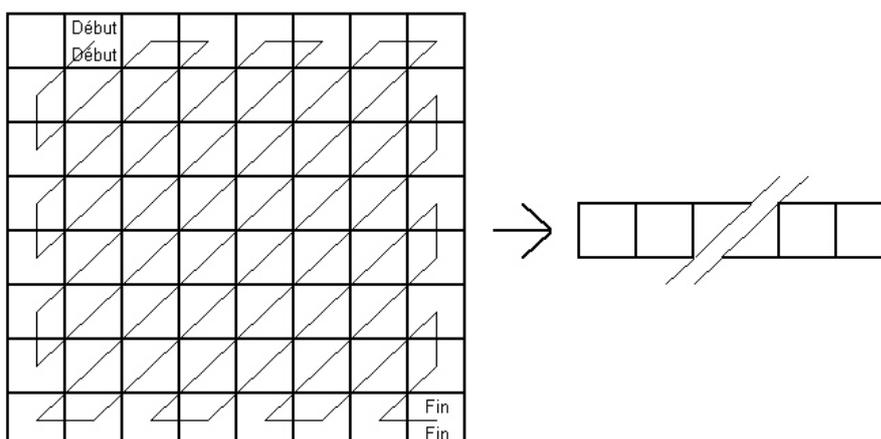


Exemple de codage d'un bloc 6x6 par l'algorithme JPEG

Au cours de l'étape de quantification, les coefficients DCT associés aux composantes fréquentielles élevées (détails fins) sont arrondis et codés avec un nombre de bits réduit au minimum exigé par l'œil. La quantification est la seule étape de l'algorithme JPEG à générer des pertes.

➤ **N°4 : Balayage en zigzag :**

La matrice obtenue après quantification est balayée en zigzag pour obtenir une forme plus commode à transporter. (La valeur moyenne n'est pas transportée avec les autres valeurs)



➤ **N°5 : Codage RLC (Run Length Coding) :**

Au lieu de coder systématiquement toutes les valeurs, on code seulement deux valeurs qui correspondent à la valeur du coefficient et au nombre de fois que celle-ci se répète. Ceci est très efficace, surtout lorsque les coefficients valent zéro.

Le codage RLC est un codage lossless (sans perte).

➤ **N°6 : Codage entropique ou VLC (Variable Length Coding) ou encore d'Huffman :**

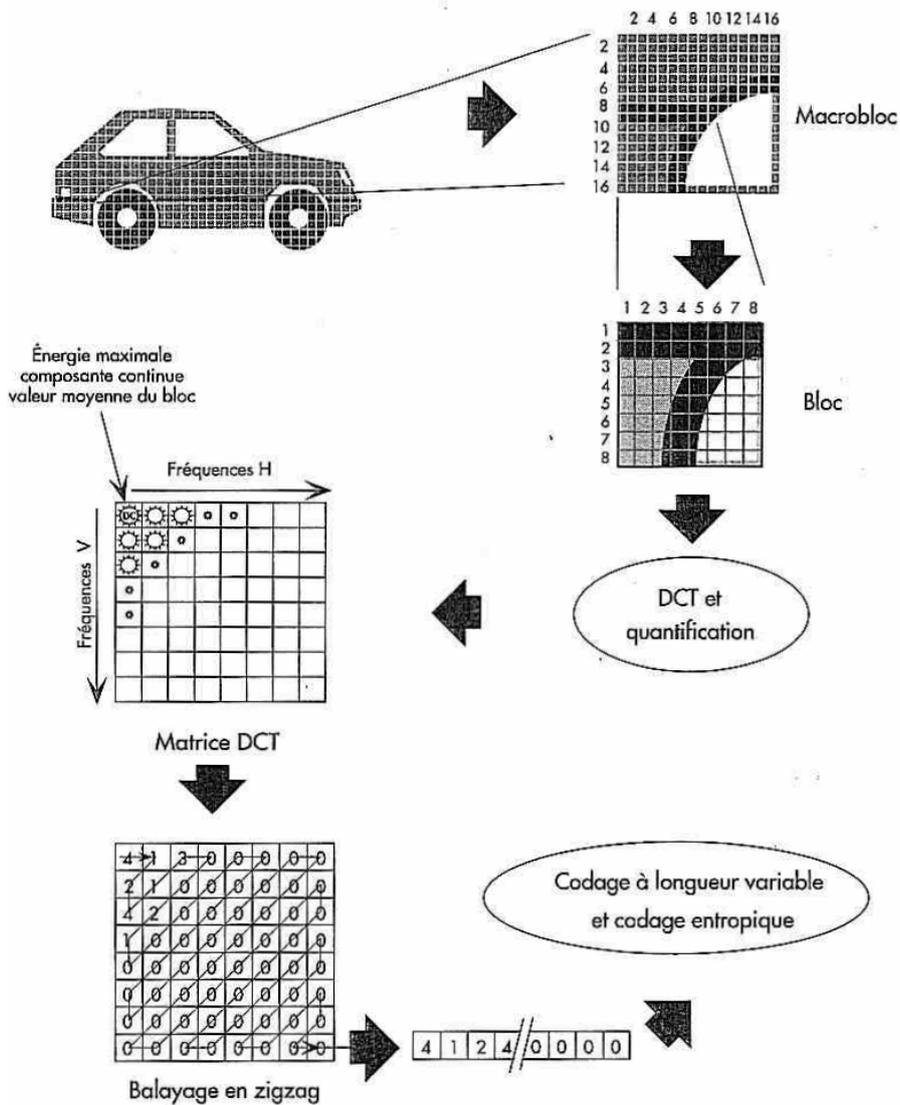
On attribue les codes les plus courts aux coefficients qui sont les plus fréquents et les codes longs à ceux qui sont les plus rares (principe du morse ou peu de signes sont utilisés pour les lettres les plus fréquentes comme le E, et plus pour les moins fréquentes comme le Q).

Le codage entropique est un codage lossless (sans perte).

Ces deux codages (RLC et entropique) ne permettent qu'un taux de compression de 2 :1.

Synoptique de compression JPEG :

Voici les différentes étapes de l'algorithmes de compression. La décompression s'effectue de manière réciproque en parcourant le dessin en chemin inverse.



b) MPEG-1 (Motion Pictures Experts Group).

C'est une norme de compression des images animées à faible résolution, destinée aux applications multimédia. Le MPEG-1 reprend l'élimination des redondances spatiales (comme le JPEG), et exploite aussi les redondances temporelles entre les images (le codage est dit inter image). Son débit est de 1,5 Mbits/s et il inclut la vidéo et le son. Une des applications du MPEG-1 est le Cd Vidéo, il est aussi devenu le format de stockage de la vidéo sur PC.

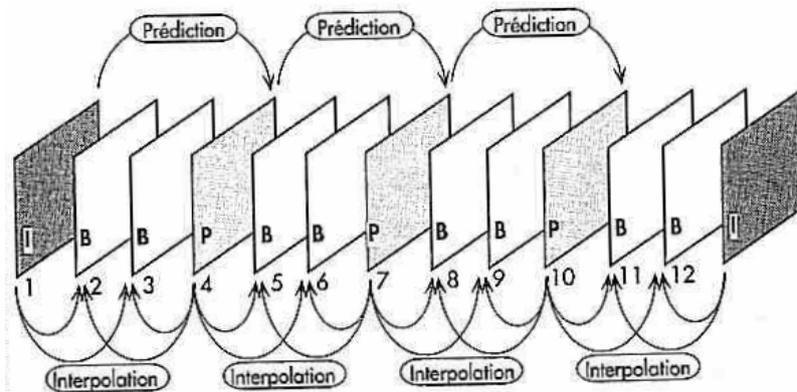
Le MPEG-1 reprend les techniques JPEG afin d'éliminer les redondances internes de chaque image, mais il fait également appel à d'autres procédés comme les images SIF.

Le format source SIF (Source Intermediate Format) se caractérise par une réduction de moitié des résolutions spatiales et temporelle. Il supprime une trame sur deux de l'image (elle sera dédoublée lors du décodage). Il élimine également un point sur deux sur chaque ligne, en luminance comme en chrominance. Ce qui nous amène à un format d'image de 360 pixels par 288 lignes (qualité VHS).

Contrairement au système M-JPEG qui traite les images indépendamment, les systèmes MPEG (1 et 2) effectuent leur traitement sur des groupes d'images appelés GOP (Groupe Of Pictures) ou codage inter image.

Les GOP se composent d'une combinaison de trois types d'images :

- L'image I (Intra) : elle est codée avec les algorithmes de JPEG en exploitant les redondances spatiale intra-image. Elle est entièrement décrite par elle-même et ne fait appel à aucune autre image. Elle ne bénéficie que d'un faible taux de compression et reste donc assez volumineuse.
- L'image P (Prédite) : elle est prédite à partir d'une image passée I ou P. Elle est codée uniquement au moyen de vecteurs mouvement indiquant les déplacements de ses éléments par rapport à l'image référence. Une image P occupe trois fois moins de place qu'une image I, mais peut transmettre des erreurs car sert aussi de référence.
- L'image B (Bidirectionnelle) : elle est construite à partir de vecteur mouvement estimé entre les images passées ou futures I ou P voisines. Elle offre le taux de compression le plus élevé, mais ne propage pas d'erreur car n'est jamais utilisée en référence. Elle occupe six fois moins de place qu'une image I.



Exemple de GOP

Un GOP commence toujours par une image I et se termine à la dernière image avant la prochaine image I. Il peut se résumer à une seule image I ou à une combinaison d'images I et P, ou I et B, ou encore I, P et B.

On ne peut briser un GOP. Comme seul l'accès aux images I est autorisée, il devient alors difficile de réaliser un montage de précision avec les GOP.

En revanche un GOP long présente l'avantage de permettre un taux de compression plus élevé qu'un GOP court pour une même qualité d'image. Autrement dit un GOP long donne une meilleure qualité d'image qu'un GOP court, à débit égal.

L'estimation de mouvement :

Le principe de l'estimation de mouvement consiste à construire une image P à partir d'une image précédente et d'informations relatives aux déplacements de ses composantes. L'estimation de mouvement n'est pas effectuée sur chaque point de l'image, ni même sur des blocs de 8x8 mais sur des macroblocs, afin de réduire le nombre d'informations à transmettre.

Si entre l'image actuelle et celle qui la précède deux macroblocs semblables sont trouvés, mais à des emplacements légèrement différents, il suffit de transmettre une seule fois ce macrobloc ainsi que sa nouvelle position au moyen d'un vecteur mouvement.

L'estimation de mouvement se décompose en 5 étapes :

- Recherche des macroblocs semblables entre une image i et l'image précédente $i-1$
 - Calcul des vecteurs mouvements caractérisant les déplacements des macroblocs
 - Construction d'une image prédite en utilisant ces vecteurs mouvements
 - Comparaison de cette image prédite avec la vraie image pour générer des données d'erreurs
 - Codage et transmission des vecteurs et des données d'erreur de prédiction
- Ce principe peut apparaître simple, c'est pourtant la phase la plus complexe du codage MPEG.

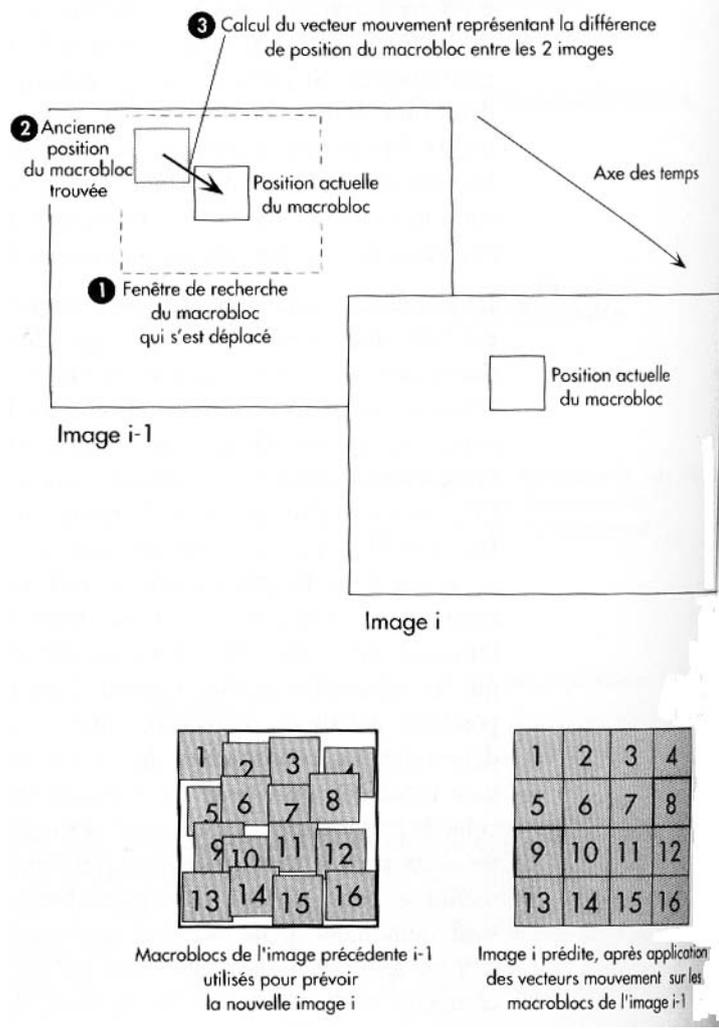
Le "block matching":

C'est la technique la plus répandue pour former le vecteur mouvement que l'on traduit par correspondance des blocs.

L'estimateur de mouvement compare l'image d'entrée, que nous considérerons comme la nouvelle image " i ", avec l'image précédente " $i-1$ ", conservée en mémoire. Cette comparaison consiste à examiner un à un les macroblocs de l'image " i ", afin de voir s'ils existaient sur l'image " $i-1$ ". La figure suivante illustre ce principe avec un macrobloc sur l'image " i ", que l'on cherche à localiser sur l'image " $i-1$ ". Une exploration est alors réalisée dans toutes les directions possibles à l'intérieur d'une fenêtre de recherche, afin d'identifier le macrobloc qui lui ressemble le plus.

Lorsqu'il est repéré la différence de position spatiale du macrobloc entre les deux images permet de déterminer les coordonnées du vecteur mouvement. Son amplitude représente la vitesse du déplacement ; sa direction indique celle de la translation. Les macroblocs occupant la même place sur les deux images sont ignorés, ce qui diminue la quantité d'information à coder.

Toutes les correspondances étroites de macroblocs sont combinées pour générer une image prédite, exactement comme le ferait le décodeur. Sauf qu'ici, cette image prédite est comparée avec la "vraie" image " i ", afin de produire des données de différence visant à compenser les erreurs ou imprécisions de l'estimateur de mouvement. Les vecteurs mouvements et les données de différence sont transmis avec l'image " $i-1$ ". Ils signaleront au décodeur comment il devra déplacer les macroblocs de cette image " $i-1$ " pour construire l'image " i ".



c) MPEG-2

Créée en 1994, il reprend les techniques de bases du MPEG-1. En revanche la qualité d'image est nettement supérieure puisqu'elle s'étend de la vidéo standard (3 à 10 Mbits/s à la haute définition 300 Mbits/s). La norme MPEG-2 est utilisée par tous les opérateurs de télévisions numériques et constitue la norme des DVD. Comme la norme MPEG-1, MPEG-2 a un codage inter image.

La norme MPEG-2 a été conçue pour couvrir toutes les applications de distribution d'images et de sons par satellite, câble et voie terrestre, ainsi que pour des support tels que le DVD. Puis cette norme a été développée pour des applications broadcast en production et post production, montage, incrustation...

Par rapport à la norme MPEG-1, MPEG-2 à pour différences :

- Support de formats d'image de qualité supérieur (jusqu'à la HD)
- Traitement de balayage entrelacé (50 trames/s) et du balayage progressif (50 images complètes/s)
- Compatibilité avec le MPEG-1 (un décodeur MPEG-2 peut décoder un signal MPEG-1)

Cependant la norme MPEG-2 se contente de décrire les outils de compression des données ainsi que la syntaxe du signal. La technique utilisée dans le codeur est établie par le constructeur et peut constamment évoluer. La qualité de l'image dépend donc aussi de la qualité du codeur et du décodeur.

Le MPEG-2 a été conçue comme une boîte à outils, organisée en profils et en niveaux. Un profil définit le jeu d'outils de compression utilisés et un niveau fixe le nombre de lignes et de pixels qu'il est possible de coder.

Profil \ Niveau	Simple	Main	4:2:2	SNR	Spatial	High
High		4:2:0 1920 x 1152 80 Mbits/s I, P, B	4:2:0, 4:2:2 1920 x 1152 300 Mbits/s I, P, B			4:2:0, 4:2:2 1920 x 1152 100 Mbits/s I, P, B
High-1440		4:2:0 1440 x 1152 60 Mbits/s I, P, B			4:2:0 1440 x 1152 60 Mbits/s I, P, B	4:2:0, 4:2:2 1440 x 1152 80 Mbits/s I, P, B
Main	4:2:0 720 x 576 15 Mbits/s I, P	4:2:0 720 x 576 15 Mbits/s I, P, B	4:2:0, 4:2:2 720 x 576 50 Mbits/s I, P, B	4:2:0 720 x 576 15 Mbits/s I, P, B		4:2:0, 4:2:2 720 x 576 20 Mbits/s I, P, B
Low		4:2:0 352 x 288 4 Mbits/s I, P, B		4:2:0 352 x 288 4 Mbits/s I, P, B		

Les profils et niveaux de la famille MPEG-2.

Dans chaque case sont indiqués respectivement le format d'entrée accepté, la résolution pixels x lignes, le débit maximal supporté et le type d'images traitées.

Deux standards MPEG-2 couvrent l'ensemble des applications courantes, le MPEG-2 MP@ML et le MPEG-2 422P@ML.

d) MPEG-2 MP@ML (Main Profile ou profil principal et Main Level ou niveau principal) :

Le MP (Main Profile) est le profil de base utilisé pour la diffusion numérique. Il traite l'image en 4 : 2 : 0 et utilise des images I, P et B.

Le ML (Main Level) est un niveau de base conforme au format d'image 4 : 2 : 2.

Ce standard est à la base de tous les bouquets de programmes numériques et a des débits s'échelonnant entre 3 et 10 Mbits/s.

e) MPEG-2 422P@ML (4 : 2 : 2 Profile @ Main Level):

Le Profil 4:2:2 étant une version améliorée du Profil principal, commençons par voir pour quelles raisons ce dernier ne pouvait convenir aux applications du studio.

Tout d'abord, le Profil principal ne traite que le signal 4:2:0, amputé de la moitié de ses échantillons de chrominance en vertical. Or, il est indéniable que la majorité des travaux de production et de post-production exigent que soit maintenue la résolution 4:2:2 de l'image. D'autre part, le débit maximal autorisé par le Profil principal, soit 15 Mbits/s, s'avère trop faible. Des tests d'évaluation effectués par le comité MPEG ont, certes, montré qu'une qualité d'image comparable au 4:2:2 peut être observée avec seulement 10 Mbits/s en première génération. Mais une si grande efficacité de compression ne peut être obtenue qu'au moyen d'un codage inter-image opérant sur des groupes (GOP) de 12 images. Or, si le fait de ne pas pouvoir briser un groupe de 12 images n'est pas particulièrement gênant en diffusion, il constitue un sérieux handicap en studio, où une grande précision de montage est indispensable. Si l'on veut modifier les paramètres de codage pour réduire la longueur du GOP, des débits dépassant largement 15 Mbits/s sont requis (50 Mbits/s avec un codage Intra). Enfin, le comportement en multigénération du Profil principal est clairement insuffisant; une perte notable de qualité peut être observée dès le troisième cycle de codage/décodage.

- *La compression inter-image MPEG-2 422@ML*

Le schéma de compression utilisé par le Betacam SX est différent des systèmes de type M-JPEG mis en œuvre dans les autres formats numériques. L'algorithme de réduction de débit employé repose en effet sur la norme MPEG-2 422@ML. Il a la particularité d'être le seul à exploiter les redondances temporelles entre les images vidéo. En fait, le codage s'effectue sur des groupes de deux images, l'une étant Intra (entièrement décrite par elle-même) et l'autre Bidirectionnelle (interpolée à partir des deux images I qui l'encadrent). L'utilisation d'une telle séquence I, B, permet d'obtenir un débit égal à 65 % de celui qui découlerait d'un codage de type M-JPEG tout intra, avec une qualité d'image équivalente. C'est cet argument qui a incité Sony à opter pour une telle solution, en dépit d'une certaine complexité engendrée pour assurer une précision de montage à l'image. Il est en effet impératif de rendre possible l'accès individuel à chaque image, malgré le fait que certaines sont codées en fonction d'autres. La technique imaginée pour résoudre ce problème consiste à toujours lire et enregistrer un groupe entier d'images I, B et de transformer, quand cela est nécessaire, une image bidirectionnelle en image monodirectionnelle (figure suivante).

Les plus du profil 4 :2 :2.

Le profil 4:2:2 reprend les outils de la norme MPEG-2 - il n'en utilise pas de supplémentaire -, mais avec de nouveaux paramètres. Le codage s'effectue tout d'abord sur des images en pleine résolution 4:2:2, bien que le traitement en 4:2:0 reste possible - certaines applications en contribution ou distribution pourront éventuellement s'en contenter. Par ailleurs, le codage ne se limite pas seulement à la partie active de l'image, mais inclut également une portion de suppression verticale. En traitant ainsi 608 lignes au lieu des seules 576 actives, ce Profil 4:2:2 ménage une zone de réserve pour véhiculer différentes données auxiliaires. La longueur maximale du GOP a, quant à elle, été réduite à seulement 2 images pour accroître la précision de montage.

De nombreuses applications, comme les systèmes de montage virtuel ou les enregistreurs sur disques durs, n'utilisent que des images I. L'efficacité de compression obtenue est, dans ce cas, environ 20 % supérieure à celle d'un système M-JPEG - même qualité avec un débit 20 % plus faible, ou qualité jugée meilleure à même débit (cela s'explique notamment par une gestion plus intelligente de l'entrelacé en MPEG-2, ainsi que par la possibilité d'utiliser des facteurs de quantification différents au niveau de chaque macrobloc). D'autres applications, comme le format d'enregistrement Betacam SX de Sony, mettent en œuvre des GOP de 2 images (I, B) offrant une meilleure efficacité de compression. Dans ce dernier cas, le problème du montage à l'image a été résolu de manière très astucieuse. A chaque fois qu'une image B doit être séparée de l'image I dont elle dépend - point de montage -, elle est recodée en incluant toutes les informations de l'image dont elle a besoin (une description plus détaillée de cette opération a été donnée dans le paragraphe " Les formats compressés ").

D'autre part, le Profil 4:2:2 assure une compatibilité descendante avec les Profils inférieurs. Autrement dit, un décodeur 422@ML est capable de traiter les flux numériques MP@ML, SP@ML, mais aussi MPEG-1. En revanche, la notion d'échelonnabilité propre aux Profils hiérarchiques ne constitue pas un besoin immédiat. Elle ne fait donc pas partie des paramètres du Profil 4:2:2 pour ne pas alourdir inutilement la complexité et le coût des circuits intégrés.

Avec un débit maximal de 50 Mbits/s, la qualité de l'image codée en MPEG-2 422@ML s'avère tout à fait convenable après des traitements de types effets spéciaux numériques ou chromakey, et se maintient jusqu'à 8 cycles de compression/décompression. Le standard MPEG-2 422P@ML est appelé à remplacer progressivement les multiples solutions M-JPEG propriétaires, dans la plupart de leurs applications. Un consortium baptisé Pro-MPEG Forum et regroupant un grand nombre de constructeurs majeurs a récemment été créé. Son but est de favoriser le développement d'équipements MPEG-2 422@ML multimarque à tous les maillons de la chaîne de production et de post-production broadcast, mais aussi et surtout de garantir leur interopérabilité. Les systèmes MPEG-2 422@ML devront cependant se partager le marché avec les systèmes basés sur le format DV et ses dérivés, qui fédère de son côté plus d'une cinquantaine de fabricants, certains appartenant même aux deux camps.

	MP@ML	422P@ML
Signal d'entrée	4:2:0	4:2:2 ou 4:2:0
Débit maximal	15 Mbits/s	50 Mbits/s
Nombre de lignes codées	576	608
Types d'images codées	I, B, P	I, B
GOP	12	2
Montage	non	oui
Capacités de multigénération	Limitées	oui
Applications	Diffusion TV Distribution sur DVD	Production Postproduction

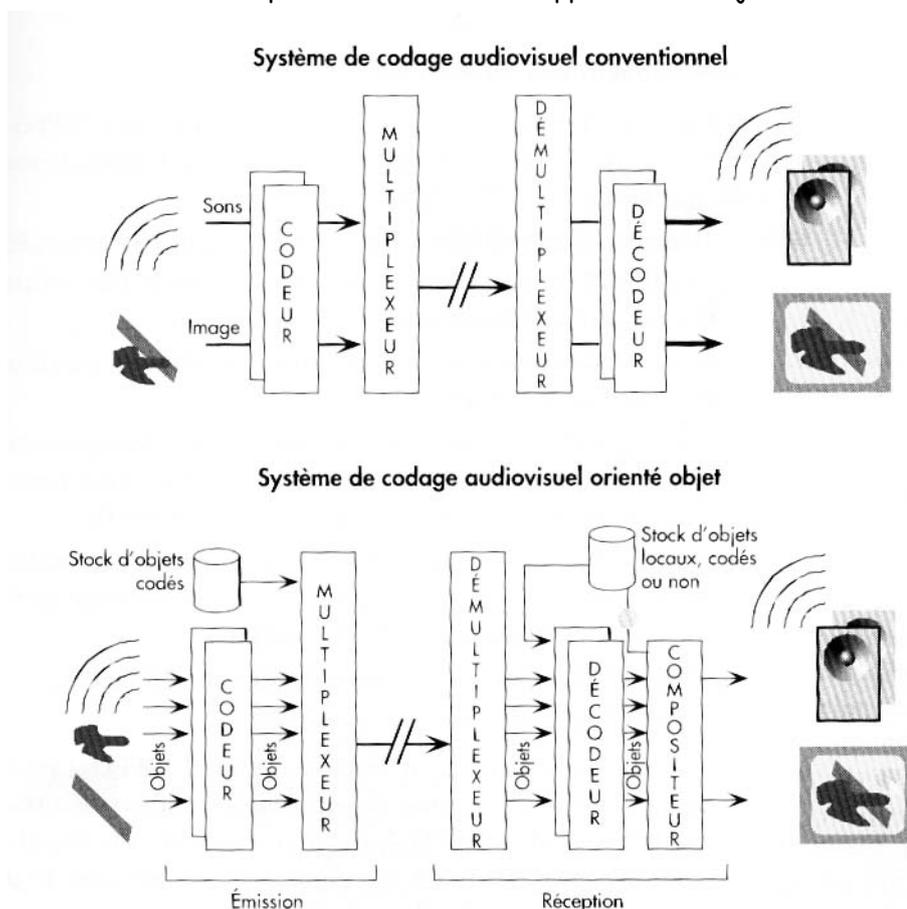
f) « The Mole » (la Taupe)

Toute manipulation d'une image codée en MPEG-2 impose de décoder le signal en 4:2:2, puis de le recoder en MPEG-2 après l'opération. Or plusieurs cycles de compression/décompression sont rapidement sources de dégradation de la qualité de l'image. Ceci est dû au fait que chaque codeur intervenant dans la chaîne réeffectue systématiquement tout le travail de compression sans tenir compte de celui effectué en amont par le codeur précédent. D'où le développement de la technologie « Mole ». Elle permet une succession de compression/décompression MPEG-2 sans détérioration du signal. Son principe consiste à extraire, lors du premier décodage, tous les paramètres relatifs à la compression effectuée en amont (GOP, vecteurs mouvement, quantification,...). Ces données sont alors insérées de manière imperceptible dans le signal vidéo numérique décompressé (d'où l'image de la taupe qui fait son chemin en souterrain, ou de l'espion). Chaque codeur qui suivra dans la chaîne reconstituera signal MPEG-2 en utilisant ces données au lieu de recalculer lui-même selon ses propres décisions. Ce processus peut s'apparenter à un système de clonage et peut être utilisé de multiples fois avec la même précision. La qualité d'image est alors conservée malgré de multiples cycles de codage/décodage. Elle est la seule à permettre le traitement multigénération de signaux MPEG-2.

g) MPEG-4: le codage des objets audiovisuels

MPEG-4 repose sur des bases radicalement différentes de celles des systèmes de codage audiovisuel traditionnels. Ces derniers se contentent en effet de traiter des séquences d'images rectangulaires de taille constante, composées chacune d'un certain nombre de lignes de pixels, et associées à une ou plusieurs pistes audio. De ce fait, le degré d'interactivité autorisé est minime, puisque restreint à l'aspect temporel du contenu: arrêt sur image, lecture en ralenti ou en accéléré, lecture arrière, etc. La composante élémentaire transmise est l'image, qui constitue de ce fait la plus petite entité sur laquelle l'utilisateur peut interagir. Il n'est pas possible d'intervenir au niveau du contenu de l'image. Cette limitation est éliminée par MPEG-4, dont la principale nouveauté est d'introduire la notion de codage orienté objet. MPEG-4 ne code pas simplement une succession d'images accompagnées de sons, mais individuellement les différentes composantes d'une scène. Une scène audiovisuelle est en effet considérée comme une association d'éléments visuels et sonores, identifiables, séparables et regroupables, ayant certaines relations dans le temps et dans l'espace. Ces éléments, appelés objets audiovisuels, peuvent être d'origine naturelle - c'est-à-dire captés par une caméra ou un microphone -, ou synthétique - issus d'un système informatique. Notons à ce sujet qu'il est aujourd'hui courant de générer des éléments naturels séparés, grâce aux techniques d'incrustation en chromakey sur fond bleu ou vert.

Par exemple, une scène basique peut être constituée de deux objets visuels, qui sont un personnage réel et un décor de synthèse, l'un étant incrusté sur l'autre. Deux objets audio peuvent leur être associés : la voix du personnage et un son d'ambiance. On peut enrichir l'ensemble avec une fenêtre en médaillon affichant une illustration, mais également un logo, un titre, etc. En fonction de leurs caractéristiques, ces différents objets sont codés individuellement et bénéficient ainsi d'outils de compression spécifiques, optimisés pour chacun d'eux. Un simple texte ne justifie pas en effet les mêmes algorithmes de compression qu'une image vidéo live ou qu'une animation 3D. L'efficacité de la réduction de débit globale est alors bien supérieure à celle qui résulterait d'une compression de type MPEG-2 appliquée à l'ensemble de l'image. L'organisation dans l'espace et le temps de ces objets au sein de la scène est décrite par un ensemble de données de composition. Celles-ci ont pour but de renseigner le décodeur sur la manière spatio-temporelle dont les objets qui lui parviennent doivent apparaître sur la scène. Les objets et les informations de composition sont protégés contre les erreurs, multiplexés et transmis sous la forme d'un flux numérique. Le débit du signal MPEG-4 peut actuellement s'étendre de 5 Kbits/s à 4 Mbits/s, mais des extensions vers le haut sont envisagées pour les applications de studio - un Profil à 16 Mbits/s est déjà prévu. Tous les éléments reçus en kit par le décodeur sont traités par un circuit appelé "compositeur". Celui-ci est chargé de les synchroniser et de les assembler, conformément aux données décrivant leur organisation sur la scène d'origine (sorte de plan de montage). L'utilisateur est cependant libre, de son côté, d'intervenir sur la composition de la scène: suppression d'objets d'intérêt secondaire, ajout



d'objets puisés dans une base locale, déplacement de certains éléments, choix du langage, etc. D'autre part, la notion d'échelonnabilité offre la possibilité, à un décodeur relativement simple, de ne traiter qu'une partie seulement du flux de données reçu, et de reproduire à sa manière une scène complexe, en réduisant par exemple les résolutions spatiale et temporelle de certains objets. MPEG-4 est ainsi capable de s'adapter aux performances variées des terminaux et systèmes d'affichage, ainsi qu'aux capacités des différents réseaux de transfert.

Les applications de MPEG-4

En autorisant l'interactivité sur le contenu audiovisuel, MPEG-4 deviendra le support d'une nouvelle génération d'applications et de services:

- communications en temps réel bidirectionnelles de type vidéo- téléphonie, sur une vaste gamme de systèmes de transmission et de configuration, avec de très courts délais;
- surveillance et contrôle d'accès, avec identification par reconnaissance visuelle ou vocale;
- réunions virtuelles entre protagonistes situés physiquement à divers endroits de la planète et matérialisés par des personnages de synthèse rassemblés dans une salle virtuelle;
- applications multimédia mobiles avec différents équipements (PDA, notebooks, stations portables, etc.), sur une large variété de canaux de transmission (GSM, satellite, etc.);
- bornes d'information et guides touristiques multimédia;
- diffusion vidéo sur Internet ou Intranet;
- applications broadcast: un train de transport MPEG-2 pourra en effet véhiculer, en plus de ses données audio/ vidéo traditionnelles, un flux MPEG-4. Celui-ci sera traité à la réception si un décodeur MPEG-4 est présent, et ignoré dans le cas contraire, sans perturber le reste du service. Parmi les applications possibles, on citera les guides de programmes évolués, les jeux interactifs, le téléachat (Pierre BELMARE), ainsi que le concept de télévision interactive utilisant un écran multifenêtre.

Balayage	Progressif et entrelacé
Résolution spatiale luminance	Toutes formes arbitraires de 8 × 8 à 2 048 × 2 048
Résolution spatiale chrominance	4:0:0, 4:2:0, et 4:2:2
Espace couleur	Monochrome, Y/ Cr/ Cb avec canal alpha
Débits	Constant ou variable Niveau faible : moins de 64 kbits/s Niveau intermédiaire : 64-384 kbits/s Niveau haut : 384 kbits/s – 4 Mbits/s

h) MPEG-7: la description des contenus multimédia

Avec la convergence des mondes de la télévision, de l'informatique et des télécommunications, la quantité d'information audiovisuelle disponible à travers le monde croît à une vitesse fulgurante. Pour pouvoir utiliser efficacement cette masse de données, encore faut-il être capable de localiser précisément celles dont on a besoin à un instant donné, sans être envahi par d'autres inutiles. Il existe actuellement des moyens de recherche basés sur une identification textuelle de l'information, comme les moteurs de recherche du web - sites les plus visités. Cependant, il n'existe pas encore de moyen, communément accepté, permettant d'identifier un document audiovisuel par une description de son contenu. MPEG-7 a pour objectif de combler ce manque, en standardisant la description de l'information multimédia quel qu'en soit le type, et indépendamment du langage. En d'autres termes, MPEG-7 spécifiera un ensemble de descripteurs caractérisant les images fixes, la vidéo animée, les graphiques, les modèles 3D, l'audio, le texte, etc. Les descripteurs MPEG-7 pourront également se rapporter à la manière dont ces éléments sont combinés dans une présentation multimédia - scénario, composition -, ainsi qu'à des informations qui peuvent être extraites du contenu, comme par exemple le contexte d'un enregistrement. Dans tous les cas, ces descripteurs seront totalement indépendants de la façon dont le contenu est codé. Ils seront soit physiquement associés au matériel décrit dans la même unité de stockage ou dans le même flux transmis, soit éloignés et reliés à lui au moyen de puissants pointeurs.

MPEG-7 fournira ainsi des outils permettant de développer des interfaces de recherche intuitives et conviviales. Au lieu de lancer des requêtes textuelles, il sera, par exemple, possible d'esquisser un dessin pour retrouver un tableau, siffloter approximativement le refrain d'un morceau musical ou jouer quelques notes sur un clavier pour obtenir les références d'un disque, décrire l'action d'une scène cinématographique et recueillir une liste de films les contenant, etc.

MPEG-7 sera capable d'indexer des objets MPEG-4, bien sûr, mais aussi des séquences MPEG-1 ou MPEG-2, et même des programmes analogiques. Plusieurs niveaux de description seront supportés en fonction des applications. Si l'on prend exemple d'un élément visuel, le niveau le plus bas pourra se restreindre à une description de sa forme, sa taille, sa texture, sa couleur, son mouvement, sa position dans l'espace. Un niveau plus élevé pourra donner une description sémantique d'une scène: "des enfants mangeant un chat dans un jardin anglais, sous un temps ensoleillé, avec un prof de TES dénudé fredonnant les tubes de F. MICKAEL." Des niveaux intermédiaires pourront également exister. Il est cependant important de souligner que MPEG-7 standardisera uniquement la description des contenus multimédia. Il ne standardisera ni l'analyse - extraction de l'information -, ni le moteur de recherche. Ces deux points sont liés à l'application et resteront donc dans le domaine compétitif, ce qui leur permettra d'être constamment améliorés, plutôt que d'être figés à un état technologique donné.

Pour résumer, MPEG-7 standardisera:

- un jeu de descripteurs, indépendant des langues pour la majorité d'entre eux. On appelle descripteur une valeur de représentation associée à une ou plusieurs caractéristiques d'un élément. Par exemple, la couleur d'un point, l'auteur d'un texte, le style d'une vidéo, la longueur d'un plan, etc.;
- des schémas de description, c'est à dire une structure et une sémantique permettant de spécifier les descriptions et leurs relations. C'est par exemple: pour un film, le titre, le réalisateur, les acteurs, les scènes; pour une scène donnée de ce film, la description de l'action, sa durée, etc.; pour une image donnée de cette scène, sa valeur de plan, sa dominante colorée, etc.;
- un langage pour les schémas de description;
- un système de codage efficace pour tous ces éléments, qui soit adapté au stockage et à la transmission.

Les applications de MPEG-7

Parmi les applications visées par MPEG-7 figurent en premier plan les bibliothèques digitales - catalogues d'images, dictionnaire musical, etc. -, les répertoires de services multimédia - pages jaunes, recherche de personnes par leur apparence physique, etc.-, la sélection de canaux de radio, T.V, ou Internet, l'édition multimédia - services de journaux électroniques personnalisés, etc. Les domaines concernés sont nombreux: éducation, tourisme, services d'investigation, médecine, architecture, archives de cinéma et T.V, journalisme, shopping, divertissement, etc. MPEG-7 devrait devenir un standard international à la fin de 2001.

i) Le MPEG-IMX

Lancé en 2000 par Sony, le MPEG-IMX est un nouveau format dans la famille Betacam, positionné à mi-chemin entre le Betacam SX (reportage news) et le Digital Betacam (postproduction haut de gamme). Résolument orienté vers l'environnement MPEG-2 actuel et futur, le MPEG-IMX ne tire pas pour autant un trait sur le passé puisqu'il offre une compatibilité en lecture avec la totalité des formats Betacam analogiques et numériques de Sony sortis au cours de ces 20 dernières années.

Un format en MPEG-2 tout intra

Le format MPEG-IMX (Interoperable Mpeg eXchange) est basé sur l'enregistrement d'un signal 4:2:2, dont le débit initial de 166 Mbits/s est réduit dans un facteur 3,3:1 à 50 Mbits/s, par un algorithme MPEG-2 422 (notons que ces valeurs sont les mêmes qu'en DVCPRO50 et qu'en D9, à la différence près que ces deux formats utilisent une compression DV). L'algorithme de compression mis en œuvre par le MPEG-IMX n'utilise, dans la boîte à outils du standard MPEG-2, que ceux opérant en mode "intra-image". Le processus de réduction de débit se restreint en effet à l'élimination des redondances spatiales présentes à l'intérieur de chaque image et ne tient pas compte des similitudes, pourtant non négligeables, entre images successives. C'est là un point qui différencie fondamentalement le MPEG-IMX du Betacam SX, l'autre format MPEG-2 de Sony, qui code les images par groupes de deux pour gagner en efficacité de compression, ce qui complexifie le traitement des signaux lors d'un montage à l'image près. Pour son nouveau format MPEG-IMX, Sony est revenu à un schéma de compression plus classique et bien mieux accepté lorsqu'il s'agit de postproduction broadcast. Il consiste tout simplement à coder individuellement chaque image en se référant uniquement à son propre contenu (codage tout-intra), comme en M-JPEG ou en DV, et à accepter en contrepartie de manipuler des débits supérieurs. Le montage à l'image près est alors possible "naturellement", sans aucun traitement particulier, et les performances en multigénération sont meilleures que celles du Betacam SX. La qualité de l'image MPEG-IMX est visuellement identique à celle du Digital Betacam jusqu'à sept cycles de décodage/encodage, ce qui devrait convenir à la plupart des productions courantes. Rappelons toutefois que le Digital Betacam affiche des caractéristiques techniques sensiblement supérieures à celles du IMX puisque la compression, également intra-image mais en M-JPEG (non normalisée), s'effectue dans un rapport de 2:1 au lieu de 3,3:1. Le Digital Betacam reste, avec le D5 de Panasonic, le format de référence pour la postproduction haut de gamme complexe et le mastering.

Compatibilité avec la famille Betacam

Pour mettre au point un magnétoscope capable de relire tous les formats Betacam analogiques et numériques, Sony a dû revoir son transport de bande 1/2". La gamme MPEG-IMX intègre donc une nouvelle plateforme 1/2", dont l'un des principaux atouts est un gain substantiel en volume. Les magnétoscopes de studio à ce format se présentent ainsi dans un coffret plus compact. Si le format MPEG-IMX est techniquement conçu pour relire tous les formats Betacam, les niveaux de compatibilité offerts varient cependant en fonction des modèles de machines. Dans tous les cas, les formats analogiques et numériques relus sont codés en MPEG-2 422 et disponibles sur la sortie compressée SDTI-CP. La commutation des circuits de lecture s'effectue automatiquement lors du chargement de la cassette.

Caractéristiques du format MPEG-IMX

Le MPEG-IMX utilise un tambour de tête de même diamètre que le digital Betacam et que le Betacam SX, soit 8,14 cm. SONY garde secrète la composition de ce tambour dont on imagine la complexité, quand on sait que chaque format lui nécessite individuellement la présence de 4 têtes supplémentaires. La vitesse de défilement de la bande est de seulement 5,4 cm/s, soit la plus faible de toute la famille Betacam, la vitesse de rotation du tambour est réduite de 75 à 50 tours/s, et la largeur des pistes vidéo n'est que de 21,7µm (une image est segmentée sur 8 pistes hélicoïdales). Le MPEG-IMX enregistre soit 4 pistes audio en 48 kHz/24 bits, soit 8 pistes audio en 48 kHz/16 bits. Il est ainsi le premier format capable de gérer le format multicanal en 5+1 (6 pistes), avec en plus la possibilité d'enregistrer un mixage stéréo classique sur les deux pistes restantes. Le MPEG-IMX offre la fonction "pre-read" et permet une lecture à vitesse variable sur une plage -1/+3x (comme le Digital Betacam et le D9). A l'instar de tous les formats Betacam, il existe deux tailles de cassettes MPEG-IMX que l'on identifie

immédiatement par la couleur verte de leur boîtier (le Beta SP est gris, le Digital Beta est bleu et le Beta SX est jaune). Du fait de la faible vitesse linéaire, les durées maximales offertes sont particulièrement élevées: 72 min sur une petite cassette (S) et 220 min sur une grande cassette (T).

j) Liaison SDTI, liaison SDI:

Normalisée en 1997, la liaison SDTI (Serial Data Transport Interface) est une extension de la liaison numérique série SDI, adaptée au transport d'un signal compressé entre deux équipements (magnétoscopes, serveurs, stations de montage non linéaire de différentes marques). Alors que la SDI véhicule uniquement le signal vidéo 4:2:2 à plein débit (avec ses 4 pistes audio), la SDTI peut transporter des données numériques compressées, structurées en paquets. Son atout est de permettre de continuer à utiliser toute l'infrastructure physique SDI d'un studio numérique (câbles, grilles de commutation..) et d'offrir une grande fiabilité de transport, supérieur à celles des liaisons informatiques et telecom. Polyvalente, la liaison SDTI est utilisée pour les échanges en mode natif (sans décompression/compression) par les systèmes DVCAM, DVCPRO, D9, Betacam SX et MPEG-IMX (à l'intérieur d'un même format!). Selon les cas, le flux compressé transite à 2 ou 4 fois la vitesse normale, profitant du débit de 270 Mb/s de la liaison SDI. L'identification du type de signal est transmise par des informations supplémentaires, en en-tête des paquets de données utiles. La version MPEG-2 de l'interface SDTI utilisée par le MPEG-IMX est baptisée SDTI-CP (Content Package).

V. LES FORMATS NUMERIQUES :

On distingue aujourd'hui trois grandes catégories de formats d'enregistrement de vidéo numériques.

➤ Les formats en composantes numériques : D1, D5 et D6

Ils enregistrent le signal en composantes 4 : 2 : 2 à plein débit, avec une quantification sur 8 ou 10 bits. Ils sont totalement compatibles 525/625 lignes et sont caractérisés par un excellent rapport S/B de l'image, ainsi que par une totale transparence en multigénération. Les puissants systèmes de correction d'erreurs corrigent de manière très efficace les pertes d'informations sur la bande, si bien que le signal relu est exempt de drop out.

➤ Les formats composites numériques : D2 et D3

Ils enregistrent le signal composite PAL ou NTSC numérisé à 4 fois la fréquence de la sous porteuse de la chrominance du système analogique, avec une quantification sur 8 bits. Comme pour les formats en composantes numériques, le rapport S/B est très élevé et les problèmes de drop out sont inexistantes. En revanche, la notion d'universalité apportée par les formats en composantes disparaît, puisque les machines PAL et NTSC sont incompatibles entre elles. De plus, les limitations inhérentes au signal composite sont conservées.

➤ Les formats compressés : Digital Betacam, Famille DV, Betacam SX, D9, etc.

Ils enregistrent, selon le cas, un signal en composantes 4 : 2 : 2 , 4 : 2 : 0 ou 4 : 1 : 1, numérisé sur 8 ou 10 bits et soumis à une réduction de débit de type M-JPEG ou MPEG-2. Ces nombreux formats, que nous allons détailler dans les pages suivantes, ont permis une véritable démocratisation de l'enregistrement numérique, aussi bien dans le domaine broadcast qu'auprès du grand public.

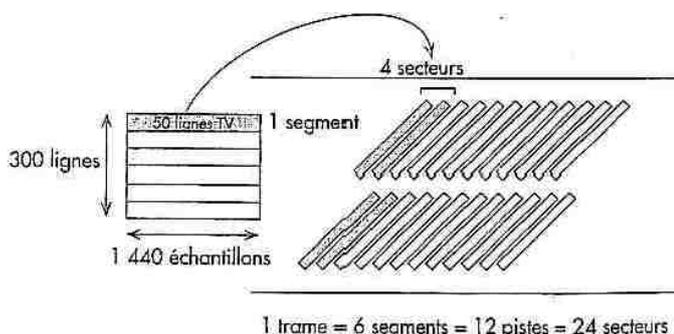
1. Les premiers formats numériques :

a) **D1 :**

Le D1 est un format d'enregistrement en composantes numériques. Il traite le signal vidéo 4:2:2 sur 8 bits, ainsi que 4 pistes audio PCM échantillonnées à 48 kHz et codées sur 20 bits. Le support magnétique est une bande oxyde de largeur 19 mm (3/4"), conditionnée dans trois tailles de cassettes (6, 34 et 94 minutes). Unique pour les standards 525 et 625 lignes, le D1 est le premier format numérique à avoir vu le jour. Mais son coût relativement élevé a été un sérieux obstacle à son utilisation, d'autant que ce format en composantes est apparu à une époque où toutes les installations étaient en composite.

➤ La segmentation :

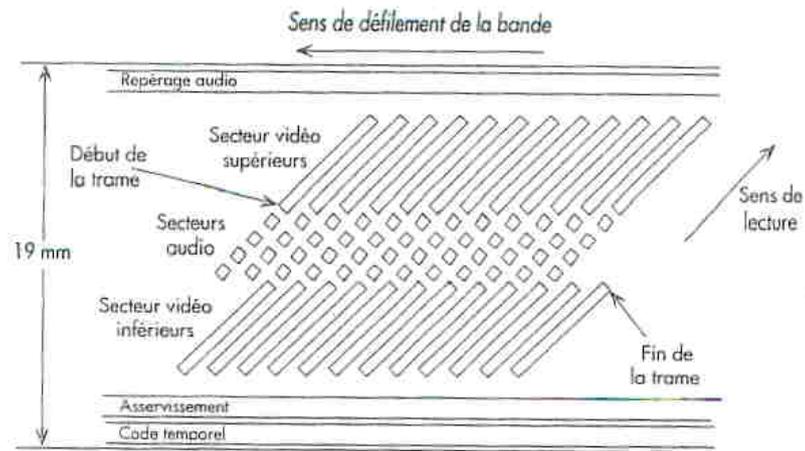
Un rapide calcul va nous donner une idée de la vitesse à laquelle doit théoriquement défiler la bande magnétique pour pouvoir enregistrer le signal numérique 4:2:2. Avec une fréquence d'échantillonnage de 13.5 MHz pour la luminance et de 6.75 MHz pour les deux signaux de différence de couleurs, on compte un total de 1440 échantillons par ligne. Avec 576 lignes utiles par image et 25 images par seconde, le débit vidéo seul est de 172.8 Mbits/s. Si l'on y ajoute les 8 Mbits/s de l'audio et les 47 Mbits/s de données supplémentaires pour la correction d'erreurs, on arrive à un débit total de 227 Mbits/s. Une trame renferme donc un total de 227/50, soit environ 4.5 Mbits. Or, pour enregistrer sur une seule piste oblique ces 4.5 Mbits, il faudrait, avec une densité d'intégration de 2 bits par micron, une piste longue de plus de 2 mètres ($4.5/2 \sim 2.2\text{m}$). Quant au tambour, son diamètre doit être de $2.2/\pi = 70\text{ cm}$. Tout cela montre qu'il est indispensable de segmenter une trame en plusieurs pistes et d'enregistrer ces pistes par plusieurs têtes.



Dans le format D1, une trame en 625/50 est répartie sur 12 pistes obliques. Chacune d'elle est divisée en deux secteurs, ou demi pistes et forme donc une trame. Un segment contenant les informations de 50 lignes est constitué du regroupement de 4 secteurs enregistrés chacun avec une tête différente. Signalons pour information qu'en 525/50, on compte seulement 10 pistes par trame donc 20 secteurs vidéo et 5 segments de 50 lignes. Les autres paramètres du format sont communs aux deux standards.

➤ La structure d'enregistrement :

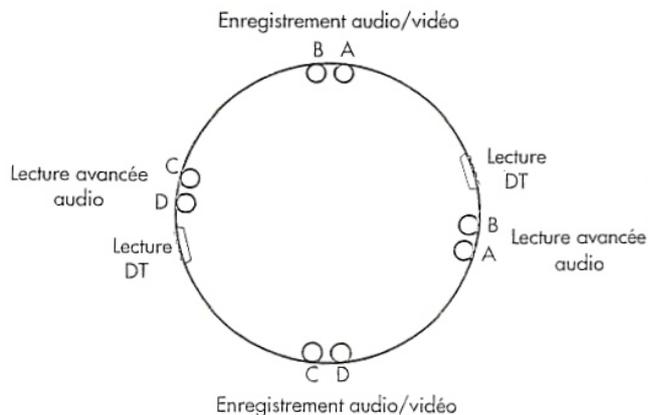
Trois pistes longitudinales sont présentes pour le repérage audio, l'asservissement du cabestan et le code temporel. La piste de repérage audio permet d'écouter le son sur une grande plage de vitesse ; elle est cependant analogique et sa qualité est inférieure à celle des pistes numériques. Du fait de la segmentation, la piste d'asservissement contient notamment des impulsions arrivant tous les deux segments, c'est-à-dire une fois à chaque rotation du tambour, soit à 150 Hz en lecture normale. Les pistes hélicoïdales sont enregistrées sans azimuth et sont donc séparées par un intervalle de garde de $5\ \mu\text{m}$ visant à éviter tout problème d'intermodulation. Les secteurs audio occupent une place privilégiée au centre de la bande (entre les secteurs vidéo), là où le contact tête/bande est optimal. Une redondance de 100% est par ailleurs appliquée au signal audio, dont les échantillons sont enregistrés deux fois à des endroits différents.



Chaque piste oblique contient successivement un secteur vidéo, quatre secteurs audio et un autre secteur vidéo. Tous les secteurs sont séparés par des intervalles de garde qui permettent d'éditer séparément l'image et les canaux audio. La longueur d'une piste est de 170 mm - pour une largeur de $40\mu\text{m}$ - et celle d'un secteur vidéo est de 78 mm.

➤ **Le tambour de têtes :**

Le tambour comporte 16 têtes : 4 pour l'effacement/enregistrement, 4 pour la lecture avancée audio, et 2 blocs de têtes de lecture Dynamic Tracking. Les têtes de lecture avancée audio permettent, en mode enregistrement, de lire le son présent sur la bande afin de le traiter ou de le mélanger avec une autre source, et de réenregistrer immédiatement le résultat de cette opération sur la piste d'origine. Les têtes DT permettent, quant à elles, une lecture audio/vidéo sur une plage de -1 à 2* la vitesse nominale.



Quatre d'entre elles permettent en outre de contrôler les signaux audio et vidéo inscrits sur la bande au cours de l'enregistrement (mode « confiance »). La bande d'épaisseur 13 ou 16 μm défile à 2806 cm/s, la vitesse d'écriture étant de 3506 m/s. Quatre têtes d'enregistrement tracent simultanément les 4 secteurs d'un segment.

➤ **Le traitement vidéo :**

Nous allons à présent passer en revue les principales étapes du traitement que subissent les données vidéo avant leur enregistrement.

Mapping vidéo (précodage de source) :

Les 8 bits de chaque échantillon vidéo sont brassés afin que soit répartie uniformément toute erreur éventuelle sur l'ensemble de l'échelle de luminance ou de chrominance. Une opération inverse est bien sûr effectuée à la lecture pour rétablir l'ordre chronologique des données.

Distribution inter secteur (brassage intra-ligne) :

Les échantillons vidéo de chaque segment de 50 lignes sont distribués sur 4 secteurs, enregistrés simultanément sur 4 têtes séparées. Chaque canal ne gère alors qu'un débit d'environ 55 Mbits/s. Un brassage intra-ligne est ainsi réalisé ; chaque secteur ne contient donc que 50 « $\frac{1}{4}$ de lignes ». C'est alors que sont calculés et ajoutés les codes de protection d'erreurs, d'abord dans le sens vertical, puis dans le sens horizontal.

Codage externe (outer code) :

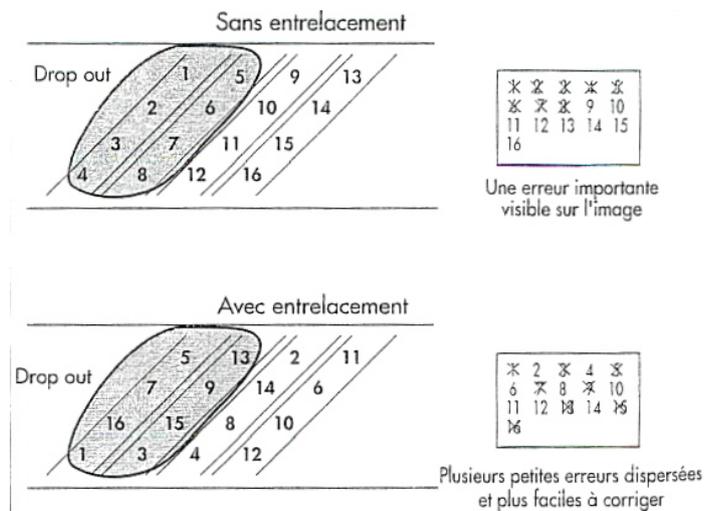
Les données de chaque secteur sont organisées sous la forme d'un tableau de 30 rangées et 600 colonnes. Pour chaque colonne sont calculés deux octets de contrôle de correction (code Reed Solomon) : c'est le codage externe qui permet de corriger les erreurs courtes.

Codage interne (inner code) :

Après avoir été soumises à un brassage intra-secteur, les données vidéo et leurs codes externes sont envoyés vers le processus de codage interne opérant cette fois dans le sens horizontal. Le tableau secteur précédent est alors découpé en 10 blocs de 60 colonnes. A chaque rangée est rajouté un mot de 4 octets. Ce codage interne permet de corriger les erreurs longues - 1 octet par ligne sur chaque bloc.

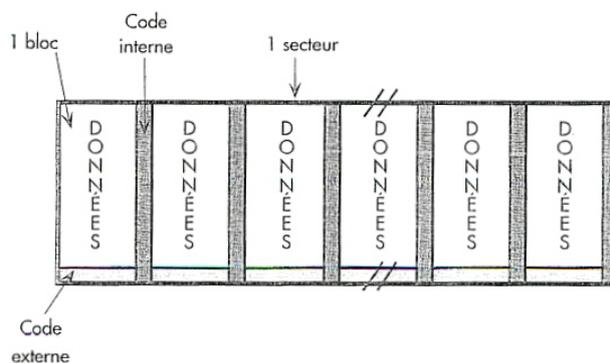
Découpage en blocs de synchronisation :

Des informations permettant de structurer les données en blocs de synchronisation et le plus petit ensemble de données enregistrées détectable. Un secteur vidéo est formé par un ensemble de 160 blocs et délimité par un préambule et un postambule.



Codage de canal :

Le codage de canal a pour rôle de mettre les données vidéo et audio sous la forme d'un signal série adapté à l'enregistrement magnétique. Le D1 utilise un code NRZ embrouillé, qui minimise les longues suite de 0 ou de 12, réduisant ainsi les composantes basses fréquences du flux numérique.



➤ **Le traitement audio :**

La densité d'intégration sur la bande est telle que chaque canal audio ne requiert que 6 secteurs par trame, soit deux fois moins que ce qui est disponible sur la bande. Il est donc possible d'enregistrer deux fois chaque échantillon audio à des emplacements différents. La protection supplémentaire ainsi apportée permet d'accroître l'efficacité de la correction d'erreurs et d'obtenir une qualité audio parfaite. Même en cas de problèmes de bruit, perte de niveau, encrassement des têtes, défauts sur la bande, etc. Si, par exemple, la tête 0 est défaillante, toutes les données de la piste 0 peuvent ainsi être retrouvées sur les pistes voisines. La redondance de 100% ainsi appliquée au signal audio porte son débit total à 8 Mbits/s.

b) D2 :

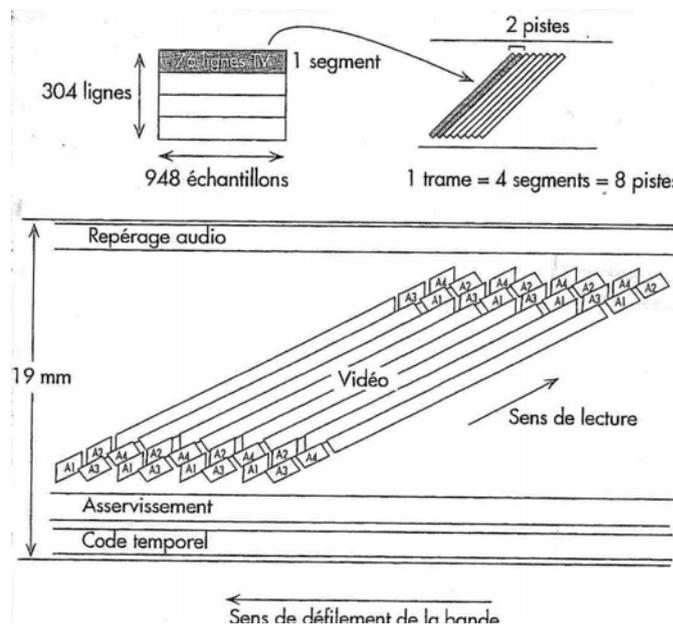
Le format D2 enregistre le signal vidéo composite numérisé sur une bande $\frac{3}{4}$ ". C'est dans le but de remplacer les machines 1" C par un équivalent numérique que la société Ampex a conçu le D2. Du fait qu'il travaille sur un signal composite, le D2 existe en deux versions incompatibles, l'une traitant le signal PAL et l'autre la signal NTSC. Les deux versions de D2 ont des empreintes de bande différentes, des débits binaires différents, mais possèdent néanmoins un grand nombre de points communs et utilisent, pour une large part, les mêmes circuits électroniques.

La fréquence d'échantillonnage a été choisie égale à 4 fois la fréquence de la sous porteuse couleur, elle-même multiple de la fréquence ligne. Or la fréquence ligne est différente dans chacun des deux standards, ce qui donne une fréquence d'échantillonnage de 17.73MHz (4×4.43) en PAL et de 14.32 MHz en NTSC (4×3.5). Une ligne de 64µs contient 1135 échantillons de 8 bits. Sa partie active n'en comporte que 948, qui sont en fait les seuls à être enregistrés. Les quatre canaux audio échantillonnés à la fréquence de 48 kHz et quantifiés sur 20 bits, avec une dynamique de 105dB.

Le débit total des informations à enregistrer, incluant les codes de correction d'erreurs, s'élève à 152 Mbits/s.

➤ **La structure d'enregistrement :**

Le D2 utilise une cassette de mêmes dimensions que le D1, mais le type de la bande et le motif d'inscription des pistes sont différents. Le D2 utilise une bande à particules métalliques, d'épaisseur 13 µm. L'enregistrement est ici réalisé avec azimuth : les pistes adjacentes de largeur 35 µm sont tracées avec un entrefer doté d'un angle d'inclinaison alterné de -15° et +15° et peuvent donc se chevaucher sans que cela n'engendre de parasitage inter piste. La vitesse de défilement linéaire est de 13.1 cm/s (en 625/50), soit un peu plus de deux fois plus inférieure à, celle du D1. Toujours en comparaison avec le D1, le débit des données du D2 est plus faible et la densité d'enregistrement est supérieure.



Les autonomies maximales offertes par les trois tailles de cassettes sont par conséquent nettement supérieures, soit respectivement : 32,94 et 208.

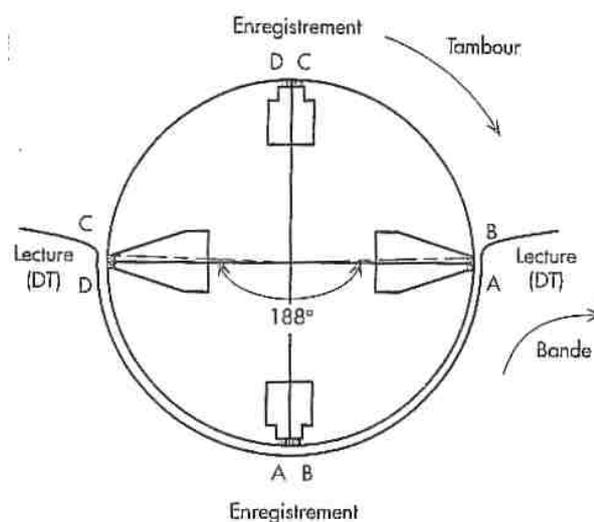
L'enregistrement est bien entendu segmenté. Une trame est divisée en 4 segments constitués de 76 lignes chacun. Un segment se subdivise en deux secteurs, dont chacun reçoit un échantillon sur deux.

Les quatre pistes audio numériques sont enregistrées sur des secteurs situés de part et d'autre de la piste vidéo, sur les bords de la bande. Tout comme le D1, chaque échantillon audio est enregistré deux fois. Les données audio sont organisées en segments de durée $\frac{1}{4}$ de trame chacun. Un segment audio est divisé en deux secteurs.

On trouve aussi sur la bande trois pistes longitudinales analogiques allouées au time code, au repérage audio et à l'asservissement.

➤ **Le tambour des têtes :**

Le tambour de têtes de diamètre 96 mm porte deux paires de têtes pour l'enregistrement et deux paires de têtes pour la lecture. Ces dernières sont à suivi dynamique de pistes et permettent une lecture sans parasite sur une plage de -1 à $+3*$ la vitesse nominale. Le tambour effectuant 100 rotations par seconde, deux tours sont nécessaires pour l'enregistrement des 8 pistes formant une trame. En enregistrement, les têtes de lecture peuvent jouer le rôle de têtes « confiance », fournissant ainsi un monitoring simultané de l'audio et de la vidéo. En mode insert, ces mêmes têtes font office de têtes de prélecture, permettant alors d'utiliser les signaux audio et vidéo présents sur la bande comme source de montage. C'est la fonction « pre-read ».



➤ **Le traitement du signal :**

Dans son principe, le traitement des données numériques dans le format D2 est similaire à celui du D1. Chacun des 4 segments d'une trame est décomposé en deux secteurs, constitués de 76 lignes de 474 échantillons. Pour ce, le premier échantillon du segment est envoyé sur le secteur 0, le second dans le secteur 1, le suivant dans le secteur 0, et ainsi de suite. Dans chaque secteur est alors opéré un brassage intra-ligne. Puis est réalisé le codage externe (vertical), qui calcule et ajoute 4 octets par colonne du tableau secteur ; sa capacité de correction est d'environ 4 erreurs par colonne. Après brassage intra-secteur, le codage interne (horizontal) consiste à ajouter, à chaque groupe de 76 colonnes, 8 octets de protection par ligne. La capacité de correction est ici de 3 octets par rangée.

c) **D3 :**

Le D3 est un format composite numérique conçu par la NHK et développé par Panasonic. Le D3 effectue globalement le même travail que le D2, mais sur une bande 1/2" (au lieu de 3/4"), ce qui a notamment permis au premier caméscopes numérique de voir le jour. Le D3 enregistre un signal vidéo PAL ou NTSC numérisé avec une fréquence d'échantillonnage égale à 4 fois la fréquence de la sous-porteuse couleur et quantifié sur 8 bits. Il traite également quatre pistes audio échantillonnées à 48 kHz et codées sur 20 bits. Le format D3 accepte trois tailles de cassettes, dont les durées maximales sont de 64,125,245 minutes. Le D3 a bénéficié de l'expérience des formats numériques qui l'ont précédé et utilise des stratégies de correction d'erreurs plus performantes, notamment en ce qui concerne l'audio.

➤ **La structure d'enregistrement :**

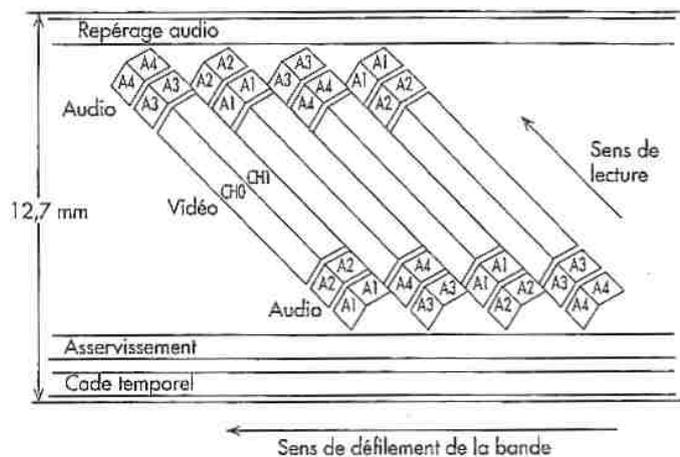
Le format D3 utilise une bande à particule métalliques de largeur 1/2"(12.7 mm), d'épaisseur 11 ou 14µm et défilant à 83.88 mm/s. Une trame est décomposée en quatre segments, un segment étant enregistré sur deux pistes. Un total de 8 pistes est donc nécessaire pour l'inscription d'une trame en PAL . Comme le D2, le D3 utilise deux canaux pour l'enregistrement : les échantillons pairs sont traités sur le canal 0 et les échantillons impairs sur le canal 1. Afin d'obtenir une plus haute densité d'informations sur la bande, le D3 utilise un angle d'azimut de plus ou moins 20° et une largeur de piste de seulement 18 µm en PAL. La densité d'enregistrement du D3 a ainsi pu être doublée par rapport à celle du D2 et triplée par rapport à celle du D1.

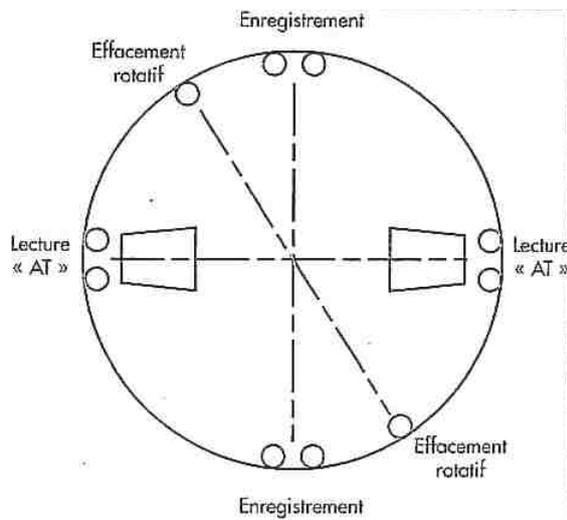
Les données relatives aux 4 canaux audio sont organisées en secteurs enregistrées de part et d'autre des pistes vidéo.

A la seule différence des précédents formats D1 et D2, qui enregistreraient deux fois les mêmes données, le D3 se contente de ne les enregistrer qu'une seule fois, en respectant une répartition particulière. Cela a été rendu possible grâce à la puissance du circuit capable de restituer les données, même si près de 50% d'entre elles sont manquantes.

➤ **Le tambour des têtes :**

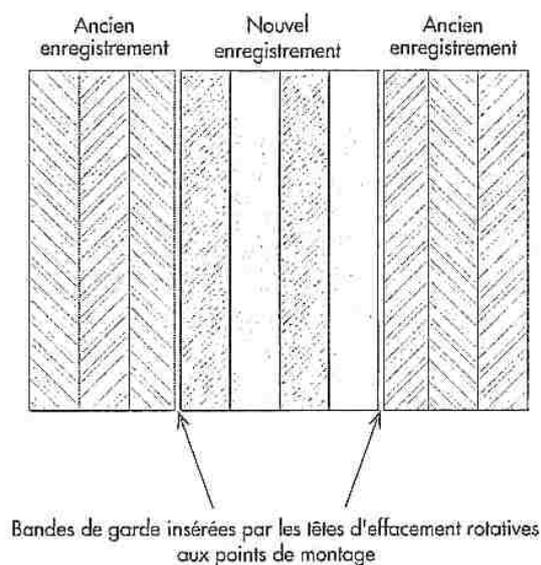
Le tambour de diamètre 76 mm porte deux paires de têtes d'enregistrement diamétralement opposées, deux paires de têtes de lecture à suivi dynamique de piste, ici appelées « Automatic Tracking » (AT), permettant une lecture de l'image de -1 à +3 fois la vitesse nominale, et deux têtes d'effacement rotatives. Chaque paires de têtes d'enregistrement trace deux piste en une rotation du tambour, une pour le canal 0, l'autre pour le canal 1.





Les têtes Automatic Tracking jouent un rôle différent selon que la machine fonctionne en mode lecture normale ou en mode confiance. En lecture normale, elles assurent une prélecture avec une phase avancée par rapport aux têtes d'enregistrement (fonction pre-read). En mode lecture confiance, les têtes AT permettent une lecture de contrôle avec une phase retardée par rapport aux têtes d'enregistrement.

Le D3 est le premier format à avoir utilisé des têtes d'effacement rotatives. Celles-ci travaillent par segment, c'est-à-dire par paire de pistes, leur largeur étant légèrement supérieure à celle de deux pistes réunies. Elles produisent de ce fait une bande de garde aux points de montage, dans le but d'éviter une hausse du taux d'erreurs engendrée par un problème de tracking et de ramener ainsi le rapport signal/bruit à sa valeur normale. De plus, les anciennes données présentes sur la bande sont totalement supprimées par les têtes rotatives d'effacement au lieu d'être simplement écrasées par les nouvelles.



Le D3 n'utilise qu'un seul guide de bande fixe avec un angle d'enroulement de bande de 37°. La tension appliquée à la bande sur le cylindre à l'entrée du tambour est quatre fois plus faible qu'en D2. Cela offre plusieurs avantages, parmi lesquels une réduction du risque d'encrassement des têtes et d'usure de la bande.

➤ **Le traitement du signal :**

Alors que dans les formats D1 et D2, les données sont entrelacées uniquement au sein de blocs de trames, dans le format D3, elles sont brassées sur une trame entière. Un tel procédé réduit d'avantage les effets de rafale d'erreurs en répartissant celles-ci sur l'ensemble de la trame, ce qui accroît l'efficacité de la correction d'erreurs. Comme pour les autres formats vidéo numériques, le D3 utilise un double code d'erreurs travaillant sur deux axes. Sa capacité de correction totale peut s'étendre à 15 lignes.

Avant l'enregistrement, les données de chaque ligne sont mélangées entre elles. Le code correcteur d'erreurs externe de 8 octets est alors ajouté verticalement ; il peut corriger jusqu'à 8 erreurs dans une colonne de segments. Le flux des données du signal vidéo traverse ensuite un séparateur de canal, qui envoie les mots alternativement sur le canal 0 et sur le canal 1. Lors d'une rotation du tambour, chaque tête enregistre deux pistes. Si une tête est s'encrasse, seule la moitié des informations est perdue. La position des octets de données est alors à nouveau réarrangée, mais cette fois ci à l'intérieur de toute la trame. Puis les signaux vidéo sont multiplexés dans le temps avec les données audio. Le code de correction d'erreur interne de 8 octets est alors ajouté horizontalement; il peut corriger jusqu'à 4 erreurs dans un bloc.

L'ensemble des données est ensuite soumis, avant l'enregistrement, à un codage de canal « 8-14 » qui, transforme tous les mots de 8 bits en mots de 14 bits. Cette opération a pour but de minimiser la composante continue et de réduire la gamme de fréquences, ce qui permet de faire circuler le signal sur un canal de bande passante plus faible.

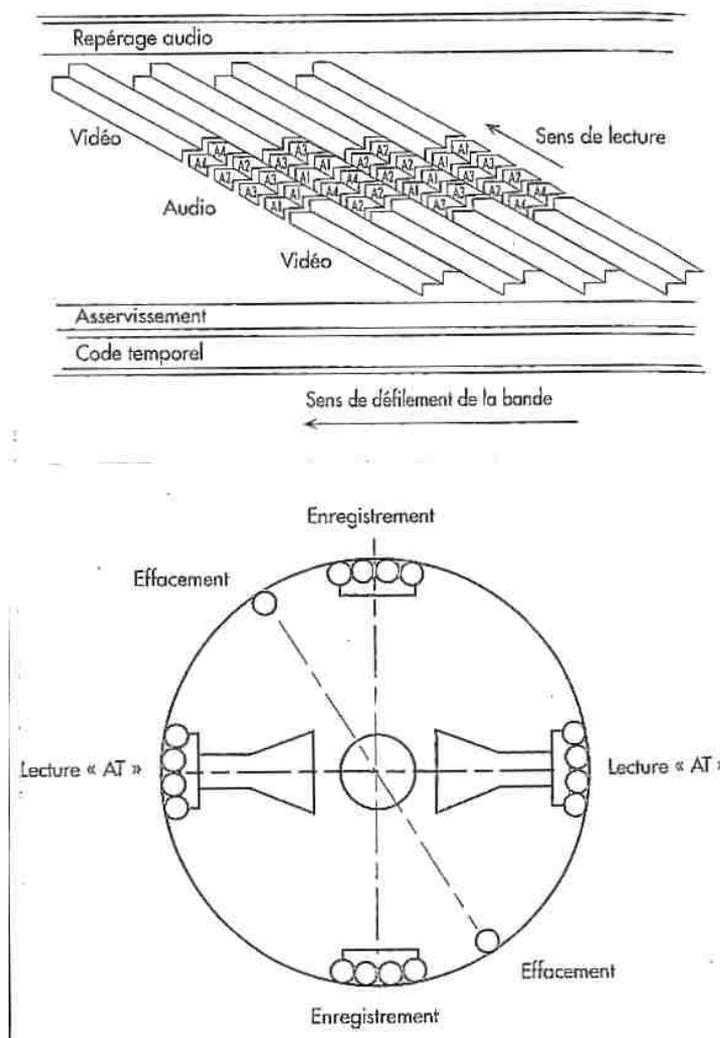
d) D5 :

Lancé en 1994 par Panasonic, le D5 est un format 1/2"composantes numériques dont la technologie et les paramètres mécaniques sont en grande partie empruntés au D3. L'utilisation, par ces deux formats de la même cassette et de la même empreinte sur la bande permet aux magnétoscopes D5 de relire des programmes composites enregistrés sur une machine D3. Le D5 est l'unique format capable d'enregistrer le format vidéo 4:2:2 sans compression avec une quantification de 10 bits par échantillon. En termes de débit, la capacité totale du D5 est de 288 Mbits/s (218 Mbits/s pour la vidéo en 4/3-13.5 MHz).

➤ **La bande et les têtes :**

Le tambour de diamètre 76 mm tourne à 100 Hz, comme en D3, mais la vitesse de défilement de la bande a été doublée pour atteindre 16.7 cm/s. Les durées maximales offertes par les trois tailles de cassettes sont par conséquent réduites de moitié par rapport au D3, soit 124,63 et 23 minutes (avec une bande d'épaisseur 11 μ m).

Le débit supérieur requis par l'enregistrement en composantes numériques a par ailleurs imposé de doubler le nombre de canaux d'enregistrement, qui passe ainsi de 2 à 4. Le tambour comporte 18 têtes : 2 groupes de 4 têtes d'enregistrement diamétralement opposés, 2 ensembles de 4 têtes à suivi de piste « AT » autorisant une lecture à vitesse variable de -1 à +2* la vitesse nominale, et 2 têtes d'effacement rotatives fonctionnant comme en D3. Le magnétoscope D5 est doté de la fonction pre-read qui consiste, en mode enregistrement, à lire le signal vidéo présent sur la bande quelques trames avant qu'il ne soit effacé et remplacé par les têtes d'enregistrement.



➤ **Le traitement du signal :**

Le D5 traite aussi bien un signal échantillonné à 13.5 MHz avec une résolution de 10 bits (pour les 1440 échantillons du format 4/3), qu'un signal échantillonné à 18 MHz avec une quantification sur 8 bits (1920 échantillons, 16/9). Il faut cependant savoir que pour des raisons de simplification, le D5 ne manipule en interne qu'un signal 18 MHz/8 bits, le signal 13.5 MHz/10 bits est par conséquent soumis, dès son entrée dans la machine, à un formatage 10-8. Les 1440 échantillons de luminance et de chrominance (en 4/3) codées sur 10 bits entrent dans une mémoire d'une ligne, dans laquelle ils sont inscrits sous la forme d'octets. A parti de ce moment, le traitement s'effectuera sur des mots de 8 bits quel que soit le type de signal d'entrée. Un brassage intra-ligne est alors effectué. Puis une séquence pseudoaleatoire est ajoutée au signal pour briser les longues suites de 0 et de 1.

Chacun des 4 canaux d'enregistrement possède deux mémoires de trame qui s'alterne en lecture et en écriture. L'entrelacement des données s'étend à l'ensemble de la trame. Les codes correcteurs d'erreurs externes et internes sont alors générés. Après ajout des données d'identification et de synchronisation, le signal est finalement soumis comme en D3, à un codage de canal 8-14 avant d'être enregistré.

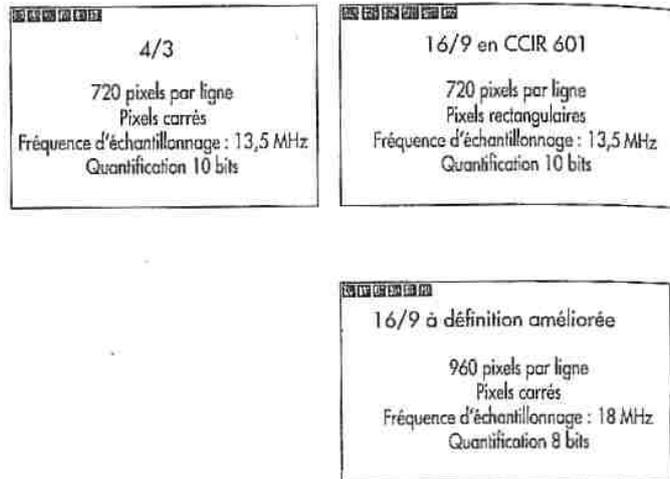
e) D5 HD :

Le D5 est en outre le seul format pouvant enregistrer un signal 16/9 échantillonné à 18 MHz. En termes de débit, la capacité totale du format D5 HD est de 288 Mbits/s dont 234 Mbits/s pour la vidéo en mode 16/9-18 MHz.

➤ Le 16/9 à définition horizontale améliorée :

Le signal numérique 4 :2 :2 peut être utilisé en 16/9 en conservant le même nombre de pixels par ligne qu'en 4/3. Cette solution, qui permet d'utiliser sans modification des équipements 4/3 pour produire en 16/9, présente néanmoins l'inconvénient de rendre rectangulaires des pixels qui sont à l'origine presque carrés. La résolution horizontale de l'image est par conséquent réduite.

Pour remédier à ce problème, il faut non pas « étirer » les pixels pour qu'ils remplissent une ligne plus longue, mais accroître leur nombre de 720 à 960 afin de conserver leur forme carrée. Pour disposer de davantage de points par ligne, il faut augmenter la fréquence d'échantillonnage et passer de 13.5 MHz à 18 MHz. L'échantillonnage ainsi fait permet d'accroître la réponse en fréquence de 5.75 MHz à 7.67 MHz. Le format D5 peut enregistrer un tel signal, avec une quantification sur 8 bits.



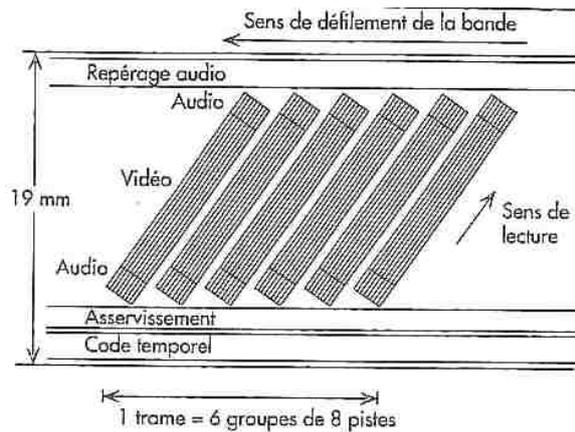
➤ Compression pour la haute définition :

Panasonic propose un compresseur/décompresseur numérique de rapport 4 :1, permettant d'enregistrer, sur une machine D5 conventionnelle, un signal à haute définition (1250 lignes, 1920 pixels), dont le débit vidéo initial de 954 Mbits/s aura été réduit à 238 Mbits/s. L'algorithme de compression, basé sur la DCT, effectue un calcul intra-image pour une précision de montage de l'image. La durée maximale d'enregistrement de 2 heures par cassettes est maintenue en haute définition.

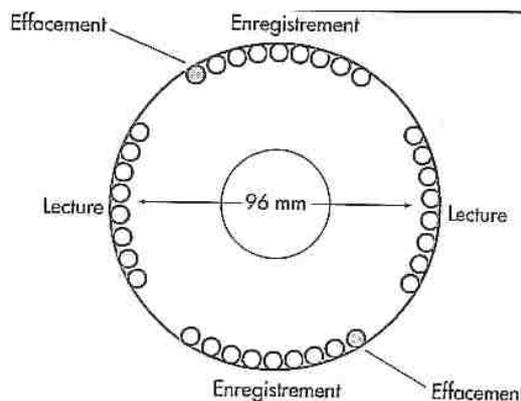
f) **D6 :**

Développé conjointement par BTS et Toshiba, le D6 est un format de haute définition en plein débit, capable de traiter les standards 1250/50 et 1125/60 (balayage entrelacé). Il enregistre sur une bande 3/4" plus d'une heure de vidéo HD sans compression, avec un débit total de 1.2 Gbits/s, dont 921 Mbits/s alloués à la vidéo.

Les données à enregistrer sont réparties sur 8 canaux parallèles, chaque canal gérant ainsi un débit de 150 Mbits/s. La vitesse d'écriture de 46 m/s est obtenue avec une vitesse linéaire de 49.7 cm/s, conjuguée à une vitesse de rotation du tambour de 150 tours par seconde. La bande s'enroule à 180° autour du tambour de diamètre 96 mm. Pas moins de 34 têtes sont requises : 16 pour l'enregistrement, 16 pour la lecture et 2 pour l'effacement rotatif. Selon l'empreinte de la bande, chaque groupe de huit têtes trace autant de pistes jointives, avec un angle d'azimut alterné de + ou - 15°. Une trame se compose de six groupes de huit pistes en 1250/50 et est inscrite en trois tours de tambour. Les groupes de huit pistes sont séparés les uns des autres par une bande de garde de 5 µm. Un groupe est entièrement effacé par une seule tête rotative. Chaque piste hélicoïdale, de largeur 21 µ, contient les données vidéo au centre et les données audio enregistrées en double de part et d'autre. Le D6 permet l'enregistrement de douze pistes audio en 1250/50 et 10 en 1125/60. On retrouve les trois pistes longitudinales utilisées de manière traditionnelle pour le repérage audio, l'asservissement et le time code.



Le traitement du signal s'effectue de manière analogue à celui des autres formats numériques : brassage des données, double code de correction d'erreurs calculé horizontalement et verticalement, puis codage par canal.



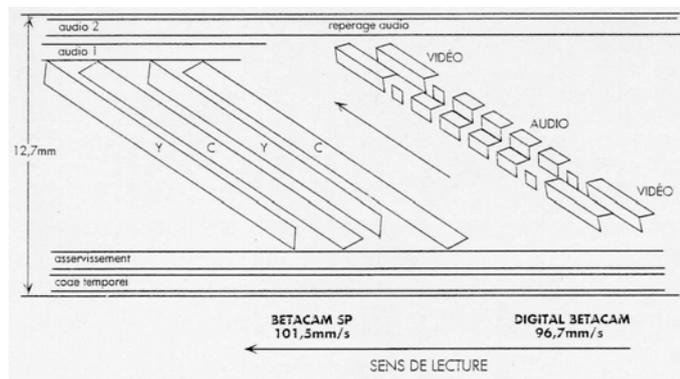
2. La famille BETACAM :

a) **DIGITAL BETACAM : (BROADCAST / POST - PRODUCTION)**

Le Digital Betacam de Sony est une synthèse des technologies développées pour les formats D1, D2 et Betacam SP. Il traite le signal vidéo 4:2:2 sur 10 bits, auquel il fait subir une réduction de débit de type M-JPEG dans un facteur 2. Les quatre pistes audio numériques sont enregistrées sans compression avec un échantillonnage à 48 kHz et un codage sur 20 bits. La gamme de machines Digital Betacam comprend des modèles compatibles en lecture avec les cassettes Betacam analogiques. Le Digital Betacam utilise une bande 1/2" à particules métalliques, optimisée pour l'enregistrement numérique. Deux tailles de cassettes sont disponibles, offrant des durées maximales d'enregistrement de 40 et 124 minutes.

➤ **La structure d'enregistrement :**

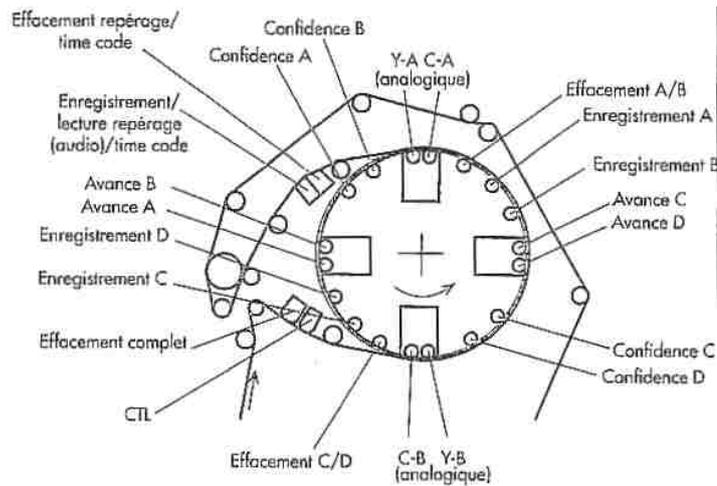
Le signal vidéo est réparti sur quatre canaux. Six pistes hélicoïdales sont nécessaires pour enregistrer une trame. La bande défile légèrement plus lentement qu'en analogique (9.6 cm/s au lieu de 10.1), alors que le tambour tourne trois fois plus vite (75 Hz). Les pistes obliques de largeur 26 μm sont tracées avec un angle d'azimut alterné de + ou - 15°, sans bande de garde entre elles. Une piste se compose d'un premier secteur vidéo, de quatre secteurs audio au centre de la bande et d'un second secteur vidéo. On retrouve les trois pistes longitudinales traditionnelles pour l'asservissement, le time code et le repérage audio. La figure nous montre que la piste audio longitudinale n°1, dite LNG, du format Betacam SP disparaît en numérique pour laisser plus de place aux pistes hélicoïdales. Une machine Digital Betacam compatible peut néanmoins lire cette piste sur une cassette analogique. Tous les paramètres propres à la lecture analogique sont sélectionnés automatiquement dès l'insertion de la cassette. La compatibilité de lecture est assurée grâce à un circuit correcteur de base de temps, qui effectue une légère expansion temporelle des signaux lus.



Le suivi de piste est réalisé par l'association de deux systèmes assurant à la machine un temps de verrouillage particulièrement court. Le premier système exploite de manière conventionnelle la piste CTL longitudinale, tandis que le second utilise deux signaux pilote, l'un de basse fréquence (400 Hz), l'autre de haute fréquence (4 MHz), enregistrés entre les secteurs audio et vidéo.

➤ **Le tambour de têtes :**

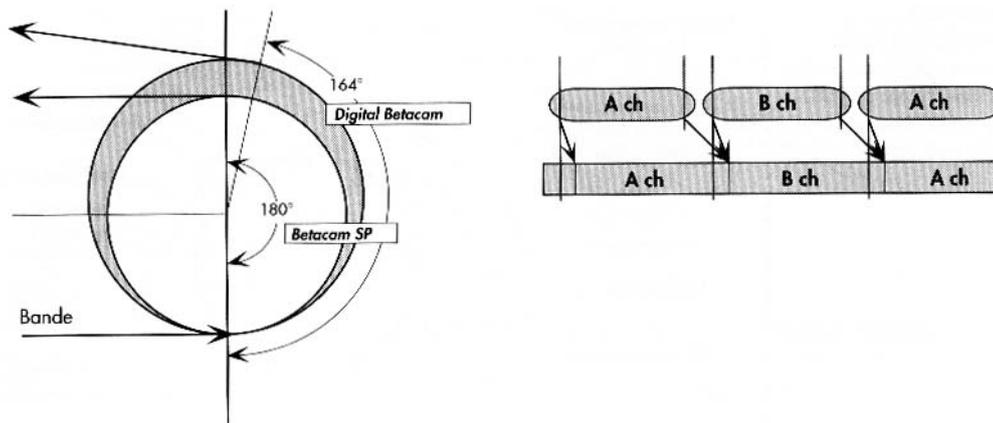
Le diamètre du tambour d'une machine Digital Betacam compatible a été porté à 81 mm contre 74.5 mm en analogique et l'angle d'enroulement de la bande est de 180°. Le tambour ne compte pas moins de 18 têtes, réparties comme suit : 4 têtes pour l'enregistrement, 4 pour la lecture (à Dynamique Tracking), 4 pour la lecture en mode confiance, 2 pour l'effacement et 4 pour la lecture des cassettes analogiques (également à Dynamique Tracking). Le tambour d'une machine non compatible ne possède que 14 têtes. Les têtes de lecture peuvent être utilisées en mode pre-read ; elles tracent alors les pistes quelques trames avant les têtes d'enregistrement. Les têtes de lecture « confiance » ont un retard d'une piste par rapport aux têtes d'enregistrement. Grâce aux têtes de lecture à alignement dynamique DT, il est possible d'effectuer une lecture sans parasite sur une plage de -1* à +3* la vitesse nominale.



➤ **Le traitement du signal :**

L'interface numérique série d'entrée reçoit le signal vidéo en composantes et les quatre canaux audio associés. Le signal analogique et le signal composite (option) sont numérisés sous forme parallèle. La compression numérique est effectuée par un unique circuit intégré, qui effectue toutes les opérations visant à réduire de moitié le débit du signal vidéo. Les codes de correction d'erreurs externe et interne sont associés aux données vidéo et audio. Puis sont ajoutées les données d'identification et de synchronisation. Le codage de canal utilisé pour l'enregistrement est ici le NRZI embrouillé.

A la lecture, les erreurs sont corrigées par les codes interne et externe, tandis que les données irrécupérables sont interpolées par les circuits de dissimulation. Les données en composantes sont converties sous forme série et multiplexées avec les données audio, l'ensemble étant délivré au format de l'interface série à 270 Mbits/s. Précisons que le format Digital Betacam ne supporte que la fréquence d'échantillonnage de 13.5 MHz, ce qui implique une résolution horizontale sensiblement plus faible en 16/9, contrairement au cas du D5.



La différence de diamètre de tambour entre le Betacam SP et le Digital Betacam impose une légère expansion temporelle des signaux analogiques, quand ils sont lus sur une machine numérique.

Tableau 7.5

Comparaison des caractéristiques principales entre le Betacam analogique et le Digital Betacam.

	Betacam SP	Digital Betacam
Enregistrement	Composantes analogiques	Composantes numériques
Nombre de têtes d'enregistrement	4	4
Rotation du tambour	25 Hz	75 Hz
Nombre de pistes par trame	2	6
Vitesse de la bande	101,51 mm/s	96,7 mm/s
Vitesse relative tête/bande	5,75 m/s	19,08 m/s
Diamètre du tambour	74,49 mm	81,4 mm

b) BETACAM SX : (NEWS / MPEG)

Le Betacam SX de Sony est un format numérique destiné à remplacer le Betacam SP. Il repose sur un transport de bande 1/2" et utilise les mêmes cassettes que le Betacam SP avec lequel il est compatible en lecture.

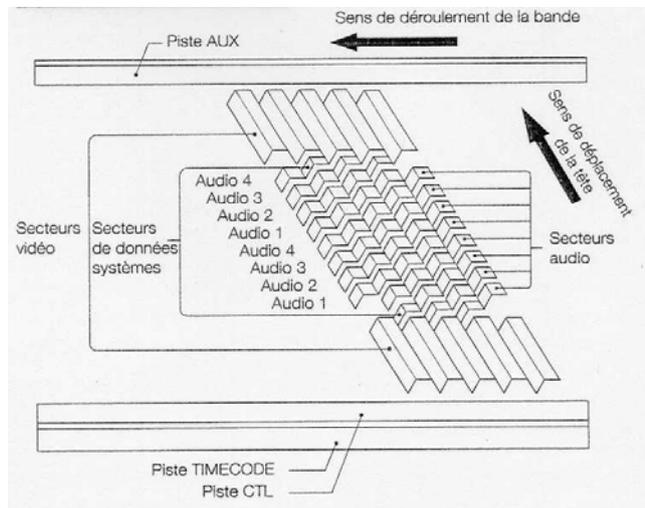
➤ La compression inter image :

L'efficacité de la compression inter image a permis d'atteindre une compression de 10 :1, réduisant à seulement 18 Mbits/s le débit du signal vidéo qui conserve la structure 4 :2 :2. Si l'on ajoute à cela les 4 pistes audio (16 bits, 48 kHz), ainsi que les données supplémentaires de correction d'erreurs, le débit total enregistré s'élève à 40 Mbits/s. L'ensemble des fonctions de codage, estimation de mouvement et décodage, a été implanté sur un jeu de 3 circuits intégrés dont on imagine la complexité et le vitesse de calcul.

➤ **Le format :**

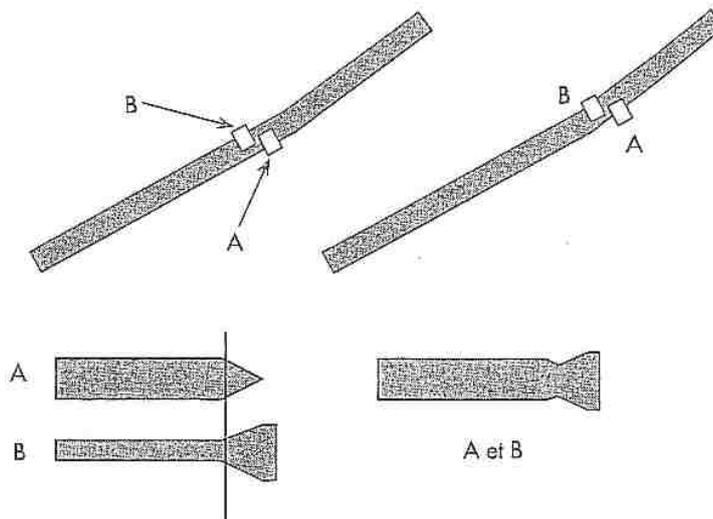
Chaque image est segmentée en 12 pistes dans les systèmes à 625 lignes et en 10 pistes dans les systèmes 525 lignes. La grande efficacité de l'algorithme de compression MPEG-2 422P@ML a permis de réduire quasiment de moitié la vitesse de défilement de la bande par rapport au Digital Betacam et au Betacam SP. Elle n'est ici que de 59.75mm/s. Le tambour effectue 75 tours par seconde, ce qui confère une vitesse d'écriture de 19 m/s. La longueur d'onde minimale enregistrée n'est pas des plus faibles, soit 0.74 μm , la bande est relativement épaisse (14.5 μm) et la largeur des pistes hélicoïdales est de 32 μm .

Tous ces paramètres conjugués à un transport 1/2" largement éprouvé font du Betacam SX un format robuste, capable de s'accommoder d'une formulation de bande assez simple et d'utiliser pour l'enregistrement les cassettes Betacam SP à particules métalliques. Sony propose cependant une nouvelle gamme de cassettes optimisées pour le format SX, se distinguant par leur boîtier jaune vif. Le coût horaire d'enregistrement est réduit de moitié par rapport à l'analogique, puisque la durée d'enregistrement est doublée en SX. Ainsi la capacité maximale est portée à 62 minutes avec une petite cassette et à 184 minutes avec une grande cassette.



➤ **Le système de tracking :**

Le Betacam SX est un format exempt de tout système de suivi de piste précis en lecture. Il utilise en effet la technique « Multiple Head Tracing » employant 2 têtes décalées qui produisent un « faisceau » de lecture beaucoup plus large que la piste parcourue. Toutes les informations de la piste sont donc récupérées, même dans le cas où celle-ci est déformée. De plus, ce système permet une lecture variable sur une plage de + ou - 1, éliminant ainsi la nécessité de recourir à des têtes DT, ce qui réduit encore le coût de fabrication du tambour.



3. HD CAM :

LA SONY HDW-F900 est conforme à la norme internationale de la haute définition avec une résolution de 1920*1080 pixels. Les trois capteurs CCD de 2 200 000 pixels sont au format 2/3 de pouce, imposé par la monture normalisée des optiques pour HD (bague Digital Beta). La caméra peut aussi recevoir des optiques film dans sa version Panavision.

Le signal est traité par un convertisseur 12 bits 74 MHz. Outre la résolution, l'originalité de la HDW-F900 est de recourir soit au balayage entrelacé, soit au balayage progressif. Les cadences de prise de vue sont de 23.976, 24, 25, 29.97, 30 (progressif), 50, 59.94 et 60 image/s (entrelacés). Le mode entrelacé est particulièrement destiné au report final sur pellicule.

1920*1080 pixels * 10 bits * 3 couleurs * 24 images = 1.49 Gbits/s, soit 1 tétraoctet pour un programme d'une heure trente. C'est la problématique de départ pour la diffusion d'un signal de TVHD numérique.

4. La famille DV :

Deux initiales ont été une révolution technologique pour les amateurs et les professionnels. Certes chère pour le grand public à ces débuts, mais plus abordables pour des professionnels confrontés au luxe prohibitif de produits comme le Digital Beta.

➤ 4 : 2 : 0 ; 4 : 2 : 1 ; 4 : 2 : 2 :

Autour d'une compression fixe de 5 : 1, les DV standards à 25 Mbits/s proposent des qualités visuelles très similaires. Le résultat final dépend donc du niveau technologique des composants électroniques, de l'optique et dans une moindre mesure des bandes utilisées. Les hautes technologies mises en œuvre pour ces dernières n'ont pas ouvert le marché de la fabrication aux cloneurs bas de gamme. En Pal, le DV est codé en 4 : 2 : 0 et le DVCPRO 25 en 4 : 1 : 1 - standard qui réduisent de moitié les informations de chrominance. Le haut de gamme vient avec le DVCPRO 50 (pour 50Mbits/s) qui revient au 4 : 2 : 2 (plus adapté aux trucages en numérique) et réduit la compression pour faire grimper la bande passante et la qualité (3.3 : 1).

➤ COMPATIBILITE :

Le numérique n'a pas gommé le vieux clivage Pal / NTSC. Seul le Secam fait les frais de la normalisation du marché. Le DV grand public a fait son apparition via la cassette mini - DV. Le DVCAM emploie des cassettes aux mêmes caractéristiques. Les bandes peuvent parfaitement être échangées dès la prise de vues. La vitesse de déroulement des bandes étant de DVCAM 1.5 fois celle du DV standard, il faut juste se méfier des durées.

Aujourd'hui tous les appareils DV et DVCAM peuvent échanger leurs cassettes en lecture. Au contraire, les caméras DVCPRO ne supportent qu'un type de cassettes.

➤ TOUT DANS LA TETE :

La qualité de la tête de la caméra est l'argument principal. Sony met en avant le traitement totalement numérique de sa DSR - 500, du capteur à l'enregistreur. Certaines caméras offrent un capteur commutable 4/3 - 16/9^{ème} sur sa surface de pixels qui offre l'avantage d'éviter d'utiliser des optiques spécifiques (Sony DSR - 500, Panasonic AJD - 900, ...) . Les nouvelles générations de capteurs sont très performantes en basse lumière et les caméras pros veillent à maintenir cet avantage par rapport aux caméras amateurs. Tous les modèles haut de gamme offrent à leur tête avancés : niveaux de noir, de détails, colorimétrie, tonalité de la peau ...

Le choix d'une caméra bi - bloc, contrairement aux modèles mono - bloc, s'adresse en priorité à ceux qui utilisent leur matériel dans des contextes de productions variés : reportage, fiction, plateau, avec la nécessité de changer les types d'enregistreur. Et en cas d'usure de la partie

magnétoscope ou d'émergence de nouvelles technologies en enregistrement, faire durer la tête est un bon pari.

a) **DV : SEMI - PROFESSIONNEL / GRAND PUBLIC**

Ce format s'appuie sur deux paramètres clés : un algorithme de compression normalisé de type MJPEG de facteur 5 :1 et une nouvelle famille de cassettes compactes utilisant une bande 1/4" (6.35 mm). Le DV offre une qualité d'image assez remarquable. Le DV est un format qui enregistre séparément les composantes de luminance et de différence de couleurs, ce qui, d'une part, garantit une bande passante plus élevée aux signaux et, d'autre part, supprime tout risque d'interférence entre eux.

Le DV est un format totalement numérique et supporte donc les copies directes sans aucune dégradation. Les problèmes de drop out sont définitivement éliminés grâce à la correction d'erreurs. L'audio n'est pas en reste puisque le format autorise au choix l'enregistrement de deux canaux pleine bande à 48 kHz/16 bits, ou quatre canaux à 32 kHz/12 bits.

Panasonic puis Sony ont rapidement pris conscience des atouts exceptionnels du DV pour les applications broadcast. Ils en proposent aujourd'hui trois déclinaisons - DVCAM (Sony), DVCPRO 25, DVCPRO 50 (Panasonic) et Digital S (ou D9 de JVC), toutes basées sur le même moteur de compression, mais bénéficiant d'améliorations plus ou moins significatives.

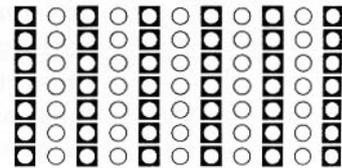
➤ Le signal vidéo du DV :

Le signal vidéo du DV est échantillonné en 4 :1 :1 dans les systèmes 525/60 et en 4 :2 :0 dans les systèmes 625/50. Dans les deux cas le sous échantillonnage des chrominances permet de réduire de 25% le débit du signal vidéo, avant même le processus de compression. Il passe ainsi de 216 Mbits/s (4 :2 :2) à 162 Mbits/s (4 :2 :0 et 4 :1 :1), avec une quantification sur 8 bits. Il sera ensuite réduit dans un facteur 5 :1 par un algorithme de type MJPEG basé sur la Transformée en Cosinus Discrète (DCT). Le débit du signal vidéo DV compressé est de 25 Mbits/s, mais le débit total enregistré, incluant l'audio et les autres données telles que la correction d'erreurs s'élève à 41.85 Mbits/s.

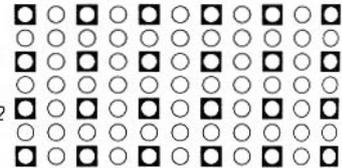
Certes, le DV aurait pu bénéficier d'une compression de type MPEG-2, bien plus efficace que le MJPEG, mais la puissance de calcul nécessaire aurait engendré une trop grande complexité de codage et de décodage.

Les structures d'échantillonnages utilisées par le DV et ses variantes.

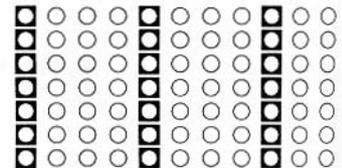
4:2:2
 DVCPRO50
 Luminance : 13,5 MHz, 720 pts/l
 Chrominance : 6,75 MHz, 360 pts/l



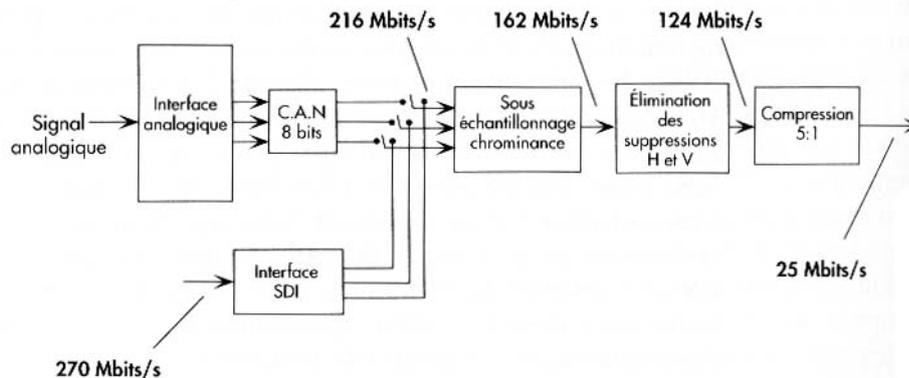
4:2:0
 DV (625/50)
 DVCAM (625/50)
 Luminance : 13,5 MHz, 720 pts/l
 Chrominance : 6,75 MHz, 360 pts/l, 1 ligne sur 2



4:1:1
 DV (525/60)
 DVCAM (525/60)
 DVCPRO25
 Luminance : 13,5 MHz, 720 pts/l
 Chrominance : 3,375 MHz, 180 pts/l



○ échantillon de luminance
 ■ échantillon de chrominance

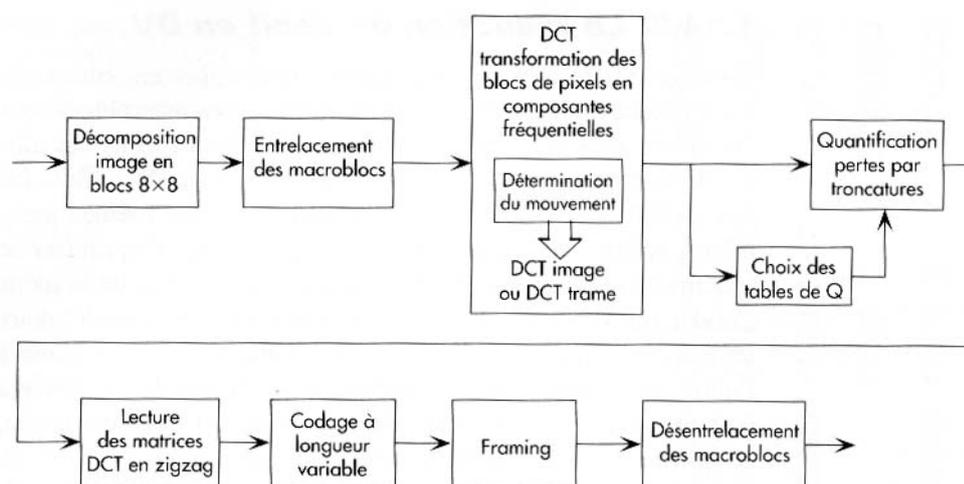


Allègement du débit vidéo avant la compression DV.

La réduction de débit en DV

L'image vidéo est préalablement découpée en blocs de 8x8 pixels, 6 blocs formant un macrobloc. Les macroblocs sont entrelacés, puis assemblés par groupes de 5 pour former des unités de compression (procédé original propre au DV). Tous les blocs d'une unité de compression utiliseront les mêmes paramètres relatifs aux tables de quantification. Pour optimiser le schéma de réduction de débit par rapport au contenu de la scène à coder, une détection de mouvement est effectuée entre les deux trames d'une image. S'il y a peu de mouvement d'une trame à l'autre, la compression est réalisée sur l'ensemble de l'image (sur des blocs de 8x8), en exploitant la redondance inter-trame, ce qui permet une meilleure efficacité de compression. Si, en revanche, le mouvement d'une trame à l'autre est trop prononcé, la compression est effectuée indépendamment sur chaque trame, sur des blocs de 8x4. Mais en aucun cas il n'est procédé à une compensation de mouvement entre images successives. Le codage DV laisse donc chaque image indépendante des autres et se prête sans aucune difficulté au

montage à l'image près. La DCT transforme chaque bloc de 8x8 pixels en une matrice de 8x8 coefficients de composantes fréquentielles afin de mettre en évidence les données redondantes. La quantification non-linéaire divise ces coefficients par des valeurs plus ou moins élevées. Elle tient compte pour cela du fait que l'œil est particulièrement tolérant devant les erreurs sur les hautes fréquences spatiales, mais qu'il est en revanche sensible aux erreurs sur les basses fréquences spatiales. Les coefficients DCT correspondant aux hautes fréquences sont donc codés avec moins de précision que ceux relatifs aux basses fréquences. La matrice obtenue après quantification est soumise à une lecture en zigzag, faisant apparaître de longues suites de zéros. Puis un codage à longueur variable est réalisé, attribuant des codes courts aux symboles les plus fréquents et réservant les codes longs aux symboles plus rares. Les données des macroblocs sont alors mises en paquet dans un bloc de synchronisation de taille fixe. Les macroblocs sont finalement désentrelacés pour retrouver leur position initiale, puis les données d'une image sont ordonnées suivant le numéro de piste auquel elles correspondent. Tout le traitement de compression/décompression DV est implémenté sur un seul circuit intégré, qui bénéficie d'une économie d'échelle du fait de sa production massive. Configurable en 4:1:1 ou en 4:2:0, il est utilisé aussi bien dans les produits DV grand public que dans les équipements broadcast aux formats DVCAM, DVC-PRO et D9.



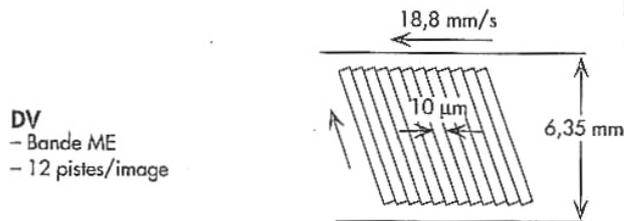
Synoptique de la compression vidéo en DV.

➤ La structure d'enregistrement :

On remarque sur l'empreinte de la bande en DV l'absence des traditionnelles pistes longitudinales d'asservissement, de time code et de repérage audio. Toutes ces informations sont en effet enregistrées sur les pistes obliques par les têtes tournantes. L'asservissement est réalisé à l'aide de deux fréquences pilotes. Les informations de time code, inscrites sur les pistes hélicoïdales, sont lisibles à toutes les vitesses.

Comme pour tous les formats vidéo numériques, le DV fait appel au processus de segmentation, qui consiste à découper une image en plusieurs segments et à enregistrer chaque segment sur une piste. Une image est découpée en 10 segments dans les systèmes à 525 lignes (elle s'étale donc sur 10 pistes) et en 12 segments dans les systèmes à 625 lignes (répartition sur 12 pistes). Dans les deux cas, 300 pistes sont inscrites en une seconde (10 pistes*30 i/s en 525 lignes, et 12 pistes*25 i/s en 625 lignes). Les pistes, dont l'angle d'azimut est de + ou - 20°, sont longues d'environ 33 mm et inclinées de 9° par rapport à l'axe de la bande. Chaque piste hélicoïdale est divisée en 4 secteurs, contenant, dans l'ordre, les données pilotes de suivi de piste ITI (Insert and Tracking Information), l'audio, la vidéo et le time code.

La bande défile très lentement, à moins de 2 cm/s, et s'enroule à 180° autour du tambour de 2.1 cm de diamètre. Le tambour porte deux têtes diamétralement opposées, qui assurent à elles seules l'enregistrement de tous les signaux et effectuent 150 rotations par seconde.



➤ **Les cassettes DV :**

La bande est composée d'une double couche de métal évaporé (cobalt pur) recouverte d'un revêtement protecteur. Deux tailles de cassettes sont disponibles, offrant des durées maximales de 60 minutes pour les caméscopes et de 270 minutes pour les magnétoscope.

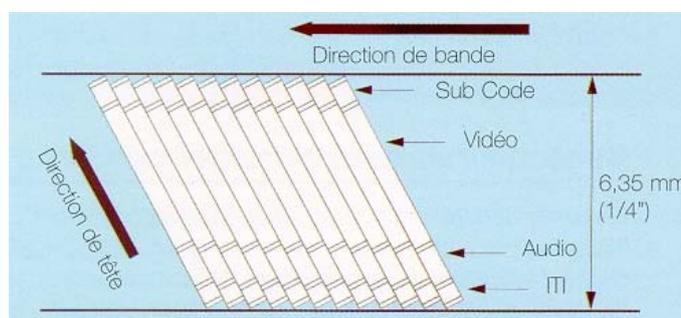
➤ **L'interface IEEE-1394 :**

Développée conjointement par Sony et Apple, l'interface haut débit IEEE-1394 permet de connecter des équipements DV entre eux, mais aussi à des ordinateurs. La liaison comprend trois paires de conducteur en cuivre : une pour les données, une pour les signaux de commande et enfin une pour l'alimentation des périphériques. Cette liaison autorise un débit maximal de 400 Mbits/s, sur une longueur ne pouvant excéder 4.5 m entre deux périphériques.

Il existe, d'après JVC, le format Professionnel DV. C'est une nouvelle ligne de produits DV qui combine l'économie et l'exécution inhérente au DV avec les produits véritablement professionnels et de haute qualité. La qualité d'enregistrement est la même. La différence est dans la technologie d'appareil photo/optique utilisée pour produire les images à enregistrer. Il garde les avantages du DV (compatibilité large, bande faible coût...) et coûte excessivement moins cher que les formats "pros" offerts par d'autres constructeurs.

b) DVCAM : INSTITUTIONNEL (SONY)

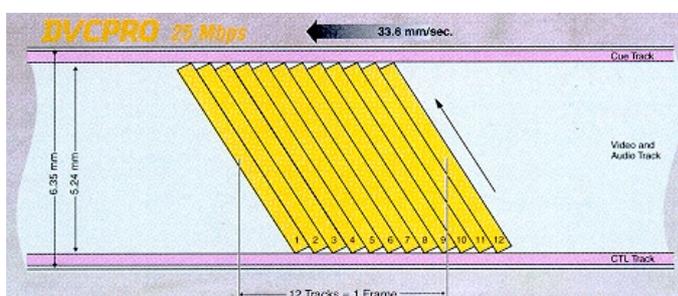
Le format DVCAM est une déclinaison du DV, proposé par Sony. Il utilise les mêmes cassettes, la même bande, la même compression 5 :1, le même échantillonnage 4 :2 :0 et la même structure d'enregistrement que le DV. Le seul point qui différencie en fait le DVCAM du DV est l'accroissement de la vitesse de défilement linéaire de la bande dans un facteur 1.5. La largeur des pistes passe ainsi de 10 à 15 µm, ce qui se traduit par une diminution de la densité d'information sur la bande. Le DVCAM est par conséquent plus robuste que le DV et mieux adapté à une utilisation intensive dans les applications institutionnelles ou lors de reportages d'actualités. La compatibilité DVCAM/DV est totale dans les deux sens en enregistrement comme en lecture. Si un enregistrement DVCAM peut être réalisé sur une cassette DV, Sony propose une famille de cassettes estampillées DVCAM. Elles se distinguent des cassettes DV par la capacité de leur puce mémoire, qui passe de 4 à 16 kbits. En DVCAM, les durées d'enregistrement sont inférieures car la bande défile plus vite, ce qui donne un maximum de 40 minutes sur les petites cassettes et de 184 minutes sur les grandes. Point important : tous les appareils DVCAM acceptent les deux tailles de cassettes DVCAM sans adaptateur.



Avec le format DVCAM est né le concept Clip Link qui permet au cadreur d'introduire lors du tournage, pour chaque plan, un indicateur de validité visant à faciliter le dérushage. Une imagerie d'index (90*60 pixels) représentative du début de chaque plan est captée et sauvegardée temporairement dans une mémoire vive du caméscope DVCAM pendant l'acquisition. Juste avant l'éjection de la cassette, toutes ces images sont copiées sur la bande.

c) DVCPRO : BROADCAST / NEWS (PANASONIC)

Le DVCPRO 25 est une déclinaison vers le haut du DV, proposée par Panasonic. S'il s'appuie sur les spécifications de base du DV, les différences du DVCPRO 25 par rapport à la version grand public sont plus nombreuses que dans le cas du DVCAM. Panasonic a en effet revu à la hausse certains paramètres clés, pour proposer un format doté d'un niveau de fiabilité et de performances élevé à celui de la production broadcast simple. Rien n'a été changé en ce qui concerne le format de la bande et le schéma de compression 5:1, mais des modifications ont été apportées sur la structure d'échantillonnage, la densité d'enregistrement et la formulation de l'enduction magnétique de la bande. Tout cela a été réalisé en maintenant une compatibilité en lecture avec les cassettes DV et DVCAM.



➤ Echantillonnage 4 : 1 : 1 :

Contrairement au DVCAM, qui suit les spécifications DV en matière de procédure d'échantillonnage (4 : 2 : 0 en 625/50 et 4 : 1 : 1 en 525/60), le DVCPRO 25 travaille uniquement en 4 : 1 : 1, quel que soit le standard de balayage. Panasonic justifie son choix en argumentant que le 4 : 1 : 1 supporte mieux la multigénération que le 4 : 2 : 0, la compression de mouvement qui s'effectue sur les deux trames étant plus efficace.

➤ Le retour des pistes longitudinales :

Alors que le DV/DVCAM est exempt de piste longitudinales, le DVCPRO 25 en traite deux, l'une pour l'asservissement, l'autre pour le repérage audio. Cette dernière est également exploitable en tant que troisième voie audio, mais il faut savoir que sa bande passante est limitée à 6 kHz par la faible vitesse de défilement linéaire. Quant au système de suivi de piste, le principe des fréquences pilotes enregistrées sur les pistes obliques n'a pas été retenu pour le DVCPRO 25. Panasonic a préféré se fier à un procédé plus classique reposant sur la traditionnelle piste longitudinale pour garantir des temps de preroll faibles en montage. Les signaux de time code sont enregistrés par les têtes rotatives dans les secteurs dédiés, comme en DV/DVCAM.

➤ La densité d'enregistrement :

La vitesse de défilement de la bande a été presque doublée par rapport au DV, la largeur des pistes passant de 10 μm à 18 μm . La diminution de la densité d'enregistrement est l'un des atouts essentiels du format, même si le prix à payer est une baisse de la durée d'enregistrement.

➤ **La bande et les cassettes :**

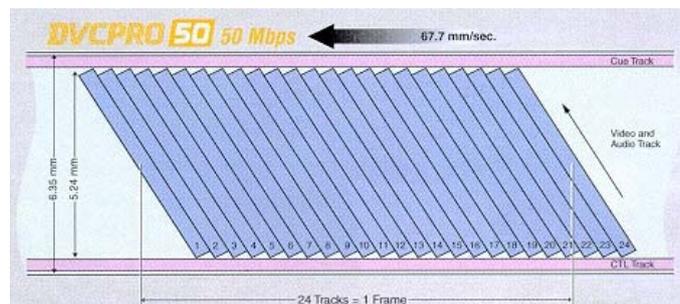
Alors que le DV et le DVCAM utilisent tous les deux les mêmes tailles de cassettes, une grande et une petite, le DVCPRO 25 ne conserve que la grande et en utilise une moyenne qui lui est propre. En raison de l'augmentation de la vitesse de défilement linéaire, la durée d'enregistrement maximale de la grande cassette est de 124 minutes, tandis que celle de la moyenne est de 63 minutes. Les magnétoscopes DVCPRO 25 peuvent cependant relire les petites cassettes DV/DVCAM grâce à un adaptateur mécanique. Mais ils ne peuvent pas enregistrer sur ces petites cassettes.

Quant à la formulation de la bande magnétique, Panasonic a préféré les particules métalliques au métal évaporé du DV/DVCAM, toujours pour des questions de robustesse. Ces particules forment une couche magnétique de seulement $0.2 \mu\text{m}$, soit la plus fine qui n'ait jamais existé sur une bande de type MP. L'épaisseur totale de la bande est supérieure de 25% à celle du DV, soit $8.8 \mu\text{m}$ au lieu de $7 \mu\text{m}$.

d) DVCPRO 50 : BROADCAST / POST - PRODUCTION (PANASONIC)

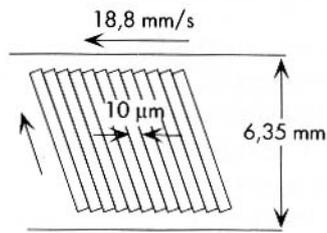
Le DVCPRO 50 est une extension haut de gamme du DVCPRO 25, visant la postproduction broadcast.

Le DVCPRO 50 reprend les caractéristiques initialement définies pour le format D9 de JVC, mais sur bande 1/4". Exit donc le 4 :1 :1, le signal vidéo retrouve en DVCPRO 50 sa structure d'échantillonnage 4 :2 :2 . La résolution en chrominance redevient donc compatible avec les opérations complexes de postproduction, faisant notamment intervenir du chromakey et de la multigénération.

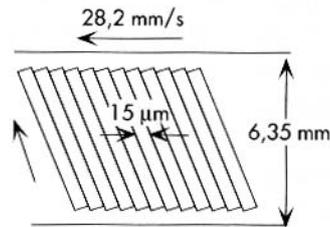


Par ailleurs, le taux de compression a été réduit de 5 :1 en DVCPRO 25 à 3.3 :1, la vitesse de défilement de la bande a été doublée par rapport au DVCPRO 25, passant de 33.8 à 67.6 mm/s, et le débit vidéo est porté de 25 à 50 Mbits/s. La largeur des pistes vidéo reste la même qu'en DVCPRO 25, soit $18 \mu\text{m}$, mais chaque image est segmentée sur 24 pistes au lieu de 12. Ainsi, le DVCPRO 50 est sensiblement plus robuste que le DVCPRO 25 et offre des performances supérieures en lecture à vitesse variable. Des tests d'évaluation effectués révèlent que la qualité d'image délivrée est identique à celle du Digital Betacam en première génération, similaire à la quatrième génération et légèrement inférieure à la septième. Pour compresser un signal 4 :2 :2 dans un facteur 3.3 :1, le DVCPRO 50 utilise en fait deux circuits intégrés DV travaillant en parallèle, chacun sur un signal 2 :1 :1. De ce fait, les équipements DVCPRO 50 sont bistandard : ils peuvent fonctionner indifféremment à 25 ou 50 Mbits/s, en enregistrement comme en lecture, par simple commutation. A 50 Mbits/s, les durées maximales d'enregistrement permises par les deux tailles de cassettes sont de 31 et 93 minutes. Le DVCPRO 50 bénéficie du système d'aide au dérushing Picture Link, qui reprend le principe du Clip Link de Sony (format DVCAM) mais sans utiliser de puce mémoire sur la cassette. Toutes les informations d'index et de validation de plans sont mémorisées pendant le tournage sur une carte PCMCIA logée dans le caméscope, puis sont transférées sur la cassette juste avant son éjection.

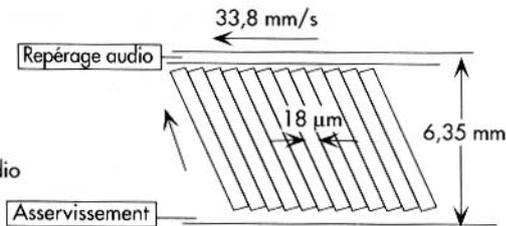
DV
 - Bande ME
 - 12 pistes/image



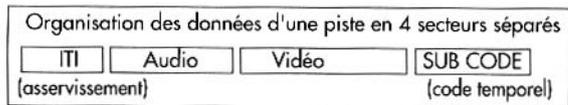
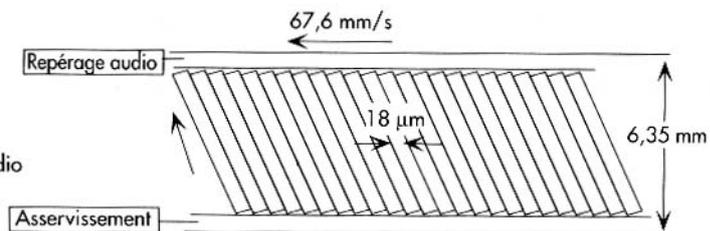
DVCAM
 - Bande ME
 - 12 pistes/image



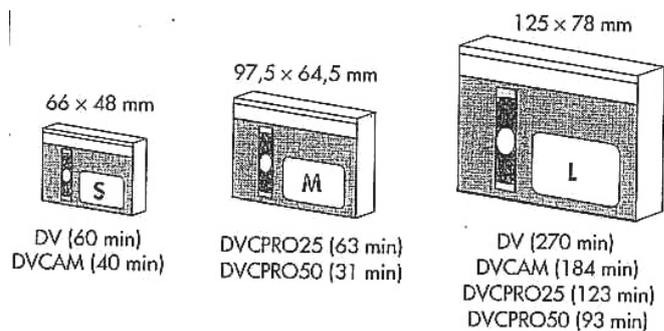
DVCPRO25
 - Bande MP
 - 12 pistes/image
 - Pistes repérage audio
 et asservissement



DVCPRO50
 - Bande MP
 - 24 pistes/image
 - Pistes repérage audio
 et asservissement



Empreinte de la bande en DV, DVCAM, DVCPRO (625/50).

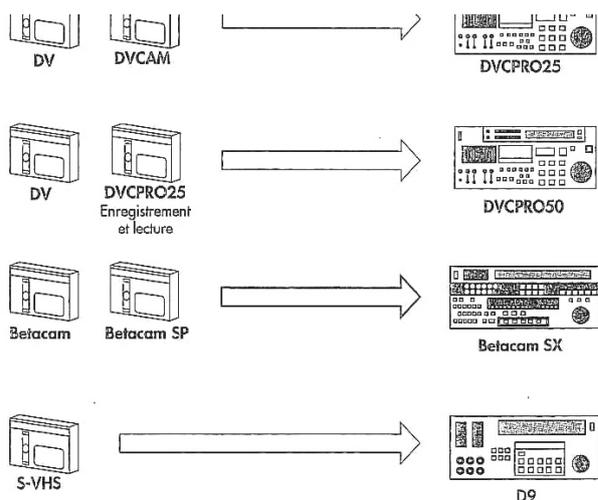
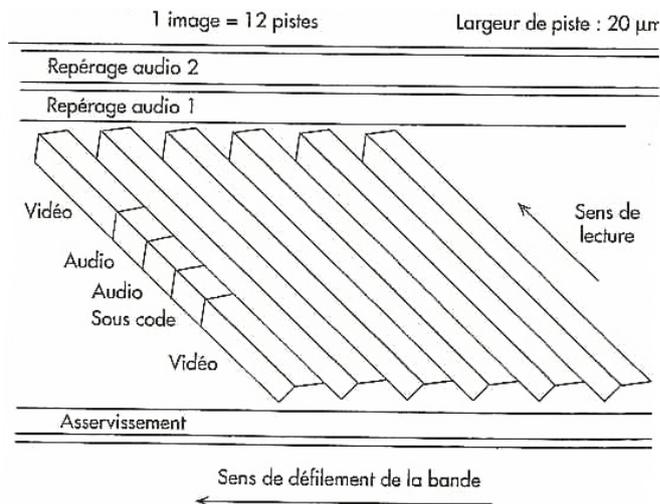


5. Digital S : D9

Le D9 de JVC est un format numérique 1/2" utilisant une réduction de débit très modérée et compatible en lecture avec le S-VHS analogique. Le signal vidéo source 4 : 2 : 2 codé sur 8 bits est soumis à une compression numérique de rapport 3.3 : 1 seulement, abaissant son débit à 50 Mbits/s. L'algorithme de réduction de type M-JPEG est le même que celui du DVCPRO 50 et fait appel à deux circuits de compression DV travaillant chacun sur un signal 2 : 1 : 1. Les 4 canaux audio échantillonnés à 48 kHz et codés sur 16 bits sont, quant à eux, enregistrés à plein débit sur les pistes obliques.

La cassette D9 a les mêmes dimensions qu'une cassette VHS ; une protection empêche toutefois son chargement dans une machine analogique. Bien que les spécifications des deux formats soient radicalement différentes, la mécanique de transport est globalement identique. La vitesse de défilement de la bande, qui est de 23.3 mm/s en S-VHS, est un peu plus de deux fois plus élevée en D9, soit exactement 57.8 mm/s. La vitesse de rotation du tambour de diamètre 62 mm (comme en S-VHS) est triplée : elle passe de 1500 à 4500 tr/min., ce qui porte la vitesse relative tête/bande de 4.84 à 14.5 m/s. Deux paires de têtes diamétralement opposées sur le tambour inscrivent 12 pistes obliques par image en 625/50 (10 en 525/50). Leur largeur est de 20 μm contre 49 μm en S-VHS, et leur enregistrement est effectué avec un angle d'azimut alterné de + ou - 15°. Chaque piste est divisée en 5 secteurs qui portent séparément les informations vidéo, audio et les données auxiliaires. Ces secteurs sont séparés par des intervalles de garde. Le montage à l'image près est ainsi possible en mode d'insertion indépendamment sur la vidéo et sur les 4 voies audio numériques. Des têtes d'effacement rotatives montées sur le tambour effacent tout le signal présent avant d'en enregistrer un nouveau. Trois pistes longitudinales sont utilisées pour enregistrer un signal d'asservissement portant une impulsion par image, ainsi que deux canaux audio optionnels pour le repérage.

Le tambour du D9 comporte des têtes avancées assurant la fonction pre-read. Le signal enregistré sur la bande peut ainsi être utilisé comme source lors d'un montage en insert.



6) Le Digital 8

Lancé en 1999 par Sony, le Digital 8 est un format grand public qui établit une passerelle intéressante entre le DV et le 8 mm/Hi8. Le Digital 8 reprend les spécifications du DV en termes de traitement vidéo et audio numérique, mais utilise comme support d'enregistrement une cassette Hi8 classique. Les deux atouts majeurs de ce nouveau format sont, d'une part, son coût réduit comparé à celui du DV et, d'autre part, sa compatibilité en lecture avec le parc de cassettes analogiques accumulées depuis 1985 (8mm) et 1989 (Hi8).

La structure d'enregistrement

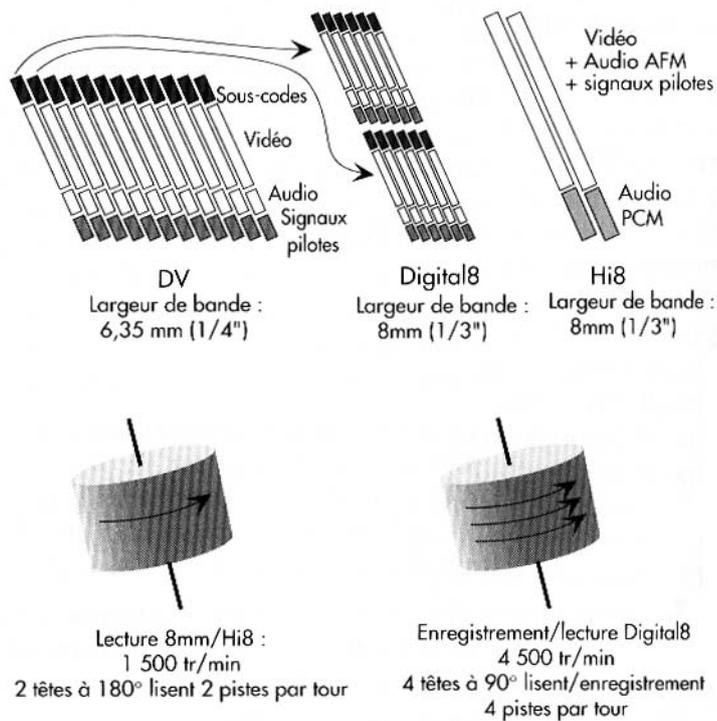
Le Digital 8 enregistre le même signal vidéo 4:2:0 que le DV, c'est à dire avec une fréquence d'échantillonnage de 13.5 MHz pour la luminance et de 6.75 MHz pour la chrominance une ligne sur deux. Une compression de facteur 5:1 réduit le débit vidéo à 25 Mbits/s. Les signaux audio sont enregistrés à plein débit en PCM (modulation par codage d'impulsion), en 16 bits/48 kHz ou en 12 bits/32 kHz. Côté mécanique, le tambour, de diamètre 40 mm (comme en 8mm), porte deux têtes diamétralement opposées pour la lecture analogique et quatre têtes à 90° pour l'enregistrement/lecture numérique. Deux régimes de rotation sont prévus pour le tambour, la sélection de l'un ou de l'autre se faisant automatiquement dès l'insertion de la cassette. En mode analogique, le tambour effectue 1500 tr/min, alors qu'en mode numérique, il tourne trois fois plus rapidement, à 4500 tr/min.

Le Digital 8 code les données identiquement au DV, avec une segmentation de l'image sur 12 pistes, mais utilise une structure d'écriture différente. Il profite en effet de la largeur supérieure de la bande Hi8 (8mm contre 6.35 en DV) pour inscrire les pistes numériques deux par deux, dans le prolongement l'une de l'autre.

Ainsi le système Digital 8 superpose six pistes sur la partie haute de la bande et six pistes sur la partie basse. Les pistes sont plus larges qu'en DV (16.34 μm au lieu de 10 μm), mais plus fines qu'en 8 mm (34.4 μm).

Par ailleurs, la vitesse de défilement de la bande est accrue d'environ 30% par rapport au 8mm/Hi8 (28.7 mm/s au lieu de 20 mm/s). La durée maximale des cassettes est donc réduite dans les mêmes proportions:

Durée mode Hi8:	30	60	90	120 minutes
Durée mode Digital 8:	20	40	60	80 minutes



Signalons que Sony recommande l'utilisation de bandes Hi8 exclusivement pour les travaux numériques, car une bande 8mm enregistrée en Digital 8 peut ne pas être relue correctement sur toutes les machines Digital 8. La compatibilité analogique numérique est totalement transparente pour l'utilisateur.

Les signaux audio/vidéo analogiques enregistrés sur une bande 8mm/Hi8 peuvent être convertis en numérique par un circuit interne et être disponibles sur la sortie DV (IEEE1394).

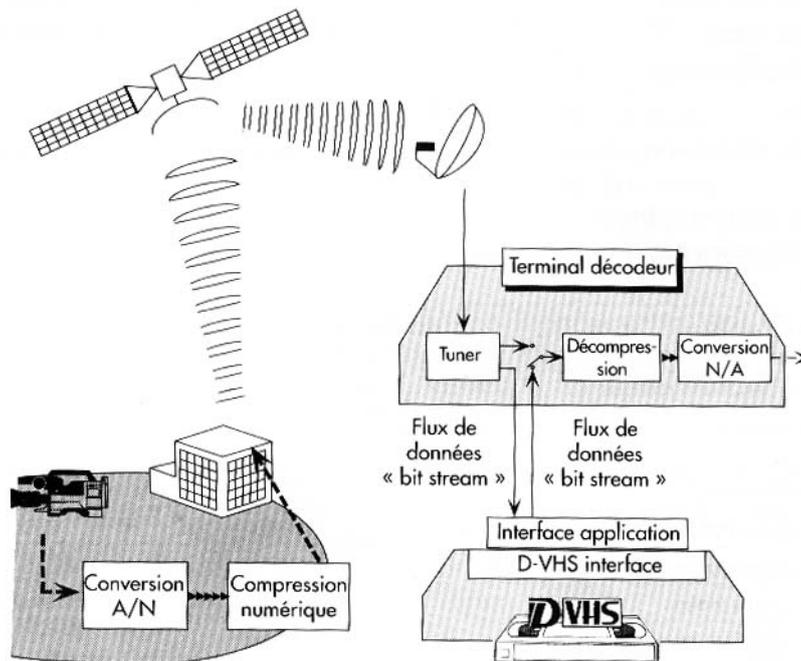
Signalons que trois fonctions offertes par le DV ne sont pas assurées par les équipements Digital 8: le doublage son sur la deuxième voie audio 12 bit/32kHz, le mode longue durée (LP), ainsi que la mémoire sur le boîtier de la cassette DV, facilitant la gestion des séquences vidéo et des photos.

	DV	Digital8	Hi8
Enregistrement luminance	Numérique		Analogique (FM)
Résolution horizontale	500 lignes environ		400 lignes environ
Échantillonnage luminance	13,5 MHz		-
Quantification	8 bits/échantillon		-
Enregistrement chrominance	Composantes numériques		Analogique (procédé under color)
Largeur de la bande chrominance	1,5 MHz environ		0,5 MHz environ
Échantillonnage chrominance	6,75 MHz		-
Débit vidéo	25 Mbits/s		-
Enregistrement audio	Numérique PCM 16 ou 12 bits		Analogique (FM)
Largeur de la bande magnétique	6,35 mm	8 mm	
Dimensions cassette	125 × 78 × 14 mm 66 × 48 × 12 mm	95 × 62 × 15 mm	
Diamètre du tambour	21,7 mm	40 mm	
Vitesse de rotation du tambour	9 000 tr/min	4 500 tr/min	1 500 tr/min
Vitesse de défilement de la bande	18,831 mm/s	28,695 mm/s	20,05 mm/s
Largeur des pistes vidéo	10 µm	16,34 µm	34,4 µm
Nombre de pistes par image	12		
Doublage audio	Oui (13 bits)	Non	Oui (si option PCM)
Mémoire sur cassette	Oui	Non	

7) Le D-VHS

Le D-VHS (Data-VHS) de JVC est un nouveau format d'enregistrement de données numériques 1/2", utilisant la technologie VHS. Sa particularité est d'enregistrer directement les données compressées, comme celles provenant d'une réception TV numérique, sous la forme d'un train d'impulsion binaires (bit stream). Le flux de données, dont le débit peut s'étendre de 2 à 28.2 Mbits/s, est en effet stocké tel quel sur la bande et restitué dans le même état. Autrement dit, le magnétoscope n'effectue aucun traitement de conversion A/N ou N/A, ni de compression/décompression.

Un magnétoscope D-VHS n'est donc pas capable, à lui seul, de reproduire les signaux audio/vidéo à destination du téléviseur. Il nécessite l'utilisation d'un décodeur externe, par exemple le même que celui chargé de traiter une réception numérique par satellite. Le format D-VHS est en effet compatible avec tous les types de diffusion numériques utilisés dans le monde - standard, haute définition, multicanal, etc. Il autorise trois niveaux de qualité d'image: HS (28.2 Mbits/s), STD (14.1 Mbits/s), LS (de 2 à 7 Mbits/s). Le mode HS peut être utilisé pour enregistrer soit un seul programme à haute définition, soit six programmes à définition standard transmis simultanément dans une diffusion multicanal, et relus ensuite un par un. Deux types de cassettes sont proposées, tous deux de la taille d'une cassette VHS. Le premier type de cassette offre une capacité de 31.7 Go, tandis que le second, utilisant une bande plus fine, offre une capacité de 44.4 Go.



Principe de l'enregistrement d'une transmission numérique avec le D-VHS.

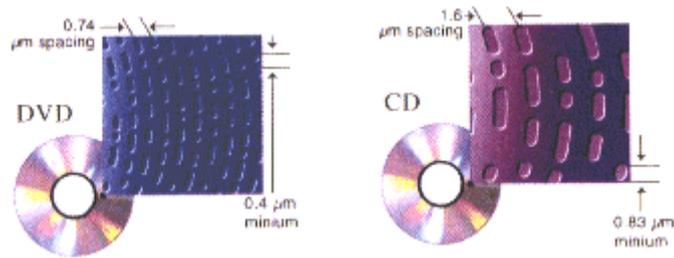
Une cassette D-VHS à bande fine pourra, par exemple, stocker 3.5 heures d'images à haute définition (mode HS), 7 heures d'images à définition standard (mode STD), ou encore 49 heures d'images à définition réduite (modes LS). Parallèlement à l'enregistrement des données numériques, le D-VHS est également capable d'enregistrer un signal analogique conventionnel et présente une compatibilité totale avec le VHS et le S-VHS, dont il reprend la technologie et la mécanique de base. Voilà qui jouera en faveur de sa pénétration et de son développement à grande échelle pour l'utilisation domestique.

Par ailleurs, le D-VHS utilise la même bande à l'oxyde que celle du S-VHS. Le D-VHS tire ainsi pleinement parti de la grande capacité alliée au faible coût de ce type de bande magnétique, et constituera un excellent support de stockage pour les applications multimédia grand public (programmes de télévision, jeux vidéo, données informatiques, DVD, etc.) et professionnelles (surveillance, archivage, etc.).

8) Le DVD

Comme sur un CD, le stockage des informations prend l'alternance des zones plates et d'alvéoles minuscules. Lorsqu'elles sont balayées par le faisceau laser, cela lui permet de reconstituer les données. Sur un DVD, l'espace entre deux pistes n'est que de 0,74mm alors qu'il est de 1,6mm sur un CD. Cela entraîne une plus grande densité d'alvéole pour le DVD. De plus, la longueur minimale d'une alvéole est de 0,4mm (contre 0,83mm pour un CD).

Tout ceci fait donc que le DVD a une capacité d'enregistrement de 6 à 7 CD. Les données sont lues à l'aide d'un faisceau plus fin que pour la lecture d'un CD : 635 ou 650 nanomètre pour le lecteur de CDROM (équivalent à une lumière rouge visible) au lieu de 780 nanomètres (lumière infrarouge).



LE DVD ROM	LE DVD RAM	SIMPLE OU DOUBLE COUCHE
------------	------------	-------------------------

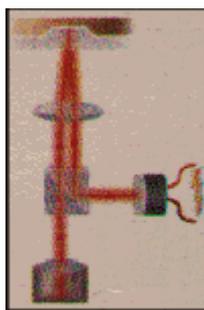
LE PRINCIPE DU DISQUE OPTIQUE :

Selon le type de DVD (enregistrable ou non), la trace physique des informations n'est pas la même : le disque comporte soit des cavités (DVD-ROM) soit des zones claires et foncées (DVD réinscriptibles). Mais le procédé de lecture est le même : une diode photosensible détecte de la lumière quand l'information codée est un zéro, et elle ne détecte rien quand l'information codée est un un.

DVD-ROM

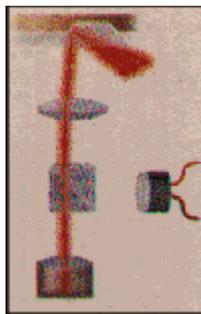
Lecture d'un zéro

Le rayon laser émis par la tête de lecture est réfléchi par le disque il repart en sens inverse vers un Prisme qui le renvoie vers la diode photosensible. Cette dernière détecte une intensité élevée du rayon reçu, ce qui signifie que l'information reçue est un zéro.



Lecture d'un 1.

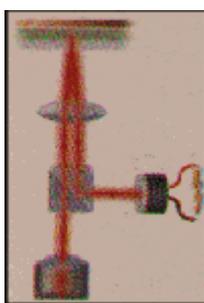
Quand il frappe une zone de passage entre une bosse et une cuvette (ou vice et versa), le rayon laser n'est plus réfléchi vers le bas. Le prisme ne reçoit aucune lumière réfléchi et la diode photosensible enregistre une intensité lumineuse nulle : elle détecte un 1.



DVD RAM

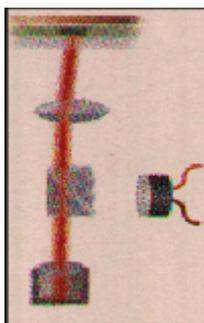
Lecture d'un zéro

Le rayonnement laser est réfléchi par la couche métallique située derrière la couche sensible du DVD. Il repart en sens inverse et le prisme le renvoie jusqu'à la diode photosensible, qui le détecte. Comme avec un DVD-ROM l'information qui est lu est un zéro.



Lecture d'un 1

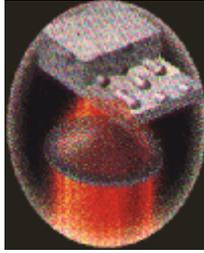
Le rayon laser est absorbé par la couche sensible du DVD-RAM. A cet endroit, en effet, elle a été altérée par une opération d'écriture qui la rends opaque pour le rayon laser. Aucun signal lumineux n'atteint la diode photosensible : l'information lue est un 1.



SIMPLE OU DOUBLE COUCHE, ENREGISTRABLE OU NON

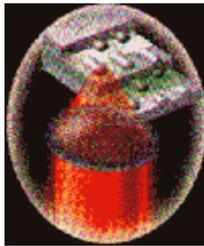
Si le DVD-ROM a une telle capacité, c'est en partie parce qu'il est gravé plus finement que le CD-ROM. Mais aussi parce que deux pistes peuvent être superposées sur chacune des faces. Le laser, focalisé par une pastille mobile, lit à volonté l'une ou l'autre. Le DVD réinscriptible, quant à lui, ne comporte jamais qu'une seule piste par face.

DVD-ROM simple couche



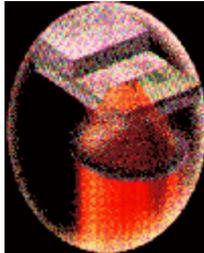
Un DVD-ROM simple couche ressemble beaucoup à un CD-ROM. La différence est la très forte densité des informations. A titre d'indicatif, les cavités d'un DVD-ROM ont une longueur minimale de 0,4 millièrne de millimètre (0.83 pour les cavités d'un CD-ROM).

DVD-ROM double couche



Sur un DVD double couche, les informations sont réparties sur deux niveaux. Le dispositif optique dont est muni le laser focalise sur l'une ou l'autre couche. La surface reflectrice qui recouvre la première couche est translucide.

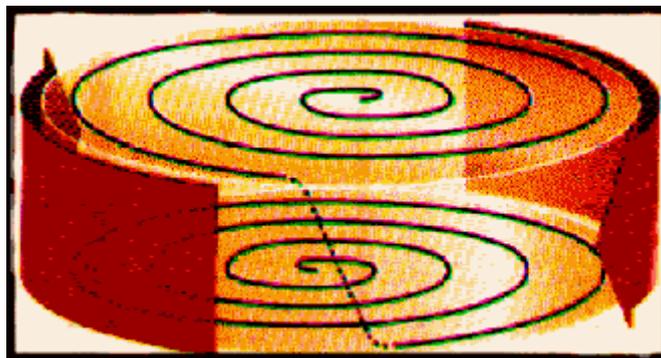
LE DVD ENREGISTRABLE



Le DVD enregistrable (DVD-RAM, RW ou R) a toujours une couche et une seule. Qu'il soit ou non protégé par un coffret, il présente toujours la même structure : une pellicule de métal réfléchissant et une couche de matière photosensible se trouvent prise en sandwich dans du plastique transparent.

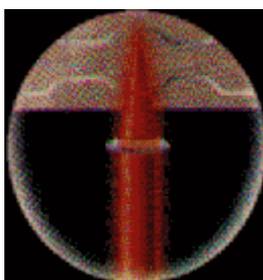
QUAND LES INFORMATIONS SONT EMPILEES SUR DEUX NIVEAUX

Pour que les données soient délivrées de manière ininterrompue, l'information est stockée sur des spirales. En outre, dans les cas des disques double couche, grâce à une astucieuse disposition, la lecture passe instantanément d'une spirale à l'autre.

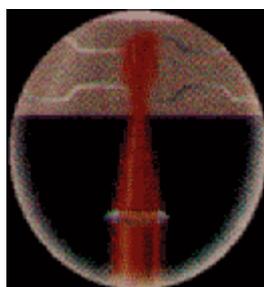


La lecture commence au centre de la première spirale. Quand le laser arrive à la fin, sur le bord du disque, il change de niveau et se met aussitôt à lire la seconde spirale, enroulée dans l'autre sens

LE FAISCEAU LASER EST MIS AU POINT SUR UNE COUCHE OU SUR L'AUTRE



Le faisceau laser est focalisé sur la couche du DVD la plus profonde, qui est recouverte par la pellicule d'aluminium. S'il rencontre une surface plate, il est réfléchi, et le lecteur de DVD détecte un zéro.



Le faisceau du laser est focalisé sur la pellicule semi-réfléchissante qui recouvre la couche superficielle du DVD. S'il rencontre une surface plate, la lumière réfléchi est détectée par la diode photosensible de la tête de lecture.

9) Le SuperVCD

Le monde des disques vidéos se satisfaisait jusqu'en 1999 du DVD et du VidéoCD. Puis, de Chine nous est arrivé en 1999 le SuperVCD

Le SuperVCD est pour simplifier l'évolution du vidéoCD et il vise à offrir une meilleur qualité vidéo que le VCD, sans rivaliser avec le DVD tout en s'appuyant sur des technologies peu couteuses (technologie CD-ROM standart, MPEG-2).

Un disque SVCD est un disque qui comporte.

9) Le MP3

Le MP3 est un format de fichier audio qui compresse les fichiers sonores traditionnels (WAV) jusqu'à 1/12 de leur taille originale en supprimant les données inaudibles pour l'oreille humaine: les ultrasons, les infrasons et les sons d'instruments inaudibles parce que joués en même temps que d'autres sons qui les masquent.

Résultat: la perte de qualité des MP3 est minime par rapport aux CD. D'où le succès de ce format audio auprès des mélomanes.

Les avantages du MP3:

- La compression énorme des chansons, ce qui facilite le stockage sur support informatique et diminue les temps de transfert sur Internet. Comptez tout de même entre 10 et 15 minutes en moyenne pour télécharger une chanson avec un modem de 56 kbps (bits par seconde). Pas de quoi grimper au rideau donc mais on rentre dans les limites du tolérable.
- L'absence de dégradation lors de la copie. Grâce à la technologie numérique, on peut reproduire à l'infini un fichier MP3, sans perte de qualité.
- Le confort d'utilisation. Grâce au MP3, vous pouvez créer vos compilations personnelles directement sur votre disque dur. Cela vous évite de devoir changer de CD à tout bout de champ sur votre ordinateur.

Les inconvénients du MP3:

- Le manque de portabilité. En théorie, vous pouvez écouter la musique en MP3 sur votre ordinateur, sur votre baladeur numérique, bientôt sur votre GSM et votre autoradio. En prévision pour 2005: un implant MP3 directement relié à votre réseau de neurones :-)

En pratique, il faut bien reconnaître qu'actuellement, peu de personnes ont la possibilité d'écouter des MP3 ailleurs que sur leur micro-ordinateur. Mais, les baladeurs numériques envahissent progressivement les boutiques et deviennent plus abordables financièrement, ce qui devrait accentuer le développement de la musique numérique.

- L'incompatibilité avec les lecteurs CD traditionnels: le format MP3 n'est pas reconnu par les lecteurs de CD audio. Vous ne pouvez donc pas (encore) écouter votre CD gravé avec plein de MP3 sur votre chaîne hi-fi.
- **Un MP3 n'a la même qualité sonore qu'un CD**
Même si l'on n'a pas l'oreille musicale, on peut déceler une (légère) différence de qualité entre un CD et un titre encodé en MP3 (taux de compression: 128 Kbs). Mais cette perte de qualité n'est pas suffisante que pour incommoder l'écoute d'une chanson, du moins pour la majorité des internautes.
- **On ne peut trouver pas n'importe quelle chanson en MP3.**
De nombreuses chansons n'ont pas encore été numérisées. On songe en particulier aux chants folkloriques et traditionnels pour lesquels il n'existe souvent pas de version digitale ni même de version analogique enregistrée sur vinyl ou cassette audio. Par ailleurs, certains artistes ont enregistré une version numérique de leur musique sous d'autres formats que le MP3, par exemple en Realaudio ou en Wav.

- **Le MP3 n'est pas le premier format audio développé pour les ordinateurs**
Les passionnés d'informatique et mélomanes avertis n'ont pas attendu le MP3 ni l'essor d'Internet auprès du grand public pour associer musique et ordinateur. Le format Midi par exemple existe depuis des temps immémoriaux (à l'échelle d'Internet, bien sûr). Alors, pourquoi le MP3 fait-il autant parler de lui ? Tout simplement en raison de son taux de compression exceptionnel. Grâce au MP3, on peut compresser très fortement des chansons (de 3 à 4 Mo en moyenne par chanson), ce qui permet de les transférer rapidement via Internet. D'où un gain considérable en temps.
- **Il est interdit diffuser des MP3 sur Internet sans l'autorisation de l'auteur. Même si l'on indique sur son site web une mention du genre "vous ne pouvez télécharger ces MP3 qu'à condition de posséder les disques originaux", on se ne protège pas contre les poursuites éventuelles devant les tribunaux.**
Ce message de mise en garde ne dispense pas le webmestre de demander l'autorisation des titulaires des droits pour diffuser leurs chansons sur son site.
- Il faut obtenir l'autorisation de l'auteur, quel que soit le format audio de la musique que l'on souhaite reproduire et diffuser en ligne et cela, même si l'on ne poursuit aucun but commercial.

10) Le DAB

Le DAB (Digital Audio Broadcasting) est la nouvelle norme numérique de radiodiffusion. La diffusion en DAB procure pour l'utilisateur de multiples avantages; le son, codé numériquement, est de qualité identique à celle d'un CD. La réception est parfaite quelle que soit la vitesse de déplacement de l'auditeur.

Les nouveaux appareils de réception possèdent un écran à cristaux liquides en plus du récepteur "classique", ce qui permet la visualisation de PAD (Program Associated Data / données associées).

Ces données associées, envoyées elles aussi en numérique, peuvent être du texte, des images ou des photos. Enfin, on peut adjoindre aux programmes radios diffusés sous forme de bouquet numérique, des NPAD (Non Program Associated Data / données auxiliaires) affectées exclusivement à la diffusion de données numériques dynamiques.

Outre la transmission par voie hertzienne des sons avec une qualité inégalée, le DAB permet donc la diffusion de textes, de données et d'images, et ce, à l'échelle d'une ville ou d'une agglomération. Les protocoles et les normes (MUSICAM) utilisés dans les techniques de compression des données et dans leur transmission sont standardisés facilitant l'interconnexion avec des réseaux existants tels qu'Internet et, par extension, tout réseau informatique privé. Enfin, le fonctionnement de la diffusion sous forme de bouquet numérique, à l'instar des opérateurs satellites, permet de diffuser plusieurs radios sur un multiplex DAB pour une ville donnée.

11) Le Div-X

Tout a débuté par la démocratisation des connections Internet à haute débit : **ADSL**, câble, etc. La vidéo sur les réseaux n'était plus un rêve mais une possibilité. Il ne restait plus qu'à trouver une norme de compression vidéo : le **MPEG-4** fut créé, et très vite, le **DivX;-)** en est né. Malheureusement, ce fabuleux progrès se limita au monde du piratage, car le **DivX;-)** n'était pas une création acharnée de quelques doués, mais simplement la modification d'un codec* de **Microsoft**. Tout emploi du **DivX;-)** était illégal, mais cette information ne dérangerait pas grand monde, car l'utilisation première du **DivX;-)** était la

conversion d'un film **DVD** en un film **AVI** pouvant tenir sur un ou deux **CD-R**. La différence entre le **DVD** et les **DivX;-)** était bien sûr frappante, mais la qualité était au rendez-vous.

A l'inverse du **MP3**, aucune utilisation commerciale ne pouvait être faite du **DivX;-)**. Qui se risquerait à développer un produit autour d'un codec Microsoft piraté ? Son point faible apparut à la face des utilisateurs : on n'aura jamais de lecteurs **DVD** compatible **DivX;-)**. Son utilisation ne quittera donc jamais celle du **PC**. Pourtant, l'espoir persistait. A la naissance du **DivX;-)**, le **Projet Mayo** (*pour mayonnaise*) démarra. *Jérôme Rota*, l'un des deux créateurs du **DivX;-)**, et l'initiateur du **Projet Mayo** rêva d'une technique de compression similaire qui n'était pas liée à la technologie de **Microsoft**. Le développement de l'**OpenDivX** commença. Les premiers travaux furent diffusés sur le net, et le résultat horrifiait tous les internautes. Il n'y a plus aucun espoir...

Finalement, l'an 2001 arriva. On n'y a jamais cru et pourtant, ils l'ont fait ! Le **DivX 4** apparut, la société **DivXNetworks** fut mise sur pied, et deux copyright furent créés : le **DivX™** et l'**OpenDivX™**. Le nouveau **DivX** ne connut pas seulement un changement de nom, il devint complètement légal, et libre d'utilisation. Actuellement, le **DivX 4** s'impose comme le codec ultime pour la diffusion de vidéos de hautes qualités via Internet. Toujours en phase bêta, il dépasse déjà de loin son petit frère, le codec hacké de **Microsoft**.

Maintenant l'espoir revient ! **DivXNetworks** s'est lancé dans la commercialisation de vidéo **DivX** par Internet, d'autres sociétés ont confiance en l'avenir du **DivX**, comme **Broadway Online** qui propose des films en **DivX**. Qui se lancera dans la création d'un lecteur **DVD** compatible **DivX** ? Il paraîtrait qu'en Asie, certains constructeurs y songent sérieusement. Affaire à suivre !

Le DivX sur la télé ?

Il est tout à fait possible de lire des **DivX** sur la télévision, mais il est impossible de visionner un film sans l'aide d'un ordinateur. Pour l'instant, seuls les ordinateurs savent décoder le **DivX**, c'est donc le **PC** qui enverra la vidéo sur la télévision à l'aide d'une sortie télé. La majorité des cartes vidéo propose ce supplément, pensez-y à l'achat.

La plupart du temps, les sorties télé sont limitées (sauf les cartes **Matrox**) à une résolution maximum de 800 pixels sur 600 ; Le passage du signal par une prise péritel diminue sensiblement la qualité de la vidéo. Si la vidéo a été compressée correctement, la qualité du **DivX** sur la télévision dépassera de loin tout résultat avec une cassette vidéo. Par contre, on ne bénéficiera pas de la superbe qualité donnée par un écran d'ordinateur. A deux mètres de distance, on ne remarquera plus cette impression numérique sur la télévision. La qualité sera un peu près la même que se soit sur l'ordinateur ou sur la télévision, et rien ne vaut le confort d'un fauteuil et d'une télévision.

Et rappelons qu'il est, pour l'instant, impossible de lire un **DivX** avec un lecteur **DVD**. Le seul moyen est de convertir votre vidéo en **MPEG-1** ou en **MPEG-2**, mais alors ce n'est plus du **DivX**.

Quelques notions à connaître :

1°) Les trois divx :

- **Divx** : pour **Digital Video Express**. C'est un système de disque **DVD-Video** destiné à la location. Le problème c'est qu'il fallait un lecteur **DVD** spécifique. Le projet a été abandonné. Il n'y a aucun rapport avec les codecs **DivX**.
- **DivX;-)** : le codec piraté de Microsoft. Le smiley en forme de clin d'œil est un hommage au défunt format de **DVD** locatif qui avait été très combattu (**Fight Divx** !) lors de son

lancement. Ce codec est divisé en deux : le **fast motion**, pour les scènes d'action, le **low motion**, pour les films plus lents.

- **DivX™** : plus connu sous le nom de **DivX 4**. C'est un codec légal, étranger à tout programme **Microsoft**. Le plus performant pour l'instant.

2°) Les autres noms du DivX;-) :

- **OpenDivX™** : le nom de code du développement du **DivX**. Pour rappel, le **DivX** est toujours en phase bêta.
- **SBC** : une méthode de compression vidéo en deux temps, basé sur le **DivX;-) 3.11 alpha**

	Type de bande magnétique	Largeur des pistes	Résolution du signal de luminance	Type d'échantillonnage vidéo	Taux de compression	Audio	Débit Vidéo enregistré sur la bande	Durée d'enregistrement
DV	Métal évaporé	10 µm	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques (8 bits) 4:2:0 en 625 L 4:1:1 en 525 L	5:1 intra image	2 canaux 16 bits/48 kHz ou 4 canaux 16 bits/32 kHz	25 Mbits/s	270 mn
DVCam	Métal évaporé	15 µm	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques (8 bits) 4:2:0 en 625 L 4:1:1 en 525 L	5:1 intra image	2 canaux 16 bits/48 kHz ou 4 canaux 16 bits/32 kHz	25 Mbits/s	184 mn
DVCPro 25	Métal particule	18 µm	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques (8 bits) 4:1:1	5:1 intra image	2 canaux 16 bits/48 kHz	25 Mbits/s	184 mn
DVCPro 50	Métal particule	18 µm	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques (8 bits) 4:2:2	3,3:1 intra image	4 canaux 16 bits/48 kHz	50 Mbits/s	93 mn

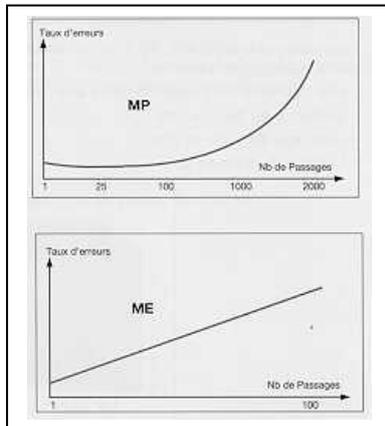
Magnétoscope	Type de cassette	Type de bande magnétique	Année	Résolution du signal de luminance	Type d'échantillonnage vidéo	Quantification	Taux de compression	Débit vidéo enregistré sur la bande	Audio	Temps d'enregistrement
D1	D1	¾" oxyde	1987	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes Numériques 4:2:2	8 bits	sans compression	166 Mbits/s	4 canaux 16/20 bits 48 kHz	94 mn avec la cassette L
D2	D2	¾" métal particule	1988	Liée au signal composite	Composite Numérique PAL ou NTSC	8 bits	sans compression	142 Mbits/s en PAL	4 canaux 20 bits 48 kHz	208 mn avec la cassette L
D3	D3/D5	½" métal particule	1991	Liée au signal composite	Composite Numérique PAL ou NTSC	8 bits	sans compression	142 Mbits/s en PAL	4 canaux 20 bits 48 kHz	245 mn avec la cassette XL
D5	D3/D5	½" métal particule	1994	720 x 576 (en 625 L) 720 x 480 (en 525 L) 960 x 576 (en 625 L 16/9) 960 x 480 (en 525 L 16/9)	Composantes numériques 4:2:2	10 bits en 4/3 8 bits en 16/9	sans compression	233 Mbits/s	4 canaux 20 bits/48 kHz	124 mn avec la cassette L
D5 HD	D3/D5	½" métal particule	1998	1920 x 1080 à diverses cadences images, progressif et entrelacé 1280 x 720P	Composantes numériques 4:2:2 et Haute définition 22:11:11	8 bits en HD	4:1	233 Mbits/s	4 canaux 20 bits/48 kHz	124 mn avec la cassette L
D6 VooDoo	D6	¾" métal particule	1995	1920 x 1080 à diverse cadences images, progressif et entrelacé	Composantes numériques Haute définition 22:11:11	10 bits pour Y 8 bits pour U et V	sans compression	1 Gbits/s	10 ou 12 canaux 20 bits/48 kHz	64 mn
Digital Betacam	Betacam	½" métal particule	1993	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques 4:2:2	10 bits	environ 2:1 intra image	126 Mbits/s	4 canaux 20 bits/48 kHz	124 mn
Betacam SX	Betacam	½" métal particule	1996	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques 4:2:2	8 bits	10:1 sur des groupes de 2 images	18 Mbits/s	4 canaux 16 bits/48 kHz	194 mn
HDCam	Betacam	½" métal particule	1998	1920 x 1080 à diverses cadences images, progressif et entrelacé, ramené à 1440 x 1080 par filtrage numérique	Composantes numériques Haute définition 22:11:11	8 bits	5:1 après filtrage numérique	126 Mbits/s	4 canaux 20 bits/48 kHz	124 mn
Digital S (D9)	VHS modifiée	½" métal particule	1995	720 x 576 (en 625 L) 720 x 480 (en 525 L)	Composantes numériques 4:2:2	8 bits	3,3:1 intra image	50 Mbits/s	2 ou 4 canaux 16 bits/48 kHz	124 mn

Caractéristiques principales des nouveaux formats numériques en 625/50.

	DV	DVCAM	DVCPRO25	DVCPRO50	D9	Betacam SX	MPEG-IMX
Constructeurs	Tous	Sony, Ikegami	Panasonic, Philips, Hitachi, Ikegami...	Panasonic, Philips, Hitachi, Ikegami...	JVC	Sony	Sony
Largeur de bande	1/4"	1/4"	1/4"	1/4"	1/2"	1/2"	1/2"
Structure d'échantillonnage	4:2:0	4:2:0	4:1:1	4:2:2	4:2:2	4:2:0	4:2:2
Résolution	8 bits	8 bits	8 bits	8 bits	8 bits	8 bits	8 bits
Compression	5:1	5:1	5:1	3,3:1	3,3:1	10:1	3,3:1
Débit vidéo	25 Mbits/s	25 Mbits/s	25 Mbits/s	50 Mbits/s	50 Mbits/s	18 Mbits/s	50 Mbits/s
Débit total	41,8 Mbits/s	41,8 Mbits/s	41,8 Mbits/s	80,9 Mbits/s	99 Mbits/s	40 Mbits/s	88 Mbits/s
Diamètre tambour	21,7 mm	21,7 mm	21,7 mm	21,7 mm	62 mm	81,4 mm	81,4 mm
Rotation tambour	150 t/s	150 t/s	150 t/s	150 t/s	75 t/s	75 t/s	50 t/s
Type de bande	Métal Évaporé	Métal Évaporé	Métal	Métal	Métal	Métal	Métal
Vitesse écriture	10,16 m/s	10,1 m/s	10,1 m/s	11,4 m/s	14,5 m/s	18,9 m/s	12,7 m/s
Vitesse bande	18,8 mm/s	28,2 mm/s	33,8 mm/s	67,6 mm/s	57,8 mm/s	59,6 mm/s	53,9 mm/s
Largeur piste	10 µm	15 µm	18 µm	18 µm	20 µm	32 µm	21,7 µm
Pistes audio	2 (48 kHz, 16 bits) ou 4 (32 kHz, 12 bits)	2 (48 kHz, 16 bits) ou 4 (32 kHz, 12 bits)	2 (48 kHz, 16 bits)	4 (48 kHz, 16 bits)	4 (48 kHz, 16 bits)	4 (48 kHz, 16 bits)	4 (48 kHz, 24 bits) ou 8 (48 kHz, 16 bits)
Preread (sur certaines machines)	non	oui	non	non	oui	oui	oui
Durée maximale des cassettes	S : 60 min L : 270 min	S : 40 min L : 184 min	M : 63 min L : 184 min	M : 31 min L : 93 min	124 min	S : 62 min L : 194 min	S : 72 min L : 220 min
Compatibilité (sur certaines machines)	• Lisible par DVCAM, DVCPRO25 et DVCPRO50	• Lit DV et DVCPRO25, • Lisible par DV, DVCPRO25 et DVCPRO50	• Lit DV et DVCAM, • Lisible par DVCPRO50 et DVCAM	• Enregistre et lit DVCPRO25 • Lit DV et DVCAM	• Lit S-VHS	• Lit Betacam et Betacam SP	• Lit Betacam, Betacam SP, Betacam SX, Digital Betacam.
Particularités	• Pas de piste linéaire d'asservissement, ni de repérage (fonctions réalisées par les pistes obliques). • Cassettes équipées d'une puce mémoire de 4 kbits stockant un catalogue des images enregistrées.	• Transfert 4x • Pas de piste linéaire d'asservissement, ni de repérage (comme en DV). • Système ClipLink permettant d'utiliser au montage des marqueurs introduits lors du tournage. • Cassettes équipées d'une puce de 16kbits	• Transfert 4x	• Transfert 2x • Mêmes spécifications techniques que le Digital-S de JVC, mais aucune compatibilité mécanique.	• Transfert 2x • Même mécanique de transport que le S-VHS, même diamètre de tambour mais tournant 3x plus vite.	• Transfert à 4x • Compression MPEG-2 422P@ML avec GOP=2 images • Format « non tracking », exempt de circuit de suivi piste. • Durées d'enregistrement 2x plus élevées qu'en Betacam analogique.	• Transfert à 2x • Compression MPEG-2 422P@ML tout intra

LEXIQUE

BANDE METAL PARTICULE ET BANDE METAL EVAPORE :



Résistance et fiabilité sont deux éléments importants dans le choix d'un format numérique professionnel. La bande Métal Particule du format DVCPRO possède une résistance à l'étirement du double de celle de la bande Métal Evaporé.

Aujourd'hui tous les formats Broadcast utilisent la bande MP. Celle-ci s'optimise au cours des premiers passages, alors que la bande ME, se dégrade très rapidement au fur et à mesure des générations, pour finalement générer des drops de lecture dus aux arrachements de particules sur le support.

CCD (COUPLED CHARGE DEVICE) :

Un capteur CCD (dispositif à couplage de charge) est un circuit intégré renfermant plusieurs centaines de milliers de cellules photosensibles. Chaque cellule accumule une quantité d'électrons proportionnelle à l'éclairement (photon) auquel elle est soumise pendant le temps d'intégration.

Il existe plusieurs structures de capteurs CCD qui se distinguent par la façon dont les charges sont transférées vers le registre de sortie :

- CCD IT : (CCD à Transfert Interligne)
Les charges accumulées sont transférées colonne par colonne dans un registre vertical qui les fait descendre, ligne par ligne, vers un registre horizontal de sortie. Le CCD IT a un gros défaut : le smear.
- CCD FT : (CCD à Transfert d'Image)
Les charges accumulées sont simultanément transmises dans une zone mémoire de capacité équivalente à la zone image. Ce mouvement est masqué par un obturateur mécanique qui permet une totale immunité au smear.
L'avantage de la structure FT est que toute la surface de la zone image est occupée par des cellules photosensibles alors que la structure IT alterne cellule photosensibles et registre à décalage vertical.
- CCD FIT : (CCD à Transfert d'Image Interligne)
Les charges intégrées sont d'abord transférées colonne par colonne dans des registres verticaux (IT) , puis dans une zone mémoire (FT). Le défaut de smear est minimisé en FIT par rapport au IT grâce à la zone mémoire, mais n'est pas nul à cause de l'absence de l'obturateur mécanique.
- CCD HAD : (Hole Accumulated Diode)
La cellule HAD est plus petite que la cellule MOS (Metal Oxyde Semi conductor) ce qui permet d'en loger plus dans la zone image du capteur. Sa surface photosensible est plus grande. L'image est formée avec un nombre de pixel plus important et l'espace entre les pixels est réduit. De plus, le défaut de smear est minimisé car les charges en surplus sont attirées par un drain d'évacuation.

- **CCD HYPER HAD :**

La cellule reprend la structure de la cellule HAD avec en plus une micro lentille qui focalise deux fois plus de lumière sur chaque capteur . La sensibilité du capteur est doublée par rapport à la structure HAD.

COMPOSANTE :

Système vidéo qui convoie les trois couleurs du signal vidéo indépendamment limitant ainsi les interférences, exemple : RVB (existence encore de diaphonie).

COMPOSITE :

La couleur est codée en un seul signal vidéo. La chrominance et la luminance sont entrelacées (exemple : Pal, NTSC).

COMPRESSION :

C'est l'application d'un algorithme ayant pour but de réduire la taille d'un signal ou la bande passante d'un signal analogique en préservant le plus possible d'informations.

COMPRESSION RATIO (TAUX DE COMPRESSION) :

C'est la rapport entre les données compressées et les données non - compressées dans un signal vidéo numérique.

CONVERSION ANALOGIQUE / NUMERIQUE :

La conversion analogique / numérique utilise 3 étapes pour numériser un signal :

- Echantillonnage (cf. ECHANTILLONNAGE) avec un filtre anti repliement qui a pour fonction de filtrer le signal analogique en ne laissant pas passer les fréquences inférieures à la fréquence d'échantillonnage divisée par deux.
- Quantification : signifie approximer la valeur du signal échantillonné par rapport au quantum. A l'origine les signaux étaient quantifiés sur 8 bits ce qui créait des problèmes de luminance ; on a donc étendu le système à 10 bits (1024 combinaisons).
- Codage : signifie écrire la valeur quantifiée en un nombre binaire régie sous certaines règles.

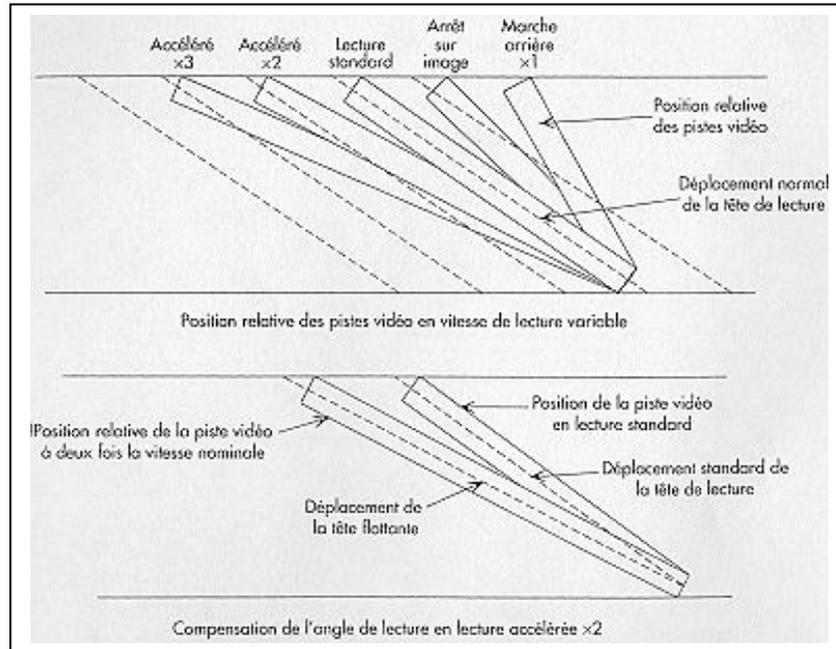
CODEC : algorithme de **codage** et de **décodage** vidéo

DCT (DISCRETE COSINE TRANSFORM) :

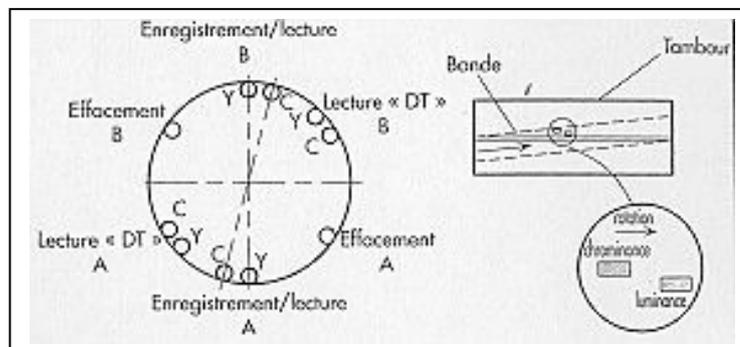
De nombreux encodeurs procèdent à un DCT sur un bloc d'information de huit par huit comme premier pas vers la compression de l'image. Le DCT convertit alors ce bloc qui est un signal de 64 points et le sépare en 64 signaux. Le résultat de l'opération est un groupe de 64 signaux appelés coefficients DCT.

DYNAMIC TRACKING (DT) : LE SUIVI DYNAMIQUE DES PISTES

Les têtes de lecture rotatives montées sur le tambour ne peuvent suivre le tracé des pistes sur la bande que si celle-ci défile à la vitesse de lecture nominale. En cas de lecture en accéléré ou au ralenti, l'angle d'inclinaison relatif des pistes change par rapport à l'angle suivi des têtes qui reste constant. Les têtes ne lisent que des morceaux de pistes et laissent apparaître sur l'image des barres de bruit.



La solution à ce problème est apportée par l'utilisation de têtes flottantes DT «Dynamic Tracking», placées sur des dispositifs piézo-électriques. Selon la tension qui leur est appliquée - tension proportionnelle à la vitesse de lecture variable -, ces derniers déplacent les têtes pour assurer un suivi correct de l'angle d'inclinaison relatif pris par les pistes. Ce déplacement s'effectue à angle droit par rapport à la piste hélicoïdale à suivre. Les têtes DT, qui prennent donc le relais des têtes de lecture traditionnelles en vitesse variable, permettent de lire des bandes en mode accéléré ou ralenti. Par ailleurs, un système de contrôle gère automatiquement les sauts de pistes en fonction de la vitesse de lecture (en vitesse *2, par exemple, seule une piste sur deux est lue).



GENLOCK :

Dispositif incorporé à une source d'images vidéo permettant de la synchroniser avec une autre source pour pouvoir réaliser des montages ou des incrustations.

JPEG (JOINT PHOTOGRAPHIC EXPERTS GROUP) :

Standard utilisant une compression DCT (cf. DCT) pour compresser des images immobiles.

MPEG (MOVING PICTURE EXPERTS GROUP) :

Comité international créé en 1988 par Leonardo Chiariglione pour définir un moyen de coder numériquement de la vidéo et de l'audio sur des disques compacts. Le MPEG définit un décodeur qui reconstruit les flux en images et en samples audio via une règle de base ou un algorithme. Contrairement aux autres moyens de compression vidéo, le MPEG ne regarde pas uniquement à l'intérieur de chaque image, mais cherche aussi les compressions possibles entre les différentes images d'une séquence vidéo.

NORMES 4 : 1 : 1, 4 : 2 : 0, 4 : 2 : 1, 4 : 2 : 2, etc... :

La fréquence d'échantillonnage a été fixée à 13.5 MHz pour le signal de luminance. Les signaux de chrominance sont échantillonnés à une fréquence deux fois plus faible (car l'œil est moins sensible à la chrominance qu'à la luminance) ; cela permet de diminuer le nombre d'échantillons $f_{DB}=f_{DR}=6.75$ MHz.

- 4 : 2 : 2 :

4 → $f_e Y = 4 * 3.375 = 13.5$ MHz, 720 échantillons par ligne.

2 → $f_e DR = 2 * 3.375 = 6.75$ MHz, 360 échantillons par ligne.

2 → $f_e DB = 2 * 3.375 = 6.75$ MHz, 360 échantillons par ligne.

Total : 1440 échantillons par ligne.

Forme : Y - Y + DR + DB - Y - Y + DR + DB - ...

- 4 : 2 : 1 :

$Y = 0.3 R + 0.55 V + 0.11 B$. On sous échantillonne le bleu qu est le moins important.

- 4 : 2 : 0 :

DR et DB sont échantillonné à 6.75 MHz une ligne sur deux.

Afin de réduire le débit d'information à diffuser et à enregistrer, il existe deux formats sous échantillonnés : le 4 : 2 : 0 et le 4 : 1 : 1. Ces formats sont notamment utilisés par les magnétoscopes DV et comportent moins d'information de chrominance. Ils correspondent aux choix historiques du NTSC (4 : 1 : 1) et du PAL - SECAM (4 : 2 : 0).

4:2:2	4:2:0	4:1:1
Fréquence d'échantillonnage : Luminance : 13,5 MHz Chrominance : 6,75 MHz	Fréquence d'échantillonnage : Luminance : 13,5 MHz Chrominance : 6,75 MHz une ligne sur 2	Fréquence d'échantillonnage : Luminance : 13,5 MHz Chrominance : 3,375 MHz

■ Echantillon de luminance et de chrominance
□ Echantillon de luminance

Fréquences d'échantillonnage et répartition des pixels pour les niveaux les plus courants

NUMERISATION:

La numérisation ouvre l'image à l'univers informatique qui permet grâce à des logiciels puissants et perfectionnés d'appliquer aux images et aux sons des traitements de plus en plus puissants. Le seul inconvénient du numérique est la quantité d'informations pour obtenir une image satisfaisante (270Mbits/s pour une image en 4:2:2 sur 10 bits). D'où la nécessité de compresser les signaux numériques (JPEG , MPEG, ...).

SDI (SERIAL DIGITAL INTERFACE) :

Sous la forme parallèle, avec 11 paires blindées pour un seul signal vidéo, il n'est pas possible d'envisager le câblage d'un studio en 4:2:2 d'une manière rationnelle. Il a donc été décidé de normaliser le signal numérique sous une forme série. Il peut alors être transporté dans un simple câble coaxial de 75 ohms au débit de 270 Mbits/s et sur une distance de 300 m. L'interface série est compatible avec les signaux codés sur 8 ou 10 bits. Les données parallèles sont stockées dans un registre à décalage. Les mots, codés sur 10 bits, sont ensuite relus à 270 Mbits/s. Le signal obtenu est embrouillé en utilisant un générateur pseudo - aléatoire qui va disperser les transitions pour assurer une bonne transmission de l'horloge.

Il existe une version 16/9 de l'interface SDI avec 960 pixels par ligne, pour conserver une résolution identique au 4/3. Cette version avec un débit de 360 Mbits/s peut être enregistrée sans compression, en 8 bits, par le magnétoscope D5.

SDTI (SERIAL DIGITAL TRANSPORT INTERFACE) :

C'est une évolution de la norme SDI, permettant de transporter des données numériques compressées, sur un support SDI et avec un débit de 270 Mbits/s à 4 fois la vitesse nominale. Les données sont mises en forme de manière à assurer une compatibilité descendante avec les grilles de commutations SDI. Les données transportées peuvent représenter un ou plusieurs flux vidéo numérique compressé, transmis en temps réel ou en accéléré.

SHUTTER :

Fonction permettant de modifier la vitesse d'obturation ou le réglage de fréquence « clear scan ».

SMEAR :

Pollution des registres verticaux des CCD dans le cas où l'on cadre une lumière intense.